# Recent Application in Biometrics

*Edited by Jucheng Yang and Norman Poh*

# STATE OF THE ART IN BIOMETRICS

Edited by **Jucheng Yang** and **Loris Nanni**

**Notice**

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

# We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

## 4,000+
Open access books available

## 116,000+
International authors and editors

## 120M+
Downloads

## 151
Countries delivered to

Our authors are among the
## Top 1%
most cited scientists

## 12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# Meet the editors

Dr. Jucheng Yang is a professor at Jiangxi University of Finance and Economics, China. He obtained his B.S. degree from South-Central University for Nationalities, China, and M.S. and Ph.D. degree from Chonbuk National University, Republic of Korea. He made his post-doc research at the Advanced Graduate Education Center of Jeonbuk for Electronics and Information Technology-BK21 (AGECJEIT-BK21), Republic of Korea. He has published over 40 papers in related international journals and conferences. He has served as one of the editors and reviewers of international journals such as IEEE Transactions on Information Forensics & Security, Neurocomputing, Information Technology Journal and as reviewer of many conferences such as VISAPP'06, ICNC'06-FSKD'06, IPCV'08, IMPRESS'09. He is the registration chair of ICMeCG'09, and the publicity chair of ICMeCG'10, ICMeCG'11. He holds a Chinese patent of biometrics too. His research interests include biometrics, image processing, pattern recognition and neural networks.

Dr. Loris Nanni received his Master Degree from the University of Bologna, Italy in 2002 and he received his Ph.D. in Computer Engineering at DEIS, University of Bologna, Italy in 2006. In December 2010, he wins a tenured position as associate researcher at Department of Information Engineering of the University of Padua, Italy. He has published more than 140 journal and conference papers, such as IEEE Transactions on Neural Networks, IEEE Transactions on Intelligent Transportation Systems, Pattern Recognition and so on. And he served as reviewer in more than 25 international journals. He teaches the course of database and information systems now. His research interests include pattern recognition, bioinformatics, and biometric systems (fingerprint classification and recognition, signature verification, face recognition).

# Contents

# Preface

Biometric recognition is one of the most widely studied problems in computer science. The use of biometrics techniques, such as face, fingerprints, iris, ears, is a solution for obtaining a secure personal identification. However, the "old" biometrics identification techniques are out of date.

The goal of this book is to provide the reader with the most up to date research performed in biometric recognition and to describe some novel methods of biometrics, emphasis on the state of the art skills.

The book consists of 15 chapters, each focusing on a most up to date issue. The chapters are divided into five sections- fingerprint recognition, face recognition, iris recognition, other biometrics and biometrics security. Section 1 collects five chapters on fingerprint recognition. Chapter 1 provides an effective fingerprint quality estimation approach in consideration of feature analysis for fingerprint quality estimation. In Chapter 2 the authors propose a novel hybrid shape and orientation descriptor that is designed for fingerprint matching. Chapter 3 gives a combined software-hardware approach to defeat fingerprint spoofing attack, and two methods are presented based on analyzing different optical properties by using optical coherence tomography (OCT) technology and the spectral analysis. In Chapter 4 the authors describe an optical information processing system for biometric authentication using the optical spatial-frequency correlation (OSC) system for the biometric authentication. Chapter 5 demonstrates that the concept of secondary fingerprint classification is feasible and consistent, and uses it to build an additional component into a fingerprint classification.

In the section 2 of face recognition, Chapter 6 gives a novel illumination normalization method simulating the performance of retina by combining two adaptive nonlinear functions, a difference of Gaussian filter and a truncation. In Chapter 7 the authors present a novel method of handling the variation caused by lip motion during speech by using temporal synchronization and normalization based on lip motion. Section 3 is a group of iris recognition articles, Chapter 8 presents an iris recognition system based on Local Binary Pattern (LBP) features extraction and selection from multiple images, in which stable features are selected to describe the iris identity while the unreliable feature points are labeled in enrolment template. In Chapter 9 a comprehensive

overview of the state-of-the-art in iris biometric cryptosystems is given. After discussing the fundamentals of iris recognition and biometric cryptosystems, existing key concepts are reviewed and implementations of different variations of iris-based fuzzy commitment are presented. Chapter 10 introduces an iris recognition method using the characteristics of orientation.

In the section of other biometrics, Gabor-Based Region covariance matrix (RCM) Features for Ear Recognition is proposed in Chapter 11. In Chapter 12 a fusion method for facial expression and gesture recognition to build a surveillance system by using Particle Swarm Optimization (PSO) and Cascaded SVMs (CSVM) classification is proposed. Chapter 13 examines the role and potential of Kansei and Kansei quality using Kansei engineering case studies, and introduces three case studies to improve Kansei quality in system design. In the last section of biometrics security, Chapter 14 deals with enhancing the efficiency of biometric by integrating it with salt value and encryption algorithms. In Chapter 15 the authors present a novel chaos-based biometrics template protection with secure authentication scheme.

The book was reviewed by editors Dr. Jucheng Yang and Dr. Loris Nanni. We deeply appreciate the efforts of our guest editors: Dr. Girija Chetty, Dr. Norman Poh, Dr. Jianjiang Feng, Dr. Dongsun Park and Dr. Sook Yoon, as well as a number of anonymous reviewers.

**Dr. Jucheng Yang**
Professor
School of Information Technology
Jiangxi University of Finance and Economics
Nanchang, Jiangxi province
China

**Dr. Loris Nanni**
Ph.D in Computer Engineering
Associate researcher
Department of Information Engineering
University of Padua
Italy

# Part 1

## Fingerprint Recognition

# Fingerprint Quality Analysis and Estimation for Fingerprint Matching

Shan Juan Xie[1], JuCheng Yang[2,1], Dong Sun Park[1],
Sook Yoon[3] and Jinwook Shin[4]
*[1]Department of Electronics and Information Engineering,*
*Chonbuk National University, Jeonju,*
*[2]School of Information Technology, Jiangxi University of Finance and Economics,*
*Nanchang,*
*[3]Dept. of Multimedia Engineering, Mokpo National University, Jeonnam ,*
*[4]Jeonbuk Technopark, Policy Planning Division, Jeonbuk,*
*[1,3,4]South Korea*
*[2]China*

## 1. Introduction

Due to their permanence and uniqueness, fingerprints are widely used in the personal identification system. In the era of information technology, fingerprint identification is popular and widely used worldwide, not only for anti-criminal, but also as a key technique to deal with personal affairs and information security. Accurate and reliable fingerprint identification is a challenging task and heavily depends on the quality of the fingerprint images. It is well-known that the fingerprint identification systems are very sensitive to the noise or to the quality degradation, since the algorithms' performance in terms of feature extraction and matching generally relies on the quality of fingerprint images. For many application cases, it is preferable to eliminate low-quality images and to replace them with acceptable higher-quality images to achieve better performance, rather than to attempt to enhance the input images firstly. To prevent these errors, it is important to understand the concepts that frequently influence the images' quality from fingerprint acquisition device and individual artifacts. Several factors determine the quality of a fingerprint image: acquisition device conditions (e.g. dirtiness, sensor and time), individual artifacts (e.g. skin environment, age, skin disease, and pressure), etc. Some of these factors cannot be avoided and some of them vary a long time.

Fingerprint quality is usually defined as a measure of the clarity of ridges and valleys and the "extractability" of the features used for identification such as minutiae, core and delta points, etc (Maltoni, et al. 2003). In good quality images, ridges and valleys flow smoothly in a locally constant direction and about 40 to 100 minutiaes are extracted for matching. Poor-quality images mostly result in spurious and missing minutiae that easily degrade the performance of identification systems.

Therefore, it is very important to estimate the quality and validity of the captured fingerprint image in advance for the fingerprint identification system. The existing

fingerprint estimation algorithms (Chen, et al. 2005; Lim, et al. 2004; Maltoni, et al. 2003;Shen, et al.2001; Tabassi, et al. 2004; Tabassi, et al.2005) can be divided into: i) those that use local features of the image; ii) those that use global features of the image; and iii) those that address the problem of quality assessment as a classification problem. The local feature based methods (Maltoni, et al. 2003; Shen, et al. 2001) usually divide the image into non-overlapped square blocks and extract features from each block. Blocks are then classified into groups of different quality. Methods that rely on global features (Chen, et al. 2005; Lim, et al. 2004) analyze the overall image and compute a global measure of quality based on the features extracted. The method that uses classifiers (Tabassi, et al. 2004; Tabassi, et al.2005) defines the quality measure as a degree of separation between the match and non-match distributions of a given fingerprint. The discrimination performance of quality measures, however, can be significantly different depending on the sensors and noise sources. In this chapter, we propose an effective fingerprint quality estimation approach. Our proposed method is not only based on the basic fingerprint properties, but also on the physical properties of the various sensors.

The chapter is organized as follows: in section 2, we firstly discuss about the factors influencing the fingerprint quality from two aspects: physical characteristics of acquisition devices and artifacts from fingers. And then, we present our proposed effective fingerprint quality estimation approach in consideration of feature analysis for fingerprint quality estimation in section 3. Finally, in section 4, we test and compare a selection of the features with a classifier for quality estimation performance evaluation on the public databases. Conclusion and further work are conducted in section 5.

## 2. Factors influencing the fingerprint quality

In this section, the concepts that frequently influence images' quality from fingerprint acquisition device and individual artifacts are first introduced. The development of fingerprint acquisition devices in common use are reviewed and analyzed with their physical principles of acquiring images, too. Due to different characteristics of capturing devices, the fingerprint quality estimation methods can be specific for each acquisition device. And we also consider various external situations reflecting individual artifacts come from users of devices, such as distortions and noises from the skin condition, the pressure, rotation, etc., which can significantly affect the fingerprint alignment and matching process.

### 2.1 Fingerprint acquisition devices

The most important part of fingerprint authentication is the fingerprint acquisition devices, which are the components where the fingerprint image is formed. The fingerprint quality would influence the matching results since the entire existed matching algorithm has their limitations. The main characteristics of a fingerprint acquisition device depend on the specific sensor mounted which in turn determines the image features (dpi, area, and dynamic range), cost, size and durability. Other feature should be taken into account when a finger scanner has to chosen for a specification use. Two main problems of fingerprint sensing are as follows: (1) Correct readout of fingerprints is impossible in certain cases, such as with shallow grooves. (2) When the skin conditions of the finger are unstable; for example, in case of a skin disorder, the finger pattern changes from readout to readout.

The principle of the fingerprint acquisition process is based on geometric properties, biological characteristics and the physical properties of ridges and valleys (Maltoni, et al.2009). The different characteristics obtained from ridges and valleys are used to reconstruct fingerprint images for different types of capture sensors.

- Geometry characteristics

The fingerprint geometry is characterized by protuberant ridges and sunken valleys. The intersection, connection and separation of ridges can generate a number of geometric patterns in fingerprints.

- Biological characteristics

The fingerprint biological characteristic means the ridge and valley have different conductivity, different dielectric constant of the air, different temperatures, and so on.

- Physical characteristics

Referring to the physical characteristics of the fingerprints, the ridges and valleys exert different pressures on the contact surface, and they have different pairs of wave impedance when they are focused on the horizontal plane.

According to these characteristics, there are two methods for capturing fingerprints. One type of sensors initially sends a detecting signal to the fingerprint, and then it analyzes the feedback signal to form a fingerprint ridge and valley pattern. Optical collection and Radio Frequency (RF) collection are two typical active collection sensors. Other fingerprint sensors are the passive ones. As the finger is placed on the fingerprint device, due to the physical or biological characteristics of the fingerprint ridges and valleys, the different sensors form different signals, and a sensor signal value is then analyzed to form a fingerprint pattern, such as in the thermal sensors, semiconductor capacitors sensors and semiconductor pressure sensors.

Fig.1. shows the development of fingerprint acquisition devices. The oldest "live-scan" readers use frustrated refraction over a glass prism (when the skin touches the glass, the light is not reflected but absorbed). The finger is illuminated from one side with a LED while the other side transmits the image through a lens to a camera. As optical sensors are based on the light reflection properties (Alonso-Fernandez, et al, 2007), which strictly impact the related gray level values, so that the gray level features-based measure quality, so Local Clarity Score ranks first for optical sensors. Optical sensors only scan the surface of the skin and don't penetrate the deep skin layer. In case that there are some spots left over or the trace from the previous acquisition of fingerprints, the resulting fingerprint may become very noisy resulting in difficulty in determining dominant ridges and orientations. This, in turn, makes the orientation certainty level of the fingerprint lower than that of a normal one. Kinetic Sciences and Cecrop/Sannaedle have proposed sweep optical sensors based on this principle. Casio + Alps Electric use a roller with the sensor inside. TST removed the prism by directly reading the fingerprint, so the finger does not touch anything (but still need a guide to get the right optical distance). Thales (formerly Thomson-CSF) also proposed the same, but with the use of a special powder to put on the finger. The BERC lab from Yonsei University (Korea) also developed a touchless sensor (2004). In 2005, TBS launch a touchless sensor with the "Surround Imaging".

A capacitive sensor uses the capacitance, which exists between any two conductive surfaces within some reasonable proximity, to acquire fingerprint images. The capacitance reflects changes in the distance between the surfaces (Overview, 2004). The orientation certainty ranks first for the capacitive sensor since capacitive sensors are sensitive to the gradient changes of ridges and valleys.

Fig. 1. The development of fingerprint acquisition devices, (a) ink (b) optical rolling devices(c) regular camera for fingerprint scan (d) silicon-capacitive scanner (e) optical touch less scanner (f) ultra sound scanner (g) thermal sensor (h) Piezo-electric material for pressure sensor

A thermal sensor is made of some pyro-electric material that generates current based on temperature differentials between ridges and valleys (Maltoni, et al.2003). The temperature differentials produce an image when the contact occurs since the thermal equilibrium is quickly reached and the pixel temperature is stabilized. However, for the sweeping thermal sensor, the equilibrium is broken as the ridges and valleys touch the sensor alternately. Some parts of the fingerprint look coarse and have poor connectivity properties.

Pressure sensor is one of the oldest ideas, because when you put your finger on something, you apply a pressure. Piezo-electric material has existed for years, but unfortunately, the sensitivity is very low. Moreover, when you add a protective coating, the resulting image is

blurred because the relief of the fingerprint is smoothed. These problems have been solved, and now some devices using pressure sensing are available. Several solutions, depending on the material, have been proposed: Conductive membrane on a CMOS silicon chip; conductive membrane on TFT, Micro-electromechanical switches on silicon chip (BMF,2011).

## 2.2 Individual artifact

In the processing of fingerprint acquisition, user's skin structure on the fingertip is captured. Some researches are focused on the possible impacts that skin characteristics such as moisture, oiliness, elasticity and temperature could have on the quality of fingerprint images.

### 2.2.1 Skin structure

For better understand the skin influence of fingerprint quality, we should know basics of our skin structure as in Fig. 2. Skin is a remarkable organ of the body, which is able to perform various vital functions. It can mould to different shapes, stretch and harden, but can also feel a delicate touch, pain, pressure, hot and cold, and is an effective communicator between the outside environment and the brain (Habif, et al.2004) .
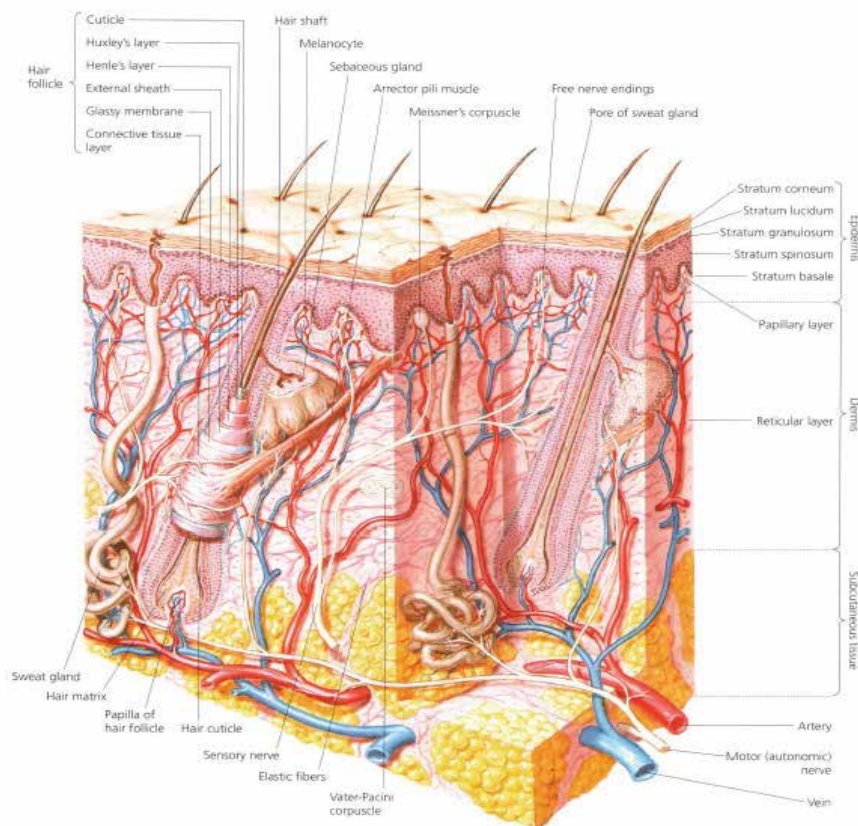


Fig. 2. Skin structure (Habif, et al.2004)

Skin is constantly being regenerated. A skin cell starts its life at the lower layer of the skin (the basal layer of the dermis), which is supplied with blood vessels and nerve endings. The cell migrates upward for about two weeks until it reaches the bottom portion of the epidermis which is the outermost skin layer. The epidermis is not supplied with blood vessels, but has nerve endings. For another 2 weeks, the cell undergoes a series of changes in the epidermis, gradually flattening out and moving toward the surface. Then it dies and is shed (Habif et al. 2004) .

### 2.2.2 Environmental factors and skin conditions

With fingerprint technology becoming a more widely used application, the effects of environmental factors and skin conditions play an integral role in overall image quality, such as air humidity, air temperature, skin moisture, elasticity, pressure and skin temperature, etc. If the finger is dry, the image includes too many light cells which will be marked for operator visual cue. On the other hand, the wet finger or the high pressure image includes more dark cells. The enrolment system will automatically reject the images that are not formed correctly. Fig.3. shows some examples of images representing three different quality conditions. The rows from top to bottom are captured by an optical sensor, capacitive sensor and thermal sensor. In each row, moving from left to right, the quality is bad, medium and good. Different factors affect diverse capture sensors.
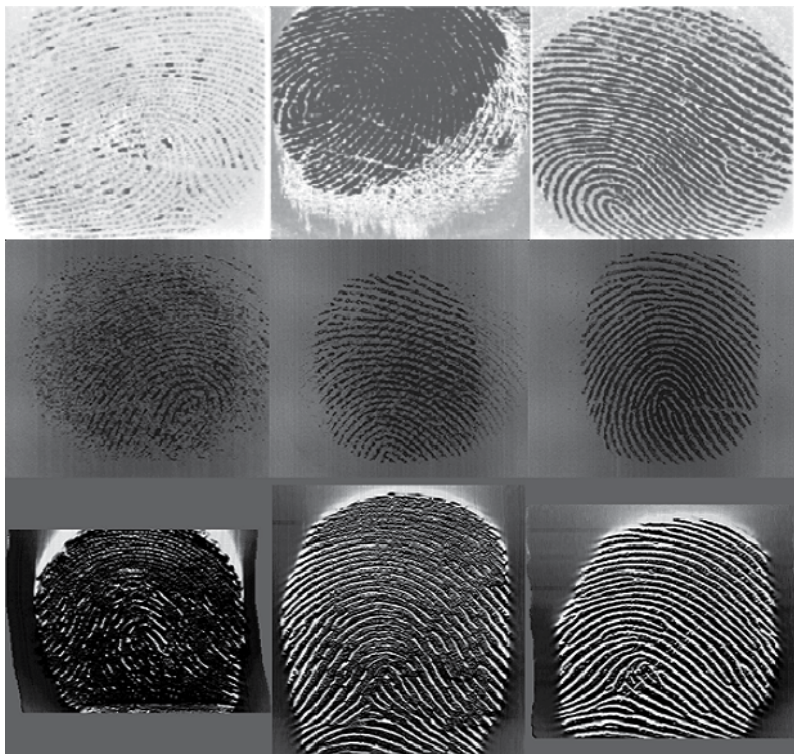


Fig. 3. Fingerprint images from different capture sensors with different environment and skin condition: (a) optical sensor, (b) Capacitive sensor and (c) Thermal sensor. (Xie,et al, 2010b)

Kang et al. (2003) researched 33 habituated cooperative subjects using optical, semi-conductor, tactile and thermal sensors throughout a year in uncontrolled environment. This study evaluates the effects that temperature and moisture have in the success of the fingerprint reader. While evaluating the fingerprints of a variety of subjects, tests determine the role of temperature and moisture in future fingerprints' applications. Each subject uses six fingers (thumb, index, and middle fingers of both hands). For each finger, the fingerprint impression is given at five levels of air temperature, three levels of pressure and skin humidity. The levels of environmental factors and skin conditions used in their experiments are listed in Table 1 (Kang, et al, 2003).

Correlation summary of the performance are conclude, for the optical sensor, it has been observed that the image quality decreases when the temperature goes below zero due to the dryness of the skin. Although all the sensors produce no major image degradation as the temperature changes, they, on the whole, give good quality images above the room temperature. This goes to the same for the air humidity. As far as the pressure is concerned, the image quality is always good with the middle level. For the optical sensor, the foreground image gets smaller for the low pressure while the fingerprint is smeared for the high pressure. The semi-conductor sensor produces good images not only with the middle pressure but also with the high pressure. It is very interesting, however, that the tactile sensor gives better images at the low pressure than at the high pressure. It is also observed that the skin humidity affects to the image quality of all the sensors except the thermal sensor which is a sweeping type. Overall, the quality of fingerprint image is more affected by the human factors such as skin humidity and pressure than the environmental factors such as air temperature and air humidity.

| Factor | | State |
|---|---|---|
| Environment Humidity | | 0~100% |
| Environment Temp(℃) | Below 0 | Winter |
| | 0~10 | Beginning of the spring or end of the fall |
| | 10-20 | Spring or fall |
| | 20-30 | Room Temperature |
| | Above 30 | Summer |
| User Pressure | High | Strongly pressing |
| | Middle | Normally pressing |
| | Low | Softly pressing |
| Skin Humidity | High | 71~100% |
| | Middle | 36~70% |
| | Low | 0~35% |

Table 1. Levels of Environmental factors and skin conditions used in experiments (Kang, et al. 2003)
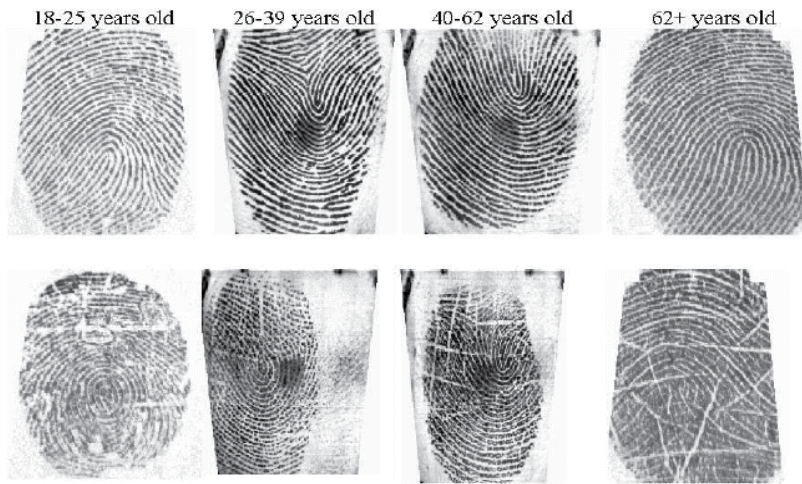
Fig. 4. Samples of high quality fingerprints (top row) and low quality fingerprint (bottom row) with different age ranges (Blomeke, et al, 2008).

### 2.2.3 Age

The Biometrics assurance group stated that it is hard to obtain good quality fingerprints from people over the age of 75 due to the lack of definition in the ridges on the pads of the fingers. Purdue University has made several inquiries into the image quality of fingerprints and fingerprint recognition sensors involving elderly fingerprints. The study compared the fingerprints of an elderly population, age 62 and older, to a young population, age 18-25 on two different recognition devices: optical and capacitive. The results were affected by the age and moisture for both the image captured by the optical sensor, but age only significantly affects the capacitive sensor. Further studies are continued by (Blomeke, et al. 2008) involving the comparison of the index fingers of 190 individual 80 years old of age and older. Fig.4. demonstrates samples of high quality fingerprints (top row) and low quality fingerprint (bottom row) with different age ranges (Blomeke, et al, 2008).

### 2.2.4 Skin diseases

Skin diseases represent a very important, but often neglected factor of the fingerprint acquirement. It is hard to account how many people suffer form skin diseases, but there are many kinds of skin disease (Habif, et al. 2004). When considering whether the fingerprint recognition technology is a perfect solution capable to resolve the security problems, we should take care about these potential skin disease patients with very poor quality fingerprints. The researchers have collected the most common skin diseases, which are psoriais, atopic eczema, verruca vulgaris and pulpitis sicca (Drahansky, et al, 2010).

Fig.5 shows some fingerprint from patients suffering under different skin diseases, either the color of the skin or the ridge lines on the fingertip could be influenced. If only the color of the skin is changed, we can avoid the problem by eliminating the optical sensor. However, the change of skin structure is very significant; the ridge lines are almost damaged. The minutiae are impossible to find for the fingerprint recognition. Even the existed image enhancement methods are helpless to reconstruct the ridge and valley structures, and the image could not be processed further more. The image will be rejected to

the fingerprint acquisition devises and fail for the enrollment since it is really poor quality due to most of the fingerprint quality estimation methodologies. The situation is unfair to the patients; they can not use the fingerprint biometrics system.



(a) Fingerprints with atopic eczema



(b) Fingerprints with psoriasis
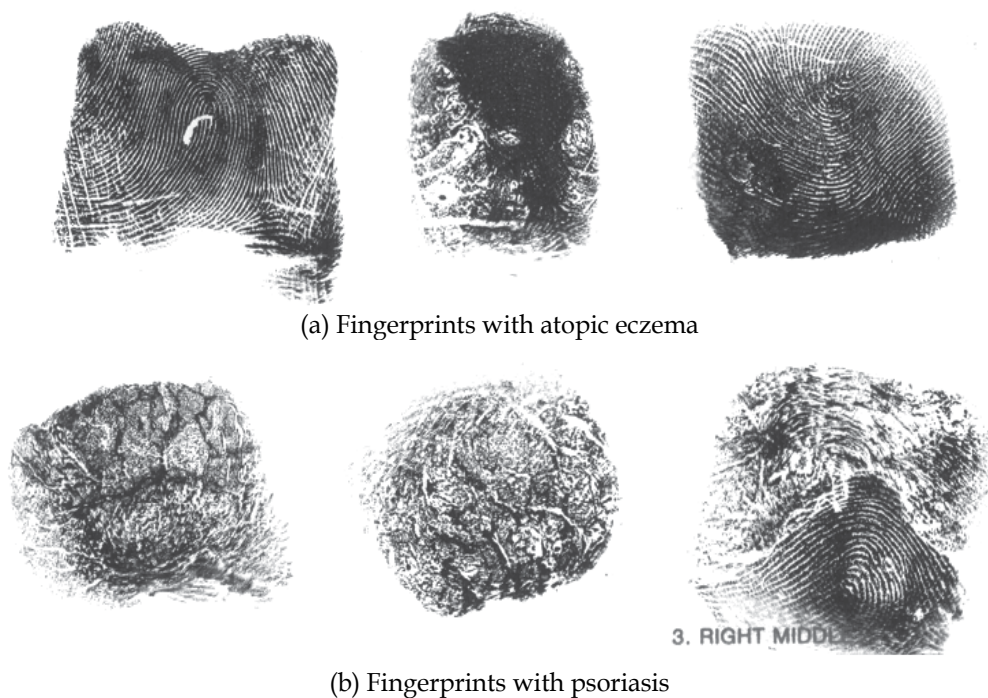
Fig. 5. Fingerprints from patients suffering under different skin diseases

For the temporary skin diseases, the users are able to use their fingers for the fingerprint authentication task after they have healed the diseases. However, for some skin disease, the irrecoverable finger damage may leave, such as the new growth of papillary lines which may cause the users can not to use their fingerprints appropriately. The disease fingerprint will be used for quality assessment, not only based on minutiae, but on finger shape, ridge, correlation, etc .Solutions are expected for the skin disease suffering patients.

## 3. Feature analysis for fingerprint quality estimation

In previous studies (Chen, et al. 2005;Lim, et al. 2004; Maltoni, et al. 2003;Shen, et al.2001; Tabassi, et al. 2004; Tabassi, et al.2005), some fingerprint quality assessments have been performed by measuring features such as ridge strength, ridge continuity, ridge directionality, ridge-valley structure or estimated verification performance. Various types of quality measures have been developed to estimate the quality of fingerprints based on these features. Existing approaches for fingerprint image quality estimation can be divided into: i) based on local features of the image; ii) based on global features of the image; and iii) based on the classifier. The local feature based methods (Maltoni, et al. 2003; Shen, et al.2001) usually divide the image into non-overlapped square blocks and extract features from each block. Methods based on global features (Chen, et al. 2005; Lim, et al. 2004) analyze the overall image and compute a global quality based on the features extracted. The method

that uses classifiers (Tabassi, et al.,2004, Tabassi, et al. 2005) defines the quality measure as a degree of separation between the match and non-match distributions of a given fingerprint.

### 3.1 Quality estimation measures based on local features

The local feature based quality estimation methods usually divide the image into non-overlapped square blocks and extract features from each block. Blocks are then classified into groups of different qualities. A local measure of quality is generated by the percentage of blocks classified with "good" or "bad" quality. Some methods assign a relative important weight to each block based on its distance from the centroid of the fingerprint image, since blocks near the centroid are supposed to provide more reliable and important information (Maltoni, et al. 2003). The local features which can indicate fingerprints quality are researched, such as orientation certainty, ridge frequency, ridge thickness and ridge to valley thickness ratio, local orientation, consistency, etc.

### 3.1.1 Orientation Certainty Level (OCL)

The orientation certainty is introduced to describe how well the orientations over a neighborhood are consistent with the dominant orientation. It measures the energy concentration along the dominant direction of ridges. It is computed as the ratio between the two eigenvalues of the covariance matrix of the gradient vector. To estimate the orientation certainty for local quality analysis, the fingerprint image is participated into non-overlapping blocks with the size of $32 \times 32$ pixels (Lim, et al.,2004; Xie, et al.,2008; Xie, et al.,2009). A second order geometry derivative, named Hessian matrix, is contributed to estimate the orientation certainty. The Hessian matrix that is constructed by H of the gradient vector for an N points image block can be expressed as in Eq. 1.

$$H = \frac{1}{N} \sum_N \left\{ \begin{bmatrix} dx \\ dy \end{bmatrix} \begin{bmatrix} dx & dy \end{bmatrix} \right\} = \begin{bmatrix} h_{xx} & h_{xy} \\ h_{yx} & h_{yy} \end{bmatrix} \tag{1}$$

In this equation, $dx$ and $dy$ are the intensity gradient of each pixel calculated by Sobel operator. Two eigenvectors of H indicate the principal directions and also the directions of pure curvature that are denoted $\lambda_a$ and $\lambda_b$ . $\lambda_a$ is the direction of the greatest curvature and $\lambda_b$ denotes the direction of least curvature.

$$Orientation\_certainty = 1 - \frac{\lambda_b}{\lambda_a} \tag{2}$$

The orientation certainty range is from 0 to 1. For a high certainty block, ridges and valleys are very clear with accordant orientation and, as the value decreases, the orientations change irregularly. When the value is 0, ridges and valleys in the block are changing consistently in the same direction. On the other hand, if the certainty value is 1, the ridges and valleys are not consistent at all. These blocks may belong to a background with no ridges and valleys.

### 3.1.2 Local Orientation Quality (LOQ)

A good quality image displays very clear local orientations. Knowing the curvature of such images with local orientations can be used to determine the core point region and invalid curvatures. Based on local orientations, LOQ is calculated by three steps (Lim, et al. 2004).

**Step 1.** Partition the sub-block.

Partition each sub-block into four quadrants and compute the absolute orientation differences of these four neighboring quadrants in clockwise direction. The absolute orientation difference is lightly greater than zero since the orientation flow in a block is gradually changed.

**Step 2.** Calculate the local orientation quality.

When the absolute orientation change is more than a certain value, in this case, 8-degrees, then the block is assumed as the invalid curvature change block. The local orientation quality of the block is determined by the sum of the four quadrants.

$$O_{mn} = \begin{cases} 0 & |ori(m) - ori(n)| \leq 8^{\circ} \\ 1 & |ori(m) - ori(n)| > 8^{\circ} \end{cases} \tag{3}$$

$$loq_1(i,j) = O_{12} + O_{23} + O_{34} + O_{41} \tag{4}$$

In the equation, ori(m) denotes the orientation value of quadrant m.

**Step 3.** Compute the preliminary local orientation quality.

The LOQ value of an image is then computed as an average change of blocks with MxN blocks in Eq. 5.

$$LOQ_1 = \sum_{i=1}^{M} \sum_{j=1}^{N} loq_1(i,j) \tag{5}$$

### 3.1.3 Ridge frequency

Fingerprint ridge distance is an important intrinsic texture property of fingerprint image and also a basic parameter to determine the fingerprint enhancement task. Ridge frequency and ridge thickness are used to detect abnormal ridges that are too close or too far whereas ridge thickness and ridge-to-valley thickness ratio are used to detect ridges that are unreasonably thick or thin. Fingerprint ridge distance is defined as the distance form a given ridge to adjacent ridges. It can be measured as the distance from the centre of one ridge to the centre of another. Both the pressure and the humidity of finger will influence the ridge distance. The ridge distance of high pressure and wet finger image is narrower than the low pressure and dry finger. Since the ridge frequency is the reciprocal of ridge distance and indicates the number of ridges within a unit length, the typical spectral analysis method is applied to measure the ridge distance in the frequency field. It transforms the representation of fingerprint images from the spatial field to the frequency field and completes the ridge distance estimation in the frequency field (Yin, et.al. 2004).

### 3.1.4 Texture feature

Shen, et al. (2001) proposed the Gabor filter to extract the fingerprint texture information to perform the evaluation (Shen, et al. 2001). Each block is filtered using a Gabor filter with different directions. If a block has good quality (i.e., strong ridge direction), one or several filter responses are larger than the others. In bad quality blocks or background blocks, the filter responses are similar. The standard deviation of the filter responses is then used to determine the quality of each block ("good" and "bad"). A quality index of the whole image

is finally computed as the percentage of foreground blocks marked as "good." Bad quality images are additionally categorized as "smudged" or "dry". If the quality is lower than a predefined threshold, the image is rejected.

## 3.2 Quality estimation measures based on global feature
### 3.2.1 Consistency Measure (CM)

Abrupt direction changes between blocks are accumulated and mapped into a global direction score. The ridge direction changes smoothly across the whole image in case of high quality. By examining the orientation change along each horizontal row and each vertical column of the image blocks, the amount of orientation changes that disobeys the smooth trend is accumulated. It is mapped into global orientation score, which has the highest quality score of 1 and the lowest quality score of 0. This provides an efficient way to investigate whether the fingerprint image posses a valid global orientation structure or not.

The consistency measure (Lim, et al, 2004) is used to represent the overall consistency of an image as a feature. To measure the consistency, an input image is binarized with optimum threshold values obtained from the Otsu's method (Ostu, N.,1979). The consistency in a pixel position is calculated by scanning the binary image with a 3×3 window as in Eq. 6. It provides a higher value if more neighborhood pixels have the same value as that of the center pixel, representing a higher consistency. The final feature for an input image can be averaged as in Eq. 7.

$$con(i,j) = \begin{cases} 0.2 \cdot (9 - sum(i,j)) \cdot (1 - c(i,j)) + c(i,j) & 4 < sum(i,j) \le 9 \\ 0.2 \cdot sum(i,j) \cdot c(i,j) + (1 - c(i,j)) & 0 \le sum(i,j) \le 4 \end{cases} \tag{6}$$

$$Consistency = \frac{\sum_{i=2}^{255} \sum_{j=2}^{255} con(i,j)}{Num} \tag{7}$$

In these equations, *Num = ImageSize/NeighborSize* , *c(i,j)* represents the consistency value of a center pixel and *sum(i,j)* sums the consistency values of the 3×3 window.

## 3.2.2 Power spectrum

Fingerprint power spectrum is analyzed by using the 2-D Discrete Fourier Transform (DFT) (Chen, et al. 2005). For a fingerprint image, the ridge frequency values lie within a certain range. A region of interest (ROI) of the spectrum is defined as an annular region with a radius ranging between the minimum and maximum typical ridge frequency values. As the fingerprint image quality increases, the energy will be more concentrated within the ROI. The fingerprint image with good quality presents strong ring patterns in the power spectrum, while a poor quality fingerprint performs a more diffused power spectrum. The global quality index will be defined in terms of the energy concentration in this ROI. Given a digital image of size M×N, the 2-D Discrete Fourier Transformation evaluated at the spatial frequency ( $\frac{2\pi t}{M}, \frac{2\pi s}{N}$ ) is given by

$$F(t,s) = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i,j) e^{-\iota 2\pi (\frac{ti}{N} + \frac{sj}{M})}, \iota = \sqrt{-1} \tag{8}$$

The global quality index defined in (Chen, et al. 2005) is a measure of the energy concentration in ring-shaped regions of the ROI. For this purpose, a set of band-pass filters is employed to extract the energy in each frequency band. High-quality images will have the energy concentrated in few bands while poor ones will have a more diffused distribution. The energy concentration is measured using the entropy.

### 3.2.3 Uniformity of the frequency field

The uniformity of the frequency field is accomplished by computing the standard deviation of the ridge-to-valley thickness ratio and mapping it into a global score, as large deviation indicates low image quality. The frequency field of the image is estimated at discrete points and arranged to a matrix, and the ridge frequency for each point is the inverse of the number of ridges per unit length along a hypothetical segment centered at the point and orthogonal to the local ridge orientation, which can be counted by the average number of pixels between two consecutive peaks of gray-levels along the direction normal to the local ridge orientation (Maltoni, et al. 2003).

### 3.3 Quality estimation measures based on classifier

Fingerprint image quality is setting as a predictor of matcher performance before a matcher algorithm is applied, which means presenting the matcher with good quality fingerprint images will result in high matcher performance, and vice versa, the matcher will perform poorly for bad quality fingerprints. Tabassi et al. uses the classifiers defines the quality measure as a degree of separation between the match and non match distributions of a given fingerprint. This can be seen as a prediction of the matcher performance. Tabassi et al. (Tabassi, et al.2004, Tabassi, et al.2005) extract the fingerprint minutiae features and then compute the quality of each extracted feature to estimate the quality of the fingerprint image into one of five levels. The similarity score of a genuine comparison corresponding to the subject, and the similarity score of an impostor comparison between subject and impostor are computed. Quality of a biometric sample is then defined as the prediction of a genuine comparison

### 3.4 Proposed quality estimation measures based on selected features and a classifier

Some interesting relationships between capture sensors and quality measure have been found in (Fernandez, et al.2007). Orientation Certainly Level (OCL) and Local Orientation Quality (LOQ) measures that rely on ridge strength or ridge continuity perform best in capacitive sensors, while they are the two worst quality measures for optical sensors. The gray value based measures rank first for optical sensors as they are based on light reflection properties that strictly impact the related gray level values repetitive. From the analysis of various quality measures of optical sensors, capacitive sensors and thermal sensors, Orientation Certainty, Local Orientation Quality and Consistency are selected to be participants in generating the features of the proposed system.

Quality assessment measures can be directly used to classify input fingerprints of a quality estimation system. The discrimination performance of quality measures, however, can be significantly different depending on the sensors and noise sources. Our proposed method is not only based on the basic fingerprint properties, but also on the physical properties of the various sensors. To construct a general estimation system that can be adaptable for various input conditions, we generate a set of features based on the analysis of quality measures.

Fig. 6 shows the overall block diagram of the proposed estimation system. The orientation certainty and local orientation quality measures are the two best measures for capacitive sensors; moreover, in this study, we develop highly improved features from these measures, along with the consistency measure, for images obtained from optical sensors and thermal sensors. The extracted features are then used to classify an input image into three classes, good, middle and poor quality, using the well-known support vector machine (SVM) (Suykens, et al. 2001) as the classifier.
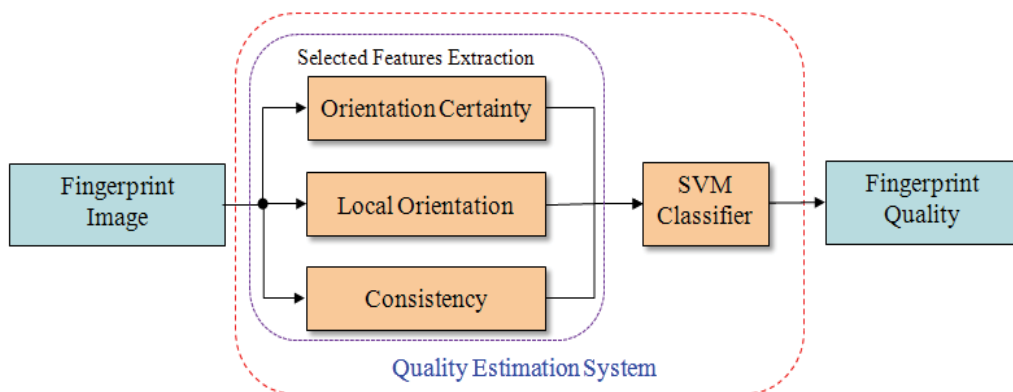


Fig. 6. Selected features and SVM classifier fused fingerprint quality estimation system (Xie et al.2010)

### 3.4.1 Improved orientation certainty level feature

The average of OCL values are used as features for their estimation system. To make the features more accurate, we introduce an optimization named as "Pareto efficient" or "Pareto optimal" (Xie et al. 2008; Xie et al. 2009, Obayashi et al.,2004) to define four classes of blocks and use the normalized number of blocks as a feature for each class. The Pareto optimality is a concept in economics with applications in engineering and social sciences, which uses the marginal rate of substitution to optimize the multi-objectives. To obtain features from OCL values for the proposed system, we classify blocks into four different classes, from good to very bad, as in Table 3, by selecting three optimal thresholds $(x_1, x_2, x_3)$. Three optimal threshold values are assumed to be located in the ranges shown in Eq. 9 and selected by resolving the multi-object optimization. We define the contrast covered by each class limited by optimal thresholds and find three thresholds that maximize the three areas at the same time.

$$
\begin{cases}
D_1 = \int_0^{x_1} \left| gOclNum(\mathrm{x}) - bOclNum(\mathrm{x}) \right| \mathrm{d}x & 0 < x \le x_1 \\
D_2 = \int_{x_1}^{x_2} \left| gOclNum(\mathrm{x}) - bOclNum(\mathrm{x}) \right| \mathrm{d}x & x_1 < x \le x_2 \\
D_3 = \int_{x_2}^{x_3} \left| gOclNum(\mathrm{x}) - bOclNum(\mathrm{x}) \right| \mathrm{d}x & x_2 < x \le x_3 \\
D_4 = \int_{x_3}^{1} \left| gOclNum(\mathrm{x}) - bOclNum(\mathrm{x}) \right| \mathrm{d}x & x_3 < x \le 1
\end{cases}
\tag{9}
$$

In this equation, *gOclNum(x)/bOclNum(x)* represents the number of blocks when the OCL value equals *x* from the good/bad-quality image. *Di* represents the contrast of a level

between good and bad quality. Obviously, if the contrast becomes larger, then it becomes easier to classify with a higher classification rate. As in Table 2, we define four classes of blocks according to their OCL values. Four OCL features of the estimation system are then defined as the normalized amount of blocks for each class. Fig. 7 shows the distribution of four features for the optical sensor in FVC2004 database. We can infer an obvious tendency that good quality images have larger values of OCL feature 1 and smaller values of OCL feature 4, and bad quality images are on the contrary.
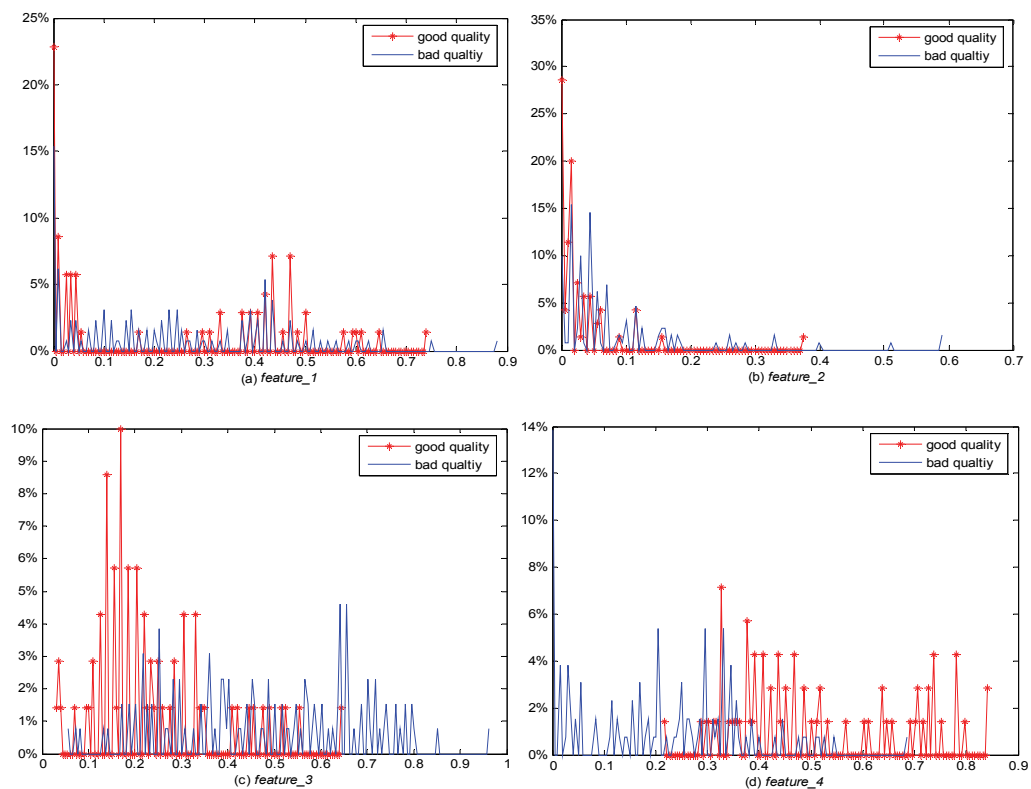


Fig. 7. The distribution of four optical sensor features (Xie,2010)

| Classify grade | Grade |
|---|---|
| 0<OCL≤0.2 | Good quality block |
| 0.2<OCL≤0.6 | Normal quality block |
| 0.6<OCL<1 | Poor quality block |
| OCL=1 | Very poor quality block or background |

Table 2. Four classes of blocks according to their OCL values

### 3.4.2 Improved local orientation quality feature

For the thermal sensor, the equilibrium is broken as the ridges and valleys touch the sensor alternately and affected by the environment temperature, sometimes the fingerprint is

coarse. Different from the poor quality image from optical and capacitive sensors, the poor thermal sensor image still has good orientation, and the ridge and valley still separate clearly. However, the consistency of the poor part obviously performs worse than the good quality one. Due to the residue from previous data acquisition or low pressure against the sensor surface, a bad quality image often carries broken ridges or valley regions; however, in a good quality image, ridges or valley regions are fairly consistent. This preliminary local orientation quality of the fingerprint may include some false positives due to the light reflection properties of optical sensors and the orientation calculation based on gray-level values. To compute the new local orientation quality of the quadrants for supplementing the artifact, we design additional steps as below. Based on the previous LOQ method, we label these block quadrants whose orientation change is more than 8 degrees. Fig.8. shows the basic concept of the improved LOQ feature. We can find the block orientation changes only in two directions: horizontal and vertical. A special label is set for each detected quadrant to avoid repeating detection. The amount of new invalid curvature blocks are set as $loq_2(i, j)$. Then, we can get $LOQ_2$ by the sum of the $loq_2(i, j)$ of unrepeated detection quadrants $(i, j)$.

$$loq_2(i, j) = (O_{25} + O_{16}) \cdot Horizontal + (O_{17} + O_{48}) \cdot Vertical \qquad (10)$$

$$LOQ_1 = \sum_{i=1}^{M} \sum_{j=1}^{N} loq_2(i, j) \qquad (11)$$

If there is orientation change in horizontal, then *Horizontal*=1, otherwise, *Horizontal*=0. *Vertical* is same as *Horizontal*. Therefore, the uniform value of *Improved LOQ* is shown as follows:

$$\text{Im} proved \ \ LOQ = \frac{LOQ_1 + LOQ_2}{4 \times (BlockNum - Num(OCL = 1))} \qquad (12)$$

Where, *BlockNum=Imagesize/Blocksize* and *Num(Ocl=1)* expresses the amount of background blocks among each sub-block partitioned into four quadrants.



Fig. 8. The basic concept of the improved LOQ feature

### 3.4.3 SVM (Support Vector Machine) classifier
The SVM is a powerful classifier with an excellent generalization capability that provides a linear separation in an augmented space by means of different kernels (Suykens, et al, 2001). Each instance in the training set contains one target value (fingerprint quality level or score)

and several attributes (extracted features). The four basic kernels are linear, polynomial, radial basis function (RBF) and sigmoid. The kernels map input data vector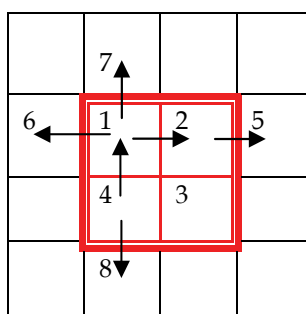s onto a high-dimensional space where a linear separation is more likely, and this process amounts to finding a non-linear frontier in the original input space. In the case, the RBF kernel is employed since it nonlinearly maps samples into a higher dimensional space, so it, unlike the linear kernel, can handle the case when the relation between class labels and attributes is nonlinear (Keerthi. and Lin, 2003). For the proposed quality estimation system, each input vector includes five features as in Eq. 12.

$$V= [OCL1, OCL2, OCL3, Consistency, Improved\ LOQ] \tag{13}$$

*OCL1*, *OCL2* and *OCL3* are three independent features chosen from four features related to the OCL measure, representing the normalized amounts of blocks for each grade. *Consistency* stands for the overall consistency, and *Improved LOQ* is the average LOQ computed from the number of blocks with invalid direction changes.

## 4. Quality estimation performance evaluation

### 4.1 Datasets

Three public Fingerprint Verification Competition (FVC) databases (FVC2000, FVC2002, FVC2004) are employed to evaluate the performance of the proposed quality estimation system. Several sets of fingerprints from various sensors are included. Table 3 shows the sensor information of fingerprint databases. There are 80 images in each Set_B database and 800 images in Set_A database. Since the proposed quality estimation system is based on the local feature, each image is divided into 64 blocks with the size of $32 \times 32$ pixels. Although the types of sensor are adopted in the database, the basis acquisition physical principle is the same for all optical, capacitive and thermal sensors.

|         | Optical sensor | Capacitive sensor | Thermal sensor |
|---------|----------------|-------------------|----------------|
| FVC2000 | DB1_B, DB3_B   | DB2_B             | -              |
| FVC2002 | DB1_A,DB2_A    | DB3_A             | -              |
| FVC2004 | DB1_A,DB2_A    | -                 | DB3_A          |

Table 3. Three public databases employed to evaluate the performance.

### 4.2 Quality benchmark

The NFIS method (Tabassi, E.,2004; Tabassi, E.,2005) is the most widely used method and typical classifier-based methods for fingerprint quality estimation. The method proposed the assumption that fingerprint quality is a predictor of matcher performance. A good quality image will result in a high matcher performance, while a bad quality image will be easily rejected. We relabel the NFIS quality from five levels into three levels, which level 1 is belong to the Good class, level 2-3 is belong to the Medium class and level 4-5 is the Bad class. Fig.9 shows the quality distribution of FVC2002 and FVC2004 by the relabeled NFIS method. As shown in Fig. 9, there are the most Good quality fingerprints in the database FVC2004_DB3 captured by thermal sensors, while the FVC2002_DB3 database captured by the capacitive sensor includes the least Good quality fingerprints. Each fingerprint assigned to a class according to the NFIS quality reclassified to three classes is used to verify the proposed method.
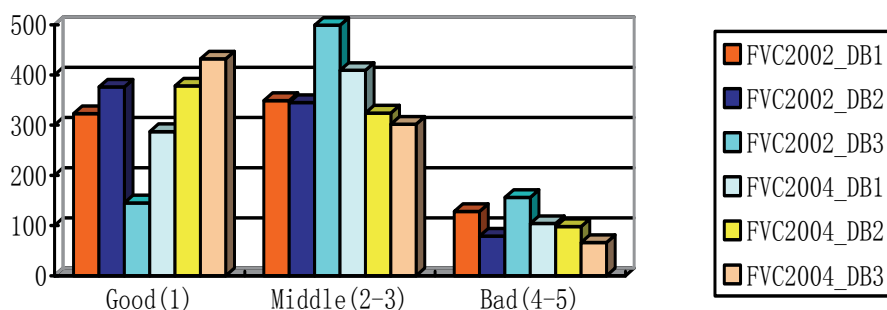
Fig. 9. Quality distribution of the databases regrouped from NFIS with five classes with level1 to level5 into three classes with 'Good', 'Middle', and 'Bad'.

### 4.3 Quality estimation performance

In the evaluation, the 10% Jackknife procedure is employed by using 90% of the images for training and 10% for testing, respectively. Four different kernels, linear, polynomial, RBF and sigmoid, are implemented for the SVM classifier to investigate the performance with different classifier conditions.

Table 4 shows the classification accuracy rate of the original OCL, CM, LOQ measures and their improved versions when they are used separately as a single quality measure. And they are the result of the simulation where the SVM classifier uses the RBF kernel which shows the better result than other kernels. In comparison with the original average OCL measure, the proposed OCL measure achieves better results for adding the optimal determining system which detects not only the local orientation stabilities but also the global ones. Moreover, for the LOQ measure, accuracy is increased after adding the further orientation step.

|           | Original Measures | | | Improved Measures | | |
|-----------|---------|---------|---------|---------|---------|---------|
|           | OCL     | CM      | LOQ     | OCL     | CM      | LOQ     |
| Optical   | 80.05%  | 81.68%  | 77.86%  | 87.50%  | 81.99%  | 81.86%  |
| Capacitive| 83.18%  | 74.62%  | 87.79%  | 90.56%  | 81.61%  | 89.10%  |
| Thermal   | 78.84%  | 77.90%  | 78.72%  | 81.96%  | 89.42%  | 83.04%  |

Table 4. Comparison of the accuracy rate of measures when they are used alone for the quality estimation.

OCL, CM and LOQ feature represent different characteristic of the fingerprint. OCL feature measures the orientation stability of the ridge. CM feature implicates the ridge connection and can detect the small noise, while the LOQ feature performs the irregular direction change of ridges. From Table 5, we can find that the classification performance is improved by combining the local measures. These different measures can make up for each other and get better results. The accuracy rate of the proposed combined measure is 95.62%, 95.50%, 96.25% for the optical, capacitive and thermal sensor, respectively. Comparing with the NFIS method, our proposed method reaches the high accuracy with fewer features. In

addition, the local features fused method reduces much computation complexity than the NFIS method, since it needn't to detect fingerprint minutiae before the quality estimation.

|            | OCL+CM  | CM+LOQ  | LOQ+OCL | LOQ+OCL+ CM |
|------------|---------|---------|---------|-------------|
| Optical    | 92.62%  | 91.25%  | 91.00%  | 95.62%      |
| Capacitive | 93.25%  | 91.88%  | 92.38%  | 95.50%      |
| Thermal    | 94.00%  | 93.95%  | 86.14%  | 96.25%      |

Table 5. Comparison of the accuracy rate of measures according to their combinations when they are used together for the quality estimation.

As residue fingerprints appear frequently in the database from the optical sensors, the problem that residue images are considered as fingerprints with the best quality cannot be ignored. In the database FVC2004 DB1_A and FVC2004 DB2_A, there are about 82 images of obvious residue. We estimate the image quality both by our proposed method and the Classifier-based method. The comparative results are shown in the Table 6. The error rate of our proposed method is 3.65%, while the error rate of Classifier based method is 12.20%. The Classifier based method mistakes the prior image as the minutiae of the remained fingerprint. The proposed system, however, can avoid this kind of residue mistaken error via the global orientation certainty.

|                    |        | Fused method |        |     | Classifier-based method |             |         |
|--------------------|--------|--------------|--------|-----|-------------------------|-------------|---------|
|                    |        | Good | Medium | Bad | Good<br>(1) | Medium<br>(2-3) | Bad<br>(4-5) |
| Subjective Quality | Good   | 44   | 1      | 0   | 43          | 2           | 0       |
|                    | Medium | 2    | 21     | 0   | 4           | 19          | 0       |
|                    | Bad    | 0    | 0      | 14  | 2           | 2           | 10      |

Table 6. Comparison of the proposed fused method to the classifier-based method on the estimation results from residue images

## 5. Conclusions and further work

In this chapter, we analyzed how the fingerprint acquisition device and individual artifacts can influence the fingerprint quality. The acquisition device developers as well as the users require objective and quantitative knowledge to get a high quality image for the fingerprint authentication. The purpose of the study is to propose the process of the image acquisition device performance evaluation under several kinds of sensors and environments. Since fingerprints have different characteristics according to the sensor technologies, the selection of features for fingerprint quality measurements is closely related to the sensors. The reprehensive quality estimation methods are reviewed including the methods based on local features of the image; methods based on global features of the image and methods based on the classifier. In order to perform well for all kinds of sensors, an effective fingerprint quality estimation method for three kinds of sensors optical, capacitive, and thermal sensors is proposed. Three improved features, OCL, CM and LOQ, are commonly used in the fingerprint estimation. The effective of using these features is verified the improvements through the simulation individually.

To improve matching performance, image processing for enhancement is essential. The quality estimation method can used for evaluate the enchantment performance. Some effective enhancement methods are proposed including a three-step using the locally normalized input images, computes the local ridge orientation and then applies a local ridge compensation filter with a rotated window to enhance the ridges by matching the local ridge orientation (Chikkerur et al. 2007; Fronthaler et al.2008; Hong et al.1998; Yang,et al. 2008c; Yang, 2011a; Yang, 2011b). However, there are five major fingerprint matching techniques: minutiae-based, ridge-based, orientation-based, texture-based and 3rd feature based matching techniques (Liu,et al.,2000; Yang,et al.2008a; Yang,et al. 2008b; Yang, 2011b). The major matching algorithms have their own proclivities of fingerprint images, and use them to verify that the presented fingerprint quality estimation approach is effective to support these matching systems appropriately. Different images are expected for the several of fingerprint matching system. Image quality is used to determine whether the captured image is acceptable for further use within the biometric system. Until now the quality estimation only based on the level 1 and level 2 features, in other words, the present quality estimation method only pay attention to the global ridge pattern and the minutiae. However, human examiners perform not only quantitative (Level 2) but also qualitative (Level 3) examination since Level 3 features are also permanent, immutable and unique (Xie,2010a; Zhao et al. ,2008). New quality estimation for the level 3 feature is expected for adopting the Level 3 based matching system.

Moreover future works include evaluation of anti-spoofing capabilities of the fingerprint readers and comparison of fingerprint image qualities with varies age. Also, skin diseases represent a very important, but often neglected factor of the fingerprint acquirement. Problems with biometrics that still lack understanding include recognition of biometric patterns with high accuracy and efficiency, assurance of infeasibility of fraudulence (Jain et al., 2004) and exploration of new features with existing biometrics and novel types of biometrics. A fingerprint recognition algorithm will be required over the fingerprint images of different levels of the quality to produce the matching score.

## 6. Acknowledgement

## 7. References

Alonso-Fernandez, F.; Fabio, R.; Fierrez, J.(2007). Ortega-Garcia, J. Comparison of fingerprint quality measures using an optical and a capacitive sensor. *In Proceedings of Biometrics:Theory, Applications, and Systems, First IEEE International Conference,* BTAS 2007,Washington, DC, USA

Biometrics Assurance Group. (2008). Fingerprints may fail elderly, warn experts. Retrieved October 19, 2008, available from
http://www.kablenet.com/kd.nsf/FrontpageRSS/4FAA013BF5C93DC48025746E0 041E01D!OpenDocument.

Blomeke, C., Modi, S., & Elliott, S. (2008). Investigating the relationship between fingerprint image quality and skin characteristics. *42nd annual 2008 IEEE international carnahan conference on security*, pp. 158-161, Purdue University: Institute of Electrical and Electronics Engineers.

BMF (Biometric Fingerprint Sensors), http://www.bm-f.com/products/overview.html,2011

Chen.Y, Dass.S, and Jain.A (2005), Fingerprint quality indices for predicting authentication performance, *in Proc. AVBPA,* pp:160–170

Chikkerur, S.; Cartwright, A. N. ; Govindaraju, V. (2007). Fingerprint enhancement using STFT analysis, *Pattern Recognition*, 40(1), pp.198-211

Drahansky Martin,  Brezinova Eva,Hejtmankova Dana  ,ORSÁG FILIP.(2010) Fingerprint Recognition Influenced by Skin Diseases. *International Journal of Bio-Science and Bio-Technology,* Daedoek, KR. ISSN 1976-118X, 2010, vol. 3, no.4, pp.11-22.

Fronthaler, H.; Kollreider, K.; Bigun, J. (2008) Local features for enhancement and minutiae extraction in fingerprints, *IEEE Transactions on Image Processing,* 17 (3), pp.354-363

FVC2000 database(2000), available from http://bias.csr.unibo.it/fvc2000/databases.asp

FVC2002 database(2002), available from http://bias.csr.unibo.it/fvc2002/databases.asp

FVC2004 database(2004), available from http://bias.csr.unibo.it/fvc2004/databases.asp

Govindaraju, V., & Shi, Z. (2004). Fingerprint image enhancement based on skin profile approximation.

Habif, T.P.:Clinical Dermatology(2004), 4th Edition, Mosby, China, 2004,p.1004, ISBN 97850532350131952.)

Hong, L.; Wang, Y.; Jain, A. K.(1998). Fingerprint image enhancement: Algorithm and performance evaluation, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 21(4), pp.777–789

Kang,H; Lee, Bongku;Kim, Hakil ; Shin, Daecheol and  Kim, Jaesung(2003); A Study on Performance Evaluation of Fingerprint Sensors, *Lecture Notes in Computer Science*, Volume 2688/2003, 1055

Keerthi, S. S. and C.-J. Lin (2003). Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation* 15 (7), 1667-1689.

Liu. J;Huang. Z; Chan. K (2000),Direct minutiae extraction from gray-level fingerprint image by relationship examination, *International Conference on Image Processing*, vol. 2, pp. 427–430.

Lee, S.H.; Lee, C.H.; Kim, J.H. (2005). Model-based quality estimation of fingerprint images. *Lect. Note.Comput. Sci.*, 3832, pp.229-235.

Lim, E.; Toh, K.A.; Suganthan, P.N.; Jiang, X.D.; Yau, W.Y.(2004). Fingerprint Image Quality Analysis.*In Proceedings of International Conference on Image Processing,* Singapore, 24–27 October,2004; Volume 2, pp. 1241-1244.

Maltoni, D.; Maio, D.; Jain, A.K.; Prabhakar, S.(2009) . Handbook of Fingerprint Recognition, 2nd ed.;Springer: New York, NY, USA, 2009, pp. 59-74.

Mansfield, T. and Wayman, J. L. (2002). Best practices in testing and reporting performance of biometric devices. on the web at www.cesg.gov.uk/biometrics.

Modi, S. K., Elliot, S. J., & Whetsone, J. (2007). Impact of age groups on fingerprint recognition. *5th IEEE Workshop on Automatic Identification Advance Technologies (AutoID2007),* Alghero, Italy. 19-23.

Obayashi, S.; Sasaki, D.; Oyama(2004). A. Finding Tradeoffs by Using Multi-Objective Optimization Algorithms. *Trans. Jpn. Son. Aeronaut. Space Sci.* 2004, 47, 51-58.

Ostu, N.(1979), A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst.*, 9, pp.62-66. Overview of Capacitive Sensors. (2010). Available online:

http://www.lionprecision.com/capacitivesensors/index.html (accessed on 23 March 2010)

Shen.L, Kot.A, and Koo.W (2001). Quality measures of fingerprint images, *in Proc. Audio Video-Based Person Authentication*, pp. 266–271.

Sickler, N. C., & Elliot, S. J. (2004). An evaluation of fingerprint quality across an elderly population vs. 18-25 year olds. *Biometric Consortium 2004 Conference*, Purdue University.

Suykens, J.A.K.(2001). Nonlinear Modeling and Support Vector Machines. *In Proceedings of IEEE Instrumentation and Measurement Technology Conference*, Budapest, Hungary, 2001;pp. 287-294.

Tabassi, E.; Wilson, C.; Watson, C.(2004). Fingerprint Image Quality, NIST research report NISTIR7151,NIST, Gaithersburg, MD, USA.

Tabassi, E.; Wilson, C.(2005). A Novel Approach to Fingerprint Image Quality. *In Proceedings of International Conference on Image Processing*, Genoa, Italy, 11–14,Vol. 2,pp. 37-40.

Wu, J.; Xie, S.J.; Song, D.H.; Lee, W.D.(2008). A New Approach for Classification of Fingerprint Image Quality. *In Proceedings of Cognitive Informatics, 7th IEEE International Conference on Digital Object Identifier, ICCI*, California, USA, 14-16 August 2008; pp. 375-383.

Xie, S.J.; Yang, J.C.; Yoon, S.; Park, D.S.(2008). An Optimal Orientation Certainty Level Approach for Fingerprint Quality Estimation. *In Proceedings of 2nd International Symposium on IITA '08,* Shanghai, China, 2008; Volume 3, pp. 722-726.

Xie, S.J.; Yoon, S.; Yang, J.C.; Park, D.S.(2009). Rule-based Fingerprint Quality Estimation System Using the Optimal Orientation Certainty Level Approach. *In Proceedings of 2nd International Conference on Biomedical Engineering and Informatics*, Tianjin, China,pp.1-5.

Xie, S.J.; Yoo, H.M.; Park, D. S. and Yoon, S.(2010a). Fingerprint reference point detemination based on a novel ridgeline feature , *17th IEEE International Conference on Image Processing (ICIP),* 2010,pp: 3073 - 3076

Xie, S.J.; Yoon, S.; J.W. Shin and D.S. Park.(2010b). Effective Fingerprint Quality Estimation for Diverse Capture Sensors,*Sensors*, 2010, 10(9), 7896-7912

Yang, J.C.; Yoon, S.; Park, D.S. (2006). Applying learning vector quantization neural network for fingerprint matching, *Lecture Notes in Artificial Intelligence (LNAI 4304)* (Springer, Berlin) ,pp.500-509

Yang, J.C.; Park, D. S. (2008a). A fingerprint verification algorithm using tessellated invariant moment features, *Neurocomputing*, 71(10-12),1939-1946

Yang, J.C.; Park, D. S. (2008b). Fingerprint Verification Based on Invariant Moment Features and Nonlinear BPNN, *International Journal of Control, Automation, and Systems,* 6(6), 800-808

Yang, J.C.; Park, D.S.; Hitchcock, R. (2008c). Effective Enhancement of Low-Quality Fingerprints with Local Ridge Compensation, *IEICE Electronics Express*, vol.5, No.23, pp.1002–1009

Yang, J.C. (2011a). A New Approach for Fingerprint Image Enhancement in Frequency Domain, *ICIC Express Letters*, Part B: Applications, vol.2, no.1, pp.171-176

Yang, J.C. (2011b). Non-minutiae based fingerprint descriptor, *book chapter, Biometrics, Intech publisher*, Vienna, Austria, ISBN: 978-953-307-618-8

Yin,Y.L.; Tian, J.; Yang, X.K(2004).Ridge Distance Estimation in Fingerprint Images: Algorithm and Performance Evaluation, *EURASIP Journal on Applied Signal Processing*, No.4, pp:495-502

# Fingerprint Matching using A Hybrid Shape and Orientation Descriptor

Joshua Abraham[1], Paul Kwan[2] and Junbin Gao[3]

[1]*School of Chemistry and Forensic Science, University of Technology Sydney*
[2]*School of Science and Technology, University of New England*
[3]*School of Computing and Mathematics, Charles Sturt University*
*Australia*

## 1. Introduction

Minutiae-based methods have been used in many commercial fingerprint matching systems. Based primarily on a point pattern matching model, these methods rely heavily on the accuracy of minutiae extraction and the detection of landmarks like core and delta for pre-alignment. Spurious and missing minutiae can both introduce errors in minutiae correspondence. Equally problematic is the inability to detect landmarks to guide pre-alignment. Taken together, these problems lead to sub-optimal matching accuracy.

Fortunately, the contextual information provided by ridge flow and orientation in the neighborhood of detected minutiae can help eliminate spurious minutiae while compensating for the absence of genuinely missing minutiae both before and during matching. In addition, coupled with a core detection algorithm that can robustly handle missing or partially available landmarks for pre-alignment, significant improvement in matching accuracy can be expected. In this chapter, we will firstly review fingerprint feature extraction, minutiae representation, and registration, which are important components of fingerprint matching algorithms. Following this, we will detail a relevant fingerprint matching algorithm based on the Shape Context descriptor found in Kwan et al. (2006). Next, we will introduce a novel hybrid shape and orientation descriptor that is designed to address the above problems. The hybrid descriptor can effectively filter out spurious or unnatural minutiae pairings while simultaneously using the additional ridge orientation cues in improving match score calculation. In addition, the proposed method can handle situations where either the cores are not well defined for detection or the fingerprints have only partial overlapping. Lastly, experiments conducted on two publicly available fingerprint databases confirm that the proposed hybrid method outperforms other methods included in our performance comparison.

### 1.1 Fingerprint recognition

An essential component of Automated Fingerprint Recognition Systems (AFRS) is the *matcher* module which makes use of fingerprint matching algorithms in order to match a test fingerprint against template fingerprint(s) for identification/verification (see Figure 1). Currently, reliable fingerprint matching is a non-trivial problem due to environmental noise and uniqueness of each impression. The accuracy of fingerprint matching algorithms depends

on the image quality, image enhancement methods, feature set extraction algorithms, and feature set pre-processing/post-processing algorithms.
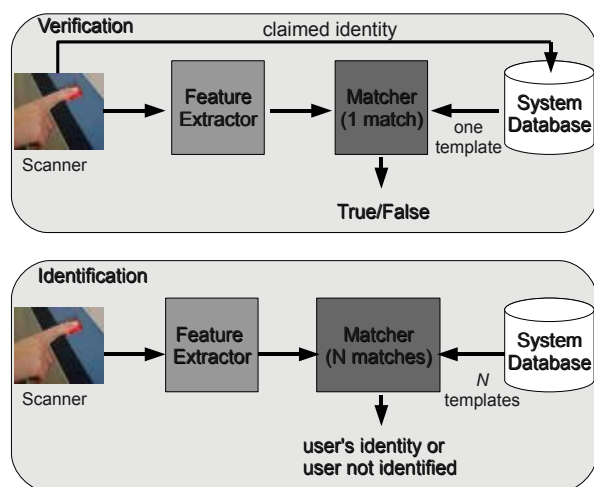


Fig. 1. Basic models for fingerprint verification and identification processes.

Noisy features introduced from environmental factors such as dust, scars, skin dryness, and scaring, are strongly desired to be removed or kept to a minimal level. Even highly robust matching algorithms will suffer from poor matching performance when inaccurate feature extraction and filtering, high noise, poor image quality, or undesirable effects from image enhancement occur.

Even without the advent of environmental noise, applied impressions of the same fingerprint are not guaranteed to be identical due to variability in displacement, rotation, scanned regions, and non-linear distortion or 'warping'. Displacement, rotation, and disjoint detected regions are obviously due to the differences in the physical placement of a finger on a scanner. Figure 2 shows different impressions of the same finger and the noticeable variability in the mentioned areas. One aspect that may be harder to see with the naked eye is non-linear distortion, which is due to both skin elasticity and angular and force variability in applied pressure.

Fingerprint matching algorithm largely follow 3 different classes: *correlation-based*, *minutiae-based*, and *non-minutiae* feature based matching. Correlation-based matching (such as Hatano et al. (2002) and Lindoso et al. (2007)) involves superimposing 2 fingerprint images together and calculating pixel-wise correlation for different displacement and rotations. Minutia-based matching uses extracted minutiae from both fingerprints in order to help perform alignment and retrieve minutiae pairings between both fingerprint minutiae sets. Minutiae-based matching can be viewed as a *point-pattern* matching problem with theoretical roots in pattern recognition and computer vision. Non-minutiae feature based matching (for example Yang & Park (2008) and Nanni & Lumini (2009)) use non-minutiae features, such as ridge shape, orientation and frequency images in order to perform alignment and matching.

Amongst all algorithm classes, minutiae-based methods are the most common due to their strict analogy with the way forensic experts compare fingerprints and legal acceptance as a proof of identity in many countries (Ratha & Bolle, 2003). Minutiae points are also known to be extremely unique from finger to finger in terms of spatial distribution, proving to be ideal features for fingerprint matching. Additionally, minutiae point sets obtain a higher level of

Fig. 2. Eight impressions of the same fingerprint from the FVC2002 database (Maio et al., 2002) with noticeable differences in region overlap, offset, orientation, and image quality.)

uniqueness versus practicality in comparison to other level types of fingerprint features, such as ridge orientation/frequency images and skin pores.

## 2. Base theory

### 2.1 Minutiae extraction
Since the vast majority of fingerprint matching algorithms rely on minutiae matching, minutiae information are regarded as highly significant features for AFRS. The two main methods of minutiae feature extraction either require the gray-scale image to be converted to a binary image, or work directly on a raw or enhanced gray-scale image.

In the binary image based method, the binarization of the gray-scale image is the initial step. This requires each gray-scale pixel intensity value to be transformed to a binary intensity of black (0) or white (1). The simplest approach is to apply a global threshold where each pixel is mapped according to

$$I(x,y) = \begin{cases} 1 & \text{if } I(x,y) \geq t, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Although novel, this method is usually not adequate since fingerprint images may have differing levels of contrast throughout the image. However, the same method can be applied with locally adaptive thresholds. Other more advanced approaches include ridge/valley edge detection techniques using Laplacian operators as in Xiao & Raafat (1991), and mathematical morphology in Gonzalez & Woods (2007).

Once produced, the binary image usually undergoes a morphological thinning operation, where ridge structures are reduced to 1-pixel thickness, referred to as the *skeleton*, in order to aid minutiae detection. The resulting thinned binary image then has each pixel, **p**, analysed in order to find minutiae location. This is achieved by having the 8-neighbourhood (pixels within $3 \times 3$ window centred at **p**) circularly traversed in an anti-clockwise manner in order to produce the Rutovitz *crossing number* introduced in Rutovitz (1966)

$$cn(\mathbf{p}) = \frac{1}{2} \sum_{i=1\ldots8} |val(\mathbf{p}_{(i \mod 8)}) - val(\mathbf{p}_{i-1})| \tag{2}$$

where $val \in \{0,1\}$ (i.e. binary image pixel intensity value). Minutiae pixel locations can now be identified, as ridge endings will have $cn = 1$ and ridge bifurcations will have $cn = 3$.
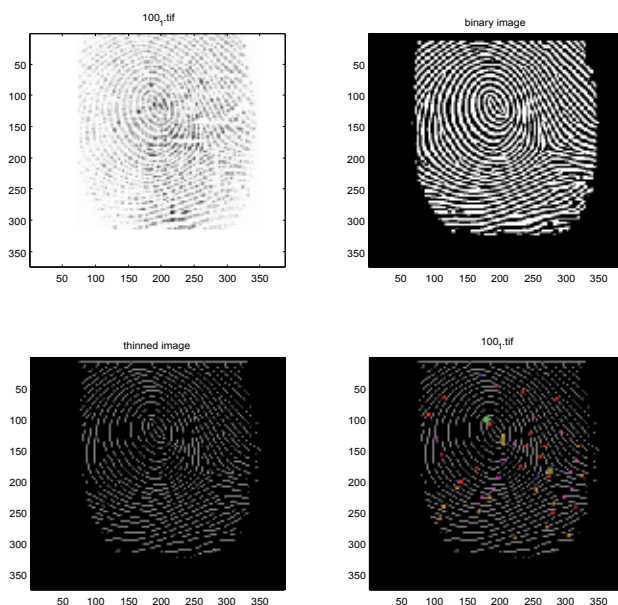


Fig. 3. **top left:** Original gray-scale image. **top right:** binary image **bottom left:** inverted skeleton (thinned) image **bottom right:** inverted skeleton image with core point (green), delta/lower core points (gold), bifurcations (blue for $\theta \in [0° - 180°)$ and purple for $\theta \in [180° - 360°)$), and ridge endings (orange for $\theta \in [0° - 180°)$ and red for $\theta \in [180° - 360°)$).

Although binarization in conjunction with morphological thinning provides a simple framework for minutiae extraction, there are a couple of problematic characteristics. Spurious minutiae (false minutiae) due to thinning algorithms (such as spurs) or irregular ridge endings. Additionally, performance is also an issue since binarization and specifically morphological thinning algorithms are known to be computationally expensive (see Figure 3).

Direct gray-scale minutiae extraction attempts to overcome the problems introduced by image binarization and thinning. One key gray-scale based method that the algorithm in Maio & Maltoni (1997) employs is ridge path following, where an initial point $(x_1, y_1)$ has a $k$ pixel length path projected toward an initial direction, $\theta_1$, and likewise, subsequent iterations have the base point $(x_{t_n}, y_{t_n})$ project the next ridge sample point $(x_{n+1}, y_{n+1})$ in the direction $\theta_n$. Analysis of the *section set $S_n$*, being a 1 dimensional cross section slice centred about $(x_{t_n}, y_{t_n})$ and orthogonal to $\theta_n$ with length $2\sigma + 1$ where $\sigma$ is the average thickness of a ridge, is used to retrieve $\theta_n$, and ultimately, $(x_{n+1}, y_{n+1})$. The path following algorithm terminates when a local maxima cannot be found at the current point's section set, giving clear indication that a ridge ending or bifurcation is reached.

In Farina et al. (1999), these structures and others were removed from the skeleton image. Minutiae were also categorised or ranked according to the degree of their meeting defined topological rules. A similar approach was used in Zhao & Tang (2007), where dot (isolated pixel) filtering, small holes filling (i.e. possibly from dominant pores) were used, in combination with other heuristics. The accuracy of a fingerprint matching algorithm was reported to be decreased by approximately 13.5% when minutiae filtering heuristics were used in comparison to no filtering.

## 2.2 Minutiae representation

Minutiae-based matching algorithms are largely dependent on extracted minutiae information. Robust minutiae-based matching algorithms have to deal with occurrences of missing and spurious minutiae, where missing minutiae can occur as a result of inaccurate feature detection, feature post-processing, or image noise obscuring minutiae detail, and spurious minutiae can be introduced by dry skin, creases, feature detection algorithms, and other potential noise causing agents. The general processes of a fingerprint matching algorithm is presented in Figure 4.
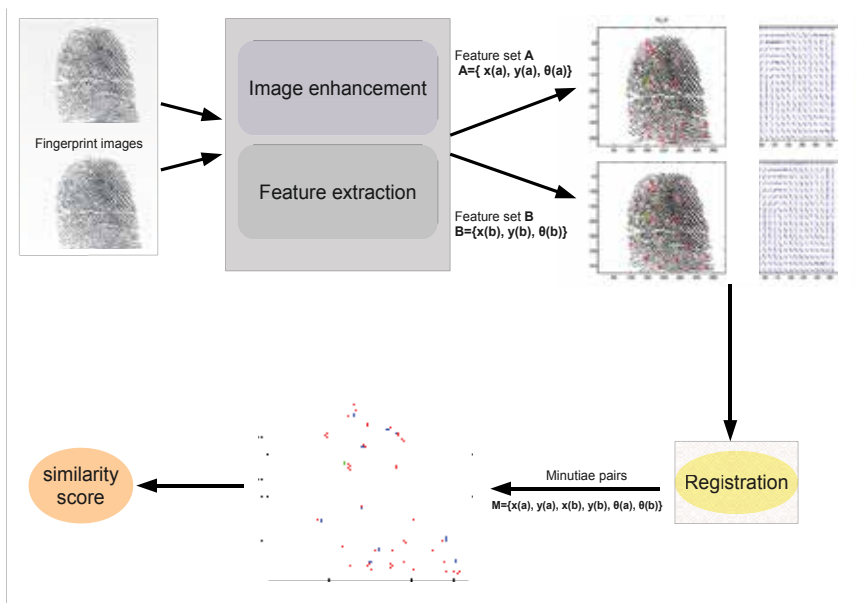


Fig. 4. General processes for minutiae-based fingerprint matching.

In minutiae-based matching, minutiae are commonly represented as minutiae structures called *minutia triplets*, where a minutia, $m_i$, is described as $m_i = \{x, y, \theta\}$ with $x,y$ representing the x-y coordinate of the minutia and $\theta$ the angular direction of the main ridge (see Figure 5 left).

The main focus of minutiae-based matching is to perform a one-to-one mapping or pairing of minutiae points from a test image minutiae set

$$A = \{m_{A_1}, m_{A_2}, \ldots, m_{A_p}\}, \text{where } m_{A_i} = \{x_{A_i}, y_{A_i}, \theta_{A_i}\} \textbf{ and } 1 \leq i \leq p \tag{3}$$

to a template image minutiae set

$$B = \{m_{B_1}, m_{B_2}, \ldots, m_{B_q}\}, \text{where } m_{B_j} = \{x_{B_j}, y_{B_j}, \theta_{B_j}\} \textbf{ and } 1 \leq j \leq q, \qquad (4)$$

forming the *minutiae pairs* $(m_{A_k}, m_{B_{\pi(k)}})$ with $\pi(k)$ as the mapping permutation of pairs from set $A$ to $B$.

Unfortunately, we cannot proceed to find minutiae pairs from triplets without some pre-processing for the following critical reasons:

- fingerprint impressions can differ in orientation, deeming the direction field in the triplet useless,

- fingerprint impressions can differ in offset, deeming the x-y fields in the triplet useless, and

- skin elasticity creates non-linear distortion or 'warping' to occur when different directional pressure is applied causing triplet x-y variations to occur.

In general, the lack of invariant characteristics of the triplet structure prohibits it to aid the process of finding minutiae pairs.

## 2.3 Registration

In order to address the issues concerning the lack of invariance of the triplet structure, *global registration* is required. Global registration concerns the alignment and overlay of the template and test fingerprints so that corresponding regions of the fingerprints have minimal geometric distance to each other. Registration can be achieved geometrically by applying (to either the test or template fingerprint minutiae set) a heuristically guided *affine transform*, where minutiae triplet field values are updated with

$$\begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix} = \begin{bmatrix} \cos(\theta_\Delta) & -\sin(\theta_\Delta) \\ \sin(\theta_\Delta) & \cos(\theta_\Delta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_\Delta \\ y_\Delta \end{bmatrix}, \qquad (5)$$

and

$$\theta_{new} = \theta - \theta_\Delta, \qquad (6)$$

where $\theta_\Delta$ is the orientation difference and $(x_\Delta, y_\Delta)$ is the displacement difference in order to super-impose one fingerprint impression on top of the other with accurate overlap and uniform direction.

Even with the advent of high distortion, minutiae points within a fingerprint image are still expected to keep their general global location in relation to the majority of other minutiae points and other key landmarks (such as cores and deltas) when alignment is achieved. Specifically speaking, the spatial distribution or geometric properties of neighbouring minutiae should have minimal difference even in distorted images. If we consider that there are clear limitations in terms of minutiae landmark relative to positioning variability (even with high distortion), while recognising that different fingerprint impressions have orientation and displacement differences, then the global registration process notably reduces the search space. This reduces algorithm complexity for finding minutiae pairs, since matching pairs are formed in smaller local neighbourhoods (i.e. constraints added for minutiae mappings) once aligned. This allows a naive brute force minutiae pairing process to be avoided.

Following the registration process, we can now produce geometric constraints for the discovery of minutiae matching pairs, including geometric distance:

$$dist_r(m_{A_i}, m_{B_j}) = \sqrt{(x_{A_i} - x_{B_j})^2 + (y_{A_i} - y_{B_j})^2} < r_\delta, \tag{7}$$

or to account for scale difference (i.e. if we are comparing images collected from different resolution scanners)

$$dist_r(m_{A_i}, m_{B_j}) = \sqrt{(x_{A_i} - k_x.x_{B_j})^2 + (y_{A_i} - k_y.y_{B_j})^2} < r_\delta, \tag{8}$$

and minutiae angle difference,

$$dist_\theta(m_{A_i}, m_{B_j}) = \min(|\theta_{A_i} - \theta_{B_j}|, 360° - |\theta_{A_i} - \theta_{B_j}|) < r_\theta. \tag{9}$$

The geometric tolerance $r_\delta$ is in place to account for distortion that may occur, whereas $r_\theta$ is the tolerance for angular differences that may arise due to orientation estimations from the ridge orientation images. Following global registration, a local search can now be performed, in order to match minutiae in the $\delta$-neighbourhood that meet the constraints in equations 7-9 (see Figure 5 right).

Once genuine minutiae pairs are produced, a metric of similarity, usually called the *similarity score*, can then be calculated. The similarity score must accurately describe how similar two fingerprints are, taking into account all of the relevant information obtained from earlier stages, such as number of genuine minutiae pairs and how similar each pair is. One similarity score given in Liang & Asano (2006) is defined as

$$sim(A, B) = \frac{n^2_{match}}{n_A n_B} \tag{10}$$

where $n_{match}$ is the number of matching minutiae pairs, and $n_A, n_B$ are the number of minutiae in the overlapped regions of the template and test fingerprints following registration.
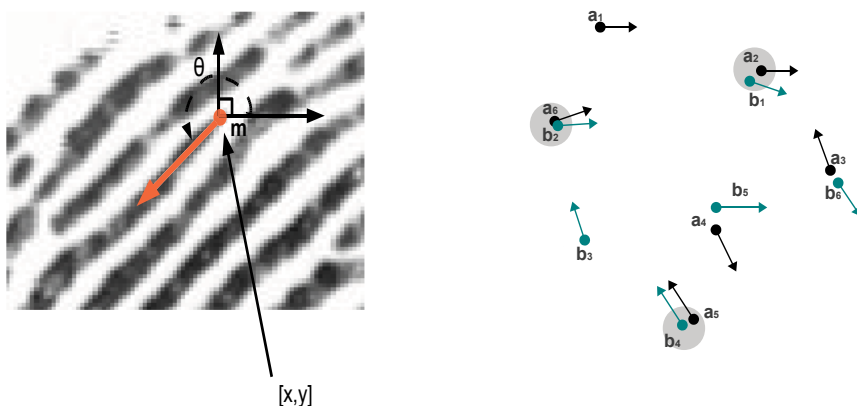


Fig. 5. **left:** minutia triplet structure representation. **right:** Minutiae points from 2 different fingerprints being mapped after registration, with gray circles representing pairs with constraints upheld (equations 7-9).

In order to effectively match fingerprints, we require that the registration used not only be computationally sound, but also perform accurate alignment. In order to achieve this, some methods use additional features (sometimes in combination with minutiae detail) for global alignment, such as cores and deltas, local or global orientation field / texture analysis, and ridge feature analysis.

Using core points for registration is known to dramatically improve the performance of a matching algorithm. In Chikkerur & Ratha (2005), a graph theory based minutiae matching algorithm reported a 43% improvement in efficiency when including the core point for registration, without adverse effect toward matching accuracy. In Zhang & Wang (2002) core points were used as key landmarks for registration. This method proved to be extremely efficient in comparison to other key registration methods. Structural features of minutiae close to the core are used to calculate the rotation needed. The core point was also used in Tian et al. (2007) for registration with the orientation that produced the minutiae pair with the minimum hilbert scanning distance. These and similar methods heavily rely on the core point for alignment. Such a dependence is not strictly robust, since not all fingerprint impressions contain core points and the inclusion of noises may effect the accuracy of core detection algorithms, possibly resulting in incorrect alignments.

In Yager & Amin (2005), the global orientation image with points divided into hexadecimal cells (see Figure 6 right) is used for registration. The steepest descent algorithm was used in order to find the affine transform $(x_\Delta, y_\Delta, \theta_\Delta)$ that minimise the cost function

$$C(P, Q') = \frac{1}{N} \sum_{p \in P, q' \in Q'} \min[(p - q'), (q' - p + \pi)], \tag{11}$$

where P is orientation image of one fingerprint and Q' is the orientation image of the second fingerprint following an affine transformation.
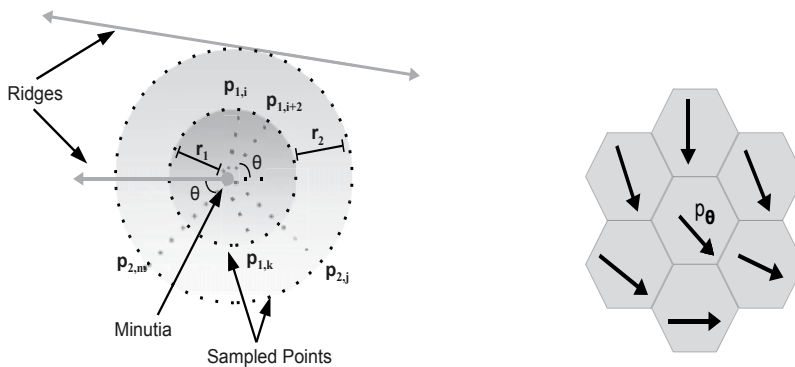


Fig. 6. **left:** The local orientation descriptor used in Tico & Kuosmanen (2003). **right:** Hexagonal orientation cells within the orientation image in Yager & Amin (2005) .

Another example which uses the global orientation image for registration can be found in Liu et al. (2006). For all possible transforms of the test fingerprint onto the template fingerprint

which has significant region overlap, the *normalised mutual information* (NMI) defined as

$$NMI(X,Y) = \frac{H(X) + H(Y)}{H(X,Y)} \tag{12}$$

is calculated, where

$$H(X) = -E_X[\log P(X)], \tag{13}$$
$$H(Y) = -E_Y[\log P(Y)], \tag{14}$$

and

$$H(X,Y) = -E_X[E_Y[\log P(X,Y)]], \tag{15}$$

where X and Y are discrete random variables representing the orientation fields, $O_x$ and $O_y$, of the template and test fingerprints, respectively, which are divided into $b$ blocks. The probabilities can be calculated as

$$P_{XY}(x,y) = \frac{n(x,y)}{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} n(i,j)}, \tag{16}$$

$$P_X(x) = \sum_{j=0}^{b-1} P(x,j), \tag{17}$$

$$P_Y(y) = \sum_{i=0}^{b-1} P(i,y), \tag{18}$$

and

$$n(x,y) = \begin{cases} 1 & \text{if } |O_x(x) - O_x(y)| \leq \lambda, \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

with $\lambda$ as a small threshold, indicating that orientation corresponding image blocks have very similar orientations. We can now find the transform which produces the maximum NMI as the global registration.

Global landmarks and features are not only used for aiding registration. Local structure sets or *descriptors* can also be used for registration. For instance, in Tico & Kuosmanen (2003), the rotation and translation invariant minutia orientation descriptor (see Figure 6 Left) is used to find minutiae pair with the maximum probabilistic value

$$[r,s] = \arg\max_{i,j} P(m_{A_i}, m_{B_j}) \tag{20}$$

with

$$P(m_{A_i}, m_{B_j}) = \frac{S(m_{A_i}, m_{B_j})^2}{\left( \sum_{k=1}^{p} S(m_{A_k}, m_{B_j}) + \sum_{l=1}^{q} S(m_{A_i}, m_{B_l}) \right)} \tag{21}$$

and $S(m_{A_i}, m_{B_j})$ is the similarity function defined as

$$S(m_{A_i}, m_{B_j}) = (1/K) \sum_{c}^{L} \sum_{d}^{K_c} \exp\left( -\frac{2\left( \min(|\theta_{c,d}^{A_i} - \theta_{c,d}^{B_j}|, \pi - |\theta_{c,d}^{A_i} - \theta_{c,d}^{B_j}|) \right)}{\pi\mu} \right) \tag{22}$$

where the orientation descriptor has a total of $K$ sample points distributed as $L$ concentric circles having $K_c$ points (i.e. possibly differing number per circle) with equidistant angular

distribution (i.e. $\frac{2\pi}{K_c}$ step size), $\theta_{c,d}^{A_i}$ is minimum angle required to rotate the $d^{th}$ sample orientation on $c^{th}$ circle to the orientation of minutia $m_{A_i}$ (likewise for $\theta_{c,d}^{B_j}$), and $\mu$ is an empirically chosen parameter. After finding the maximum pair index, $[r,s]$, the affine transform is performed on the set $B = \{m_{B_1}, m_{B_2}, \ldots, m_{B_q}\}$, with $\theta_\Delta = \theta_{A_r} - \theta_{B_s}$ and $[x_\Delta \quad y_\Delta]^T = [x_{A_r} - x_{B_s} \quad y_{A_r} - y_{B_s}]^T$ as the transformation parameters. The additional local texture information contained in the orientation-based descriptor is then used in the similarity score to give

$$sim(A,B) = \frac{\left(\sum_{(i,j)\in C} S(m_{A_i}, m_{B_j})\right)^2}{n_A n_B} \tag{23}$$

where $S(m_{A_i}, m_{B_i})$ is the function defined in equation 22, $C$ is the set of minutiae pairs, and $A_i, B_j$ are the template/test minutiae list indexes, respectively.

Unlike most algorithms that have global registration preceding local registration or minutiae pairing, the proposed method in Bazen & Gerez (2003) finds a list of minutiae pairs prior to performing global registration. Each minutia in the template and test fingerprints have an extended triplet structure defined as a 2-neighbourhood structure in the form of $\{x, y, \theta, r_1, \theta_1, r_2, \theta_2\}$, where $r_1$ and $\theta_1$ are the polar co-ordinates of the closest minutia, and likewise for the second closest minutia, $r_2$ and $\theta_2$. The list is then built by finding pairs from after aligning each minutiae structure and then comparing the similarity. This initial minutiae list may contain false pairs. Using the largest group of pairs that use approximately the same transform parameters for alignment, a least squares approach is then used to find the optimal registration. To aid highly distorted fingerprints the non-affine transformation model based on the Thin Plate Spline (T.P.S) (defined in section 3.1.1) is applied to model distortion, with minutiae pair correspondences as anchor points. Such a model allows the minutiae pair restrictions of equations 7-9 to be more rigorously set, helping reduce an algorithms FAR (False Accentance Rate).

There exist algorithms that bypass global registration all together. In Chikkerur & Govindaraju (2006), a proposed local neighbourhood minutia structure called *K-plet* uses a graph theory based consolidation process in combination with dynamic programming for local matching (i.e. minutiae pairing). Another example of a matching algorithm that does not require registration can be found in Kisel et al. (2008), which opts to use translation invariant minutia structures with neighbourhood information for finding genuine minutiae correspondences.

For the majority of algorithms that use global registration, local minutiae matching is then performed. In order to aid local matching, structures based on triplets and other *shape descriptors*, which are shape descriptive data sets employed for the geometric analysis of shapes (that may have been previously utilised in the registration process), can be used to measure minutiae similarity. For instance, a greedy algorithm is used in Tico & Kuosmanen (2003), where subsequent pairs are selected in order of descending probability values (i.e. equation 21) in conjunction with equations 7-9. A similar methodology can be found in Qi et al. (2004), where a greedy algorithm and textural minutia-based descriptor is similarly used.

## 3. Hybrid shape and orientation descriptor

In this section, a brief theoretical foundation concerning the Thin Plate Spline (T.P.S) and shape context will initially be established. Following this, a fingerprint matching algorithm using

the enhanced shape context descriptor introduced in Kwan et al. (2006) is reviewed. Next, the proposed hybrid descriptor method which uses the enhanced shape context descriptor in conjunction with the orientation-based descriptor described in Tico & Kuosmanen (2003) will be introduced. Experimental results will be reported in Section 3.2.3.

### 3.1 Enhanced shape context

Shape matching algorithms that use contour based descriptors based on point samples (or point pattern matching algorithms) are analogous to minutiae-based fingerprint matching algorithms, as they usually combine the use of descriptors with dynamic programming, greedy, simulated annealing, and energy minimization based algorithms, in order to register shapes and compute a similarity measure. Hence, like fingerprint matching algorithms, desirable characteristics of shape matching methods are invariance to rotation and scale, while achieving robustness toward small amounts of distortion and outlier point samples.

The Shape Context descriptor in Belongie et al. (2000) and Belongie et al. (2002) is a robust contour based shape descriptor used for calculating shape similarity and the recovering of point correspondences. Recently, Kwan et al. (2006) proposed a fingerprint matching based variant of the shape context, the *Enhanced Shape Context*, utilising additional contextual information from minutiae sets.

Initially, we are given $n$ and $m$ minutiae from test and template fingerprints, $P$ and $Q$, respectively. For each minutia, $p_i \in P$, we are to find the best matching minutia, $q_j \in Q$. When the shape context descriptor is constructed for a particular minutia, a coarse histogram

$$h_{p_i}(k) = \# \left\{ p_j \neq p_i \ : \ (p_j - p_i) \in bin(k) \right\}. \tag{24}$$

involving the remaining $n-1$ minutiae of $P$ is built as the *shape context* of minutia $p_i$. Each bin corresponds to the tally of minutiae in a particular spatial region with distance $r_l \leq d \leq r_h$ and direction $\theta$.
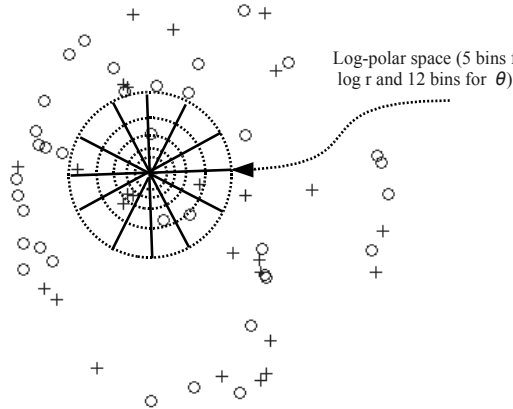


Fig. 7. Log-polar histogram bins used to create shape context histogram for a minutia point. Bifurcations and ridge endings are denoted by '+' and 'o', respectively.

The spatial geometric regions are divided to be uniform in log-polar space, where the log-polar transformation is defined as the mapping from the Cartesian plane (x,y) to the log-polar plane $(\xi, \eta)$ with

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} \log r \\ \theta \end{bmatrix} = \begin{bmatrix} \log \sqrt{x^2 + y^2} \\ \arctan \frac{y}{x} \end{bmatrix}. \tag{25}$$

The shape context descriptor is then constructed for each minutiae $p_i \in P$, and likewise, each $q_j \in Q$, providing a localised spatial survey of the minutiae distributions for each fingerprint. We can now consider the cost of matching two minutiae, which we can later use to find the optimal mapping of minutiae. Let $C_{ij} = C(p_i, q_j)$ denote the cost of matching minutia $p_i \in P$ with $q_j \in Q$.

Since the shape context are distributed as histograms, we can use a modification of the $\chi^2$ statistic:

$$C_{ij} \equiv C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{\left[ h_{p_i}(k) - h_{q_j}(k) \right]^2}{h_{p_i}(k) + h_{q_j}(k)} \tag{26}$$

where $h_{p_i}$ and $h_{q_j}$ denote the $K$-bin histograms of points $p_i$ and $q_j$, respectively. This cost can be modified to include application specific weighting and additional costs, in order to add extra relevant information, and hence, improve accuracy.

In order to improve overall accuracy, the enhanced shape context cost value was modified to include contextual information, such as minutia type (i.e., bifurcation and ridge endings, as shown in Figure 7) and minutia angle. This produced the modified log-polar histogram cost as

$$C_{ij}^* \equiv C^*(p_i, q_j) = \left( 1 - \gamma C_{ij}^{type} C_{ij}^{angle} \right) \cdot \left( \frac{1}{2} \sum_{k=1}^{K} \frac{\left[ h_{p_i}(k) - h_{q_j}(k) \right]^2}{h_{p_i}(k) + h_{q_j}(k)} \right) \tag{27}$$

with $0 \leq \gamma \leq 1$,

$$C_{ij}^{type} = \begin{cases} -1 & \text{if } type(p_i) = type(q_j), \\ 0 & \text{if } type(p_i) \neq type(q_j) \end{cases} \tag{28}$$

and

$$C_{ij}^{angle} = -\frac{1}{2} \left( 1 + \cos((\angle_{initial-warped})) \right) \tag{29}$$

where $\angle_{initial-warped}$ is the absolute value of the angle difference in the ridge orientation tangent at $p_i$ and $q_j$ in the beginning and after each iterative warping (see section 3.1.1). If $\angle_{initial-warped}$ is greater than $\pi$, it is adjusted as $2\pi - \angle_{initial-warped}$ so it will less than or equal to $\pi$.

After computing the cost $C_{ij}^*$ for all possible $n \times m$ pairs, the mapping permutation (one-to-one), $\pi$, that minimises the total matching cost

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)}) \tag{30}$$

which can be computed via the *Hungarian* algorithm as in Jonker & Volgenant (1987). To conform to a one-to-one mapping, $|n - m|$ dummy points can be added to the smaller fingerprint minutiae set. Minutiae that are mapped to these dummy points can be considered to be outliers. For more robust handling, dummy point mappings can be extended to minutiae that have a minimum cost greater than a desired threshold $\epsilon_d$.

### 3.1.1 Registration using the Thin Plate Spline

After finding the minutiae correspondences, the Thin Plate Spline (T.P.S) can be used to register the point correspondences together, accounting for a rigid global transformation and local non-linear transformation.

T.P.S is a mathematical model based on algebraically expressing the physical bending energy of a thin metal plate on point constraints. T.P.S is both a simple and sufficient model for non-rigid surface registration with notable applications in medical imaging. T.P.S was first introduced in Bookstein (1989) for the accurate modelling of surfaces that undergo *natural warping*, where no significant folds or twists occur (i.e., where a diffeomorphism exists).

Two sets of landmark points (i.e. minutiae) from two $\mathbb{R}^2$ surfaces are paired in order to provide an interpolation map on $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. T.P.S decomposes the interpolation into a linear component with an affine transformation for a global coarse registration and a non-linear component with smaller non-affine transformations. In other words, the linear component or affine transform can be considered as the transformation that expresses the global geometric dependence of the point sets, whereas the non-affine transform component identifies individual transform components in order to fine tune the interpolation of the point sets. In addition, the affine transform component allows T.P.S to be invariant under both rotation and scale.

In the general two dimensional T.P.S case, we have $n$ control points

$$\{\mathbf{p}_1 = (x_1, y_1), \mathbf{p}_2 = (x_2, y_2), \dots, \mathbf{p}_n = (x_n, y_n)\} \tag{31}$$

from an input $\mathbb{R}^2$ image and target control points

$$\{\mathbf{p}'_1 = (x'_1, y'_1), \mathbf{p}'_2 = (x'_2, y'_2), \dots, \mathbf{p}'_n = (x'_n, y'_n)\} \tag{32}$$

from a target $\mathbb{R}^2$ image. To set up the required algebra of the general T.P.S case, we define the following matrices

$$\mathbf{K} = \begin{bmatrix} 0 & U(r_{12}) & \dots & U(r_{1n}) \\ U(r_{21}) & 0 & \dots & U(r_{2n}) \\ \dots & \dots & \dots & \dots \\ U(r_{n1}) & U(r_{n2}) & \dots & 0 \end{bmatrix}, n \times n; \tag{33}$$

where $U(r) = r^2 \log r^2$ with $r$ as the Euclidean distance, $r_{ij} = \|p_i - p_j\|$,

$$\mathbf{P} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_n \end{bmatrix}, 3 \times n; \tag{34}$$

$$\mathbf{V} = \begin{bmatrix} x'_1 & x'_2 & \dots & x'_n \\ y'_1 & y'_2 & \dots & y'_n \end{bmatrix}, 2 \times n; \tag{35}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{V} & \mathbf{0}_{2\times3} \end{bmatrix}^T, (n+3) \times 2; \tag{36}$$

and

$$\mathbf{L} = \begin{bmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P^T} & \mathbf{0}_{3\times3} \end{bmatrix}, (n+3) \times (n+3); \tag{37}$$

We can now find the vector $W = (w_1, w_2, \dots, w_n)$ and the coefficients $a_1$, $a_x$, $a_y$, by the equation

$$\mathbf{L}^{-1}\mathbf{Y} = (\mathbf{W}|\ a_1\ a_x\ a_y)^T \tag{38}$$

which can then have its elements used to define the T.P.S interpolation function

$$f(x,y) = \left[f_x(x,y), f_y(x,y)\right], \tag{39}$$

returning the coordinates $[x_{res}, y_{res}]$ compiled from the first column of $\mathbf{L}^{-1}\mathbf{Y}$ giving

$$f_x(x,y) = a_{1,x} + a_{x,x}x + a_{y,x}y + \sum_{i=1}^{n} w_{i,x}U(\|\mathbf{p}_i - (x,y)\|). \tag{40}$$

where $\left[a_{1,x}\ a_{x,x}\ a_{y,x}\right]^T$ is the affine transform component for $x$, and likewise for the second column, where

$$f_y(x,y) = a_{1,y} + a_{x,y}x + a_{y,y}y + \sum_{i=1}^{n} w_{i,y}U(\|\mathbf{p}_i - (x,y)\|). \tag{41}$$

with $\left[a_{1,y}\ a_{x,y}\ a_{y,y}\right]^T$ as the affine component for $y$. Each point (or minutia location) can now be updated as

$$x_{new} = f_x(x,y) = x_{res} \tag{42}$$

$$y_{new} = f_y(x,y) = y_{res}. \tag{43}$$

It can be shown that the function $f(x,y)$ is the interpolation that minimises

$$I_f \propto \mathbf{W}\mathbf{K}\mathbf{W}^T = \mathbf{V}(\mathbf{L_n}^{-1}\mathbf{K}\mathbf{L_n}^{-1})\mathbf{V}^T, \tag{44}$$

where $I_f$ is the *bending energy* measure

$$I_f = \int\int_{\mathbb{R}^2} \left(\frac{\partial^2 z}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 z}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 z}{\partial y^2}\right)^2 dxdy \tag{45}$$

and $\mathbf{L_n}$ is the $n \times n$ sub-matrix of $\mathbf{L}$.

Since *ill-posed* mappings of control points which violate mapping existence, uniqueness, or continuity, can readily exist in real world examples, the use of a *Regularization* term like Wahba (1990), $\lambda.I_f$, can be included in order to smooth the performed interpolation. Thus, the minimization of the error term

$$H[f] = \sum_{i=1}^{n} (v_i - f(x_i,y_i))^2 + \lambda.I_f \tag{46}$$

is performed, where the matrix $\mathbf{K}$ is replaced with $\mathbf{K} + \lambda.\mathbf{I}$. One should note that $\lambda = 0$ results in exact interpolation.

Using the regularized T.P.S transformation method, $n$ iterations are applied, where each iteration has minutiae mappings reassigned and transformation re-estimated using the previous minutiae set transformed state of the test fingerprint (note: the template remains static).

### 3.1.2 Similarity score

Once the $n$ iterations are performed, the final pairs have now been established. From this, the shape similarity distance measure can be calculated as

$$D_{sc}(P, Q) = \frac{1}{n} \sum_{p \in P} \arg \min_{q \in Q} C(p, T(q)) + \frac{1}{m} \sum_{q \in Q} \arg \min_{p \in P} C(p, T(q)) \qquad (47)$$

where $T(.)$ denotes the T.P.S transformed representative of the contour point $q$. In addition, an appearance term, $D_{ac}(P, Q)$, measuring pixel intensity similarity and a bending energy term, $D_{be}(P, Q) = I_f$, can be added to the similarity score.

Afterward, the similarity measure was modified as

$$D_{sc}^* = D_{sc} + \beta D_{be}. \qquad (48)$$

Although this measure does not take into account the strict one-to-one mapping of minutiae; through experimentation, this method proved to be sound, providing acceptable performance in fingerprint similarity assessment. However, the resulting minutiae mapping from the application of the Hungarian algorithm on the contextually based cost histograms produced some un-natural pairs, as illustrated in Figure 8 (top). This is due to the lack of a minutiae pair pruning procedure. The existence of un-natural pairs could potentially skew the Thin Plate Spline (T.P.S) linear transform performed. In addition, such pairs generally increase the bending energy substantially, thus leading to invalid matching results, particularly for genuine matches.

### 3.2 Proposed matching method

Recently, hybrid matching algorithms have been used for fingerprint matching. Although minutiae detail alone can produce a highly discriminant set of information, the combination of level 1 feature, such as orientation and frequency, and other level features, can increase discriminant information, and hence, increase matching accuracy, as illustrated in Benhammadi et al. (2007), Youssif et al. (2007), Reisman et al. (2002), and Qi et al. (2004).

The detailing of the proposed hybrid matching algorithm based on a modified version of the enhanced shape context method in Kwan et al. (2006), along with the integration of the orientation-based descriptor of Tico & Kuosmanen (2003) is given here, illustrating a significant performance improvement over the enhanced shape context method of Kwan et al. (2006). The main objective of the integration is two-fold, firstly to prune outlier minutiae pairs, and secondly to provide more information to use in similarity assessment.

As briefly described earlier in section 2.3, the orientation-based descriptor in Tico & Kuosmanen (2003) utilises the orientation image to provided local samples of orientation around minutiae in a concentric layout. Each orientation sample point is calculated as

$$\theta_{c,d}^{A_i} = \min(|\theta_{c,d}^{s_A} - \theta_{A_i}|, \pi - |\theta_{c,d}^{s_A} - \theta_{A_i}|) \qquad (49)$$

being the $d^{th}$ sample on the $c^{th}$ concentric circle with distance $r_c$ away from the minutiae point $m_{A_i}$, where $\theta_{A_i}$ and $\theta_{c,d}^{s_A}$ are the minutia and sample point orientation estimations, respectively. The orientation distance of equation 22 is used to prune outlier pairs resulting from the Hungarian algorithm which produced the mapping permutation of equation 30. Although the shape context defines the similarity measure in equation 47-48 with no strict one-to-one correspondence, minutiae should have a more strict assessment based on the optimal mapping, since minutiae are key landmarks as opposed to randomly sampled contour
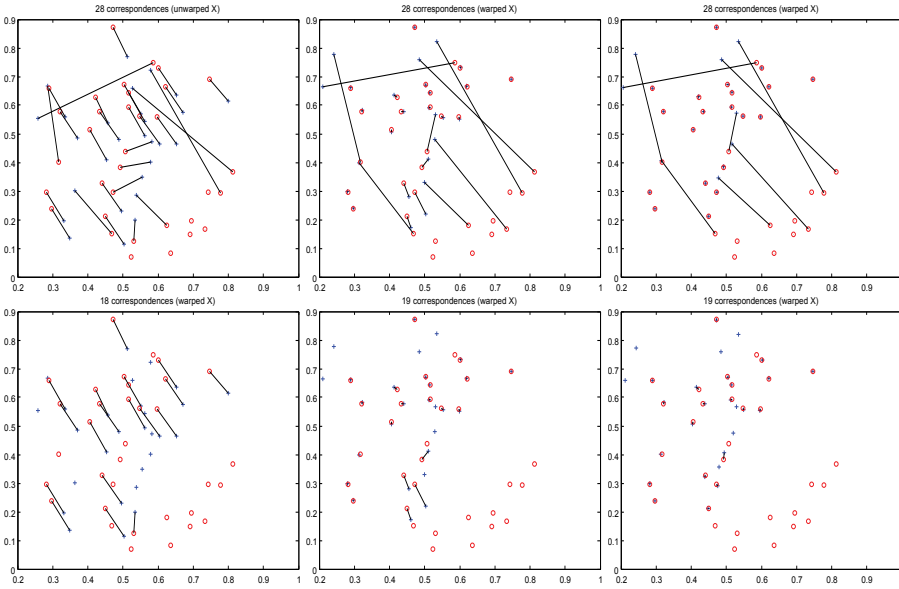
Fig. 8. **top:** The Enhanced Shape Context method producing initial minutiae pair correspondences. The next two images are for the following iterations. One should note that there are some clear un-natural pairs produced, where *plate foldings* are evident (i.e. pair correspondences that cross other pair correspondences). **bottom:** The proposed hybrid method initial minutiae pair correspondences along with following iterations of produced correspondences.

points. Thus, equation 48 can be modified to only score the pairs with the optimal one-to-one mapping as

$$D_{sc}^{**}(P, Q) = \frac{1}{n} \sum_{p_i \in P | D_o(p_i, q_{\pi(i)}) < \delta} C(p_i, q_{\pi(i)}) + \Lambda D_o(p_i, q_{\pi(i)}) + \beta D_{be} \qquad (50)$$

with an addition term in the summation to account for orientation distance scaled by the tunable parameter, $\Lambda$ with range $[0, 1]$.

In terms of what concentric circle radii and sample configuration should be used, the method explained in Tico & Kuosmanen (2003) prescribes that the radius for circle $K_l$ be $r_l = 2l \times \tau$, where $\tau$ denotes the average ridge period, and for the sample configuration, the circle $K_l$ should have roughly $\lceil \frac{\pi r_l}{\tau} \rceil$. As the average ridge period was recorded to be 0.463 mm in Stoney (1988), for a fingerprint image with dots per inch (dpi) equal to $R$, the previous formula for the configuration can be expressed as $K_l = \lceil \frac{172.r_l}{R} \rceil$. However, this was really only used as a rough guide for the configuration used in the experimentation of Tico & Kuosmanen (2003). This was also used as a rough guide for our implementation of the orientation-based descriptor.

Although the log-polar space adequately provides extra importance toward neighbouring sample points, additional emphasis on local points may be desirable, since minutiae sets are largely incomplete and do not entirely overlap. For a given minutia, non-local bin regions may be partially or largely outside the segmented region of interest for one fingerprint and not the other. In addition, the non-local bins are spatially larger, and hence, have a higher probability

of containing false minutia caused by noises. Such issues can potentially result in incorrect minutiae pairs. Thus, the enhanced shape context's log-polar histogram cost in equation 27 is adjusted to contain a tunable Gaussian weighting of histogram bin totals, depending on their distances away from the centre (reference minutia), with

$$C^{**}(p_i, q_j) = \left(1 - \gamma C_{ij}^{type} C_{ij}^{angle}\right) \cdot \left(\frac{1}{2} \sum_{k=1}^{K} \frac{\left[h_{p_i}(k) - h_{q_j}(k)\right]^2}{h_{p_i}(k) + h_{q_j}(k)} \times \exp\left(-\frac{(r_k - r_{min})^2}{2\sigma^2}\right)\right) \quad (51)$$

where $r_{min}$ is the outer boundary of the closest bin, $r_k$ is the current bin outer boundary distance, and $\sigma^2$ is a tunable parameter (see Figure 9).
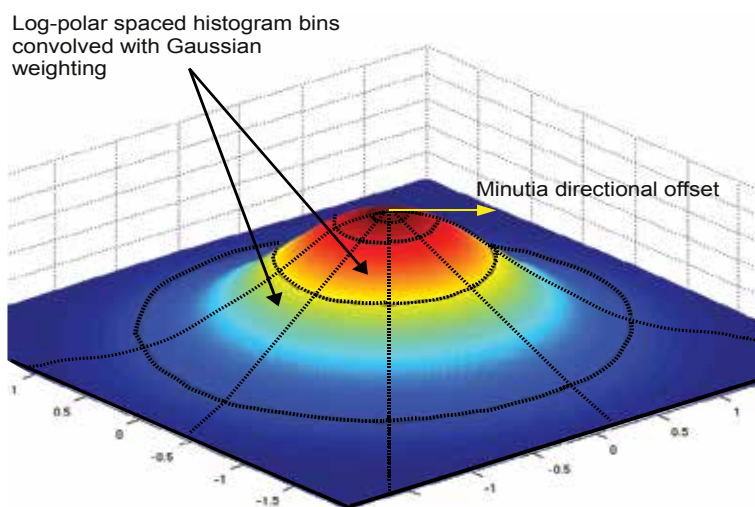


Fig. 9. The log-polar sample space convolved with a two dimensional Gaussian kernel, resulting in each bin to be weighted according to its distance from the origin (minutia). The histogram cost calculation uses the direction of the reference minutia as the directional offset for the bin order, making the descriptor invariant to rotation.

### 3.2.1 Adaptive greedy registration

If we re-examine the registration of the enhanced shape context method of Kwan et al. (2006) discussed in section 3.1, we can summarise the iterative process as updating minutiae pairs with the enhanced shape context descriptor, followed by using T.P.S to perform a global alignment with the linear transform component, and then performing non-linear transform to model warping caused by skin elasticity. This method does not utilise any singularities for the registration process, solely relying on the T.P.S framework for registration. Hence, registration is largely reliant on the accuracy of the spatial distribution of minutiae relative to a given reference minutia, which is often inadequate.

The proposed method does not use the T.P.S transform to perform the initial affine transform. Instead, a greedy method similar to Tico & Kuosmanen (2003) (of equations 20-23) is used.

Using the minutiae set representation of equations 3 and 4, each possible pair $(m_{A_i}, m_{B_j})$ of the affine transform, $T$, is calculated by using the orientation differences of each minutiae direction as the rotation component, and the difference in x-y coordinates as the offset component (as illustrated in equations 7-9). The heuristic that requires maximising is

$$H(A, B) = \arg\max_{\psi} \frac{\sum_i S(m_{A_i}, m_{B_{T(i)}})}{n(\Phi_\psi)} \tag{52}$$

where $S(.)$ is previously defined in equation 22 as the similarity function based on the orientation-based descriptor, $\psi$ is the index pair reference for the transform $T$ (i.e. $(i, j)$ with $m_{A_i} = m_{B_{T(j)}}$), $\Phi_\psi$ is the *anchor point* set (with size $n(\Phi_\psi)$) of minutiae pairs $(m_{A_p}, m_{B_q}) \in \Phi_\psi$ which have each other as closest points from the opposite minutiae sets with distance less than an empirically set limit, $\delta_M$, after applying the given transform. In addition, all anchor point minutiae pairs must have similar orientation, hence meeting the constraint

$$\min\left(|\theta_{A_i} - \theta_{B_j}|, \pi - |\theta_{A_i} - \theta_{B_j}|\right) < \delta_\theta \tag{53}$$

(see Figure 5 left). From the given definition, we can easily verify that $(i, j) \in \Phi_\psi$.

If the primary core point exists in both the test and template feature sets, provided that the core point detection is highly accurate with a given test dataset, then we have good reason to ignore affine transforms that cause the core points to be greater than a fixed distance, $\delta_D$, away from each other. Additional pruning can be achieved by only allowing transforms where the reference minutiae pair has a orientation-based descriptor similarity score to meet

$$S\left(p_i, q_{\pi(i)}\right) < \delta_S \tag{54}$$

where $\delta_S$ is empirically set. Such restrictions will not only improve performance, but also make the registration process much more accurate, since we are not just blindly maximising the heuristic of equation 52.

Many registration algorithms use singularities, texture, and minutiae pair information as tools for finding the most likely alignment. However, the overlapped region shape similarity retrieved from minutiae spatial distribution information provides an additional important criteria. After finding the bounding box (overlapping region) of a possible affine transform meeting all prior restrictions, we can then measure shape dissimilarity via the application of the shape context to all interior points $P \subseteq A$ and $Q \subseteq B$ with equation 50, giving additional criteria for affine transforms to meet $D_{sc}(P, Q) < \delta_{sc}$ for an empirically set threshold, $\delta_{sc}$.

For a candidate transform, we have the corresponding anchor point set $\Phi_\psi$. We will now define two additional nearest neighbour sets of the interior points in the overlap region (as depicted in Figure 10) as

$$\Phi_\alpha = \left\{i \mid \sqrt{(x_{A_i} - x_{B_j})^2 + (y_{A_i} - y_{B_j})^2} < \delta_N\right\} \tag{55}$$

where $\Phi_\alpha$ contains all interior nearest neighbour indices from fingerprint $A$, and likewise, $\Phi_\beta$, containing interior nearest neighbour indices for fingerprint $B$. Thus, we have

$$\Phi_\psi \subseteq \left\{(i, j) \mid i \in \Phi_\alpha \textbf{ and } j \in \Phi_\beta\right\}. \tag{56}$$

with $1 \leq i \leq p$ and $1 \leq j \leq q$. The affine transform method is detailed in algorithm 1.

If a candidate affine transform meets the heuristic of equation 52 with the given constraints,

---

**Algorithm 1** Proposed registration algorithm (Affine component)

---

**Require:** minutiae set $P$ and $Q$ (representing fingerprints A and B, respectively).

$H \leftarrow 0$

$\psi \leftarrow$ NIL

$\Phi_\psi \leftarrow$ NIL

$Q_T \leftarrow$ NIL

**for all** possible minutiae pairs $\left(p_i, q_j\right)$ **do**

$$S\left(p_i, q_j\right) \leftarrow (1/K) \sum_c^L \sum_d^{K_c} \exp\left(-\frac{2\left(\min(|\theta_{c,d}^{p_i} - \theta_{c,d}^{q_j}|, \pi - |\theta_{c,d}^{p_i} - \theta_{c,d}^{q_j}|)\right)}{\pi\mu}\right)$$

**if** $S\left(p_i, q_j\right) < \delta_S$ **then**

continue

**end if**

{Perform transform from minutiae pair with offset and orientation parameters from minutiae x-y and $\theta$ differences}

$Q' \leftarrow T(Q)$

{Make sure core points are roughly around the same region, provided both cores exist}

**if** $BothCoresExist(P, Q')$ **and** $CoreDist(P, Q') > \delta_D$ **then**

continue

**end if**

{Get T.P.S cost matrix (see algorithm 2)}

$C^{**} \leftarrow TPS_{cost}(P, Q')$

$D_{sc} \leftarrow \frac{1}{n} \sum_{p \in P} \arg\min_{q \in Q'} C^{**}(p, q) + \frac{1}{m} \sum_{q \in Q'} \arg\min_{p \in P} C^{**}(p, q)$

**if** $D_{sc} > \delta_{sc}$ **then**

continue

**end if**

$\Phi_{\psi_P} \leftarrow \left\{(a, b) \mid \arg\min_{a \in P} \sqrt{(x_{A_a} - x_{B_b})^2 + (y_{A_a} - y_{B_b})^2}\right\}$

$\Phi_{\psi_Q} \leftarrow \left\{(a, b) \mid \arg\min_{b \in Q} \sqrt{(x_{A_a} - x_{B_b})^2 + (y_{A_a} - y_{B_b})^2}\right\}$

$\Phi'_\psi \leftarrow \left\{(a, b) \mid (a, b) \in \Phi_{\psi_P} \cap \Phi_{\psi_Q} \text{ and } \sqrt{(x_{A_i} - x_{B_j})^2 + (y_{A_i} - y_{B_j})^2} < \delta_M\right\}$

$H_{test} \leftarrow \frac{\sum_{(a,b) \in \Phi'_\psi} S(p_a, T(q'_b))}{n(\Phi'_\psi)}$

**if** $H_{test} > H$ **then**

$H \leftarrow H_{test}$

$\psi \leftarrow (i, j)$

$\Phi_\psi \leftarrow \Phi'_\psi$

$Q_T \leftarrow Q'$

**end if**

**end for**

$\Phi_\alpha \leftarrow \left\{i \mid \sqrt{(x_{A_i} - x_{B_j})^2 + (y_{A_i} - y_{B_j})^2} < \delta_N\right\}$

$\Phi_\beta \leftarrow \left\{j \mid \sqrt{(x_{A_i} - x_{B_j})^2 + (y_{A_i} - y_{B_j})^2} < \delta_N\right\}$

**return** $\Phi_\psi$, $\Phi_\alpha$, $\Phi_\beta$, $Q_T$

---

---

**Algorithm 2** $TPS_{cost}$: Calculate TPS cost matrix

---

**Require:** minutiae set $P$ and $Q$ (representing fingerprints A and B, respectively).

 **for all** minutiae $q_j \in Q$ **do**

  **for** $k = 1$ to $K$ **do**

   $h_{q_j}(k) \leftarrow \# \left\{ q_j \neq q_i \ : \ (q_j - q_i) \in bin(k) \right\}$

  **end for**

 **end for**

 **for all** minutiae $p_i \in P$ **do**

  **for** $k = 1$ to $K$ **do**

   $h_{p_i}(k) \leftarrow \# \left\{ p_j \neq p_i \ : \ (p_j - p_i) \in bin(k) \right\}$

  **end for**

 **end for**

 **for all** minutiae $p_i \in P$ and $q_j \in Q$ **do**

  $C_{ij}^{type} \leftarrow \begin{cases} -1 & \text{if } type(p_i) = type(q_j), \\ 0 & \text{if } type(p_i) \neq type(q_j) \end{cases}$

  $C_{ij}^{angle} \leftarrow -\frac{1}{2} \left( 1 + \cos((\angle_{initial-warped})) \right)$

  $C^{**}(p_i, q_j) \leftarrow \left( 1 - \gamma C_{ij}^{type} C_{ij}^{angle} \right) \cdot \left( \frac{1}{2} \sum_{k=1}^{K} \frac{\left[ h_{p_i}(k) - h_{q_j}(k) \right]^2}{h_{p_i}(k) + h_{q_j}(k)} \times \exp \left( -\frac{(r_k - r_{min})^2}{2\sigma^2} \right) \right)$

 **end for**

 **return** $C^{**}$

---

we can then focus on the non-affine aspect of registration. The T.P.S non-affine component adequately modeled the warping caused by skin elasticity in the previously proposed method. With this in mind, it will be desirable to utilise the non-linear component.

The T.P.S has its point correspondence method modified through having anchor point correspondences remain static throughout the iterative process, thus attempting to restrict the affine transform component of T.P.S while finding new correspondences from $\Phi_\alpha$ on to $\Phi_\beta$ that do not exist in $\Phi_\psi$. The shape context's log-polar histogram cost is modified from equation 51 as

$$C^\gamma(p_i, q_j) = \gamma_d \gamma_\theta C^{**}(p_i, q_j) \tag{57}$$

where $\gamma_d$ is defined as

$$\gamma_d = \begin{cases} 1 & \text{if } dist(p_i, q_j) < \delta_{max}, \\ \infty & \text{otherwise} \end{cases} \tag{58}$$

with $\delta_{max}$ set as the maximum feasible distance caused by warping after applying the candidate affine transform, and similarly, $\gamma_\theta$, is defined as

$$\gamma_\theta = \begin{cases} 1 & \text{if } \min(|\theta_{A_i} - \theta_{B_j}|, \pi - |\theta_{A_i} - \theta_{B_j}|) < \theta_{max}, \\ \infty & \text{otherwise} \end{cases} \tag{59}$$

with $\theta_{max}$ set as the maximum feasible orientation difference caused by orientation estimation error in the extraction process.
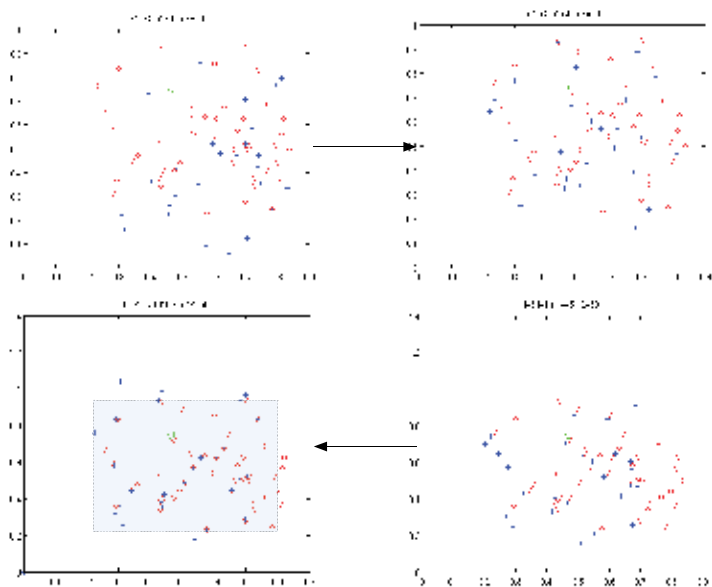
Fig. 10. An example affine transform candidate search sequence with final state and corresponding bounding box for the overlapped region.

>From the new cost function, the Hungarian algorithm is then used to find additional one-to-one correspondences to what already existed in the anchor point set. The candidate affine transform in comparison to the affine component of T.P.S,

$$\mathbf{A_T} = \begin{bmatrix} a_{x,x} & a_{x,y} & a_{1,x} \\ a_{y,x} & a_{y,y} & a_{1,y} \\ 0 & 0 & 1 \end{bmatrix},$$ (60)

is used to assess the accuracy of the additional correspondences found. The T.P.S affine component would not have prominent translation, rotation, and shear parameters since the candidate transform should have already adequately dealt with the affine registration task as the Euclidean constraints on the anchor point worked to keep the global transform rigid. In addition, the inclusion of additional natural minutiae correspondences should not significantly alter the affine registration required. Thus, the lack of prominent translation, rotation, and shear parameters for the T.P.S affine transform indicates an existing agreement between both affine transforms. Evaluation of the translation distance is given by

$$r_{affine} = \sqrt{a_{1,x}^2 + a_{1,y}^2} < r_{max}.$$ (61)

Using the Singular Value Decomposition ($SVD$) of the non-translation components of $\mathbf{A_T}$:

$$SVD\left(\begin{bmatrix} a_{x,x} & a_{x,y} \\ a_{y,x} & a_{y,y} \end{bmatrix}\right) = UDV^T$$ (62)

where $U, V^T \in SO(2, \mathbb{R})$ (i.e. 2x2 dimension rotation matrices with angles $\theta_\alpha$ and $\theta_\beta$, respectively) and $D$ is a 2x2 diagonal matrix representing scaling along the rotated coordinate

**Algorithm 3** Proposed registration algorithm (Non-affine component)

**Require:** minutiae set $P$ and $Q_T$ (for fingerprints A and transform B), anchor set $\Phi_\psi$, nearest neighbourhood sets $\Phi_\alpha$ and $\Phi_\beta$.

$\quad P' \leftarrow \{p_i \mid i \in \Phi_\alpha\}, Q_T \leftarrow \left\{q_j \mid j \in \Phi_\beta\right\}, \omega_{affine} \leftarrow 0, D_{be} \leftarrow 0$

$\quad$ **for all** minutiae $q_j \in Q_T$ **do**

$\quad\quad$ **for** $k = 1$ to $K$ **do**

$\quad\quad\quad h_{q_j}(k) \leftarrow \# \left\{q_j \neq q_i \ : \ (q_j - q_i) \in bin(k)\right\}$

$\quad\quad$ **end for**

$\quad$ **end for**

$\quad$ **for** $iter = 1$ to $n$ **do**

$\quad\quad$ **for all** minutiae $p_i \in P$ **do**

$\quad\quad\quad$ **for** $k = 1$ to $K$ **do**

$\quad\quad\quad\quad h_{p_i}(k) \leftarrow \# \left\{p_j \neq p_i \ : \ (p_j - p_i) \in bin(k)\right\}$

$\quad\quad\quad$ **end for**

$\quad\quad$ **end for**

$\quad\quad$ **for all** minutiae $p_i \in \Phi_\alpha$ and $q_j \in \Phi_\beta$ **do**

$\quad\quad\quad C_{ij}^{type} \leftarrow \begin{cases} -1 & \text{if } type(p_i) = type(q_j), \\ 0 & \text{if } type(p_i) \neq type(q_j) \end{cases}$

$\quad\quad\quad C_{ij}^{angle} \leftarrow -\frac{1}{2}\left(1 + \cos((\angle_{initial-warped}))\right)$

$\quad\quad\quad C^{**}(p_i, q_j) \leftarrow \left(1 - \gamma C_{ij}^{type} C_{ij}^{angle}\right) \cdot \left(\frac{1}{2}\sum_{k=1}^{K} \frac{\left[h_{p_i}(k) - h_{q_j}(k)\right]^2}{h_{p_i}(k) + h_{q_j}(k)} \times \exp\left(-\frac{(r_k - r_{min})^2}{2\sigma^2}\right)\right)$

$\quad\quad\quad C^{\gamma}(p_i, q_j) \leftarrow \gamma_d \gamma_\theta C^{**}(p_i, q_j)$

$\quad\quad$ **end for**

$\quad\quad$ {calculate and add minutiae pairs additional to the anchor set pairs.}

$\quad\quad \Phi_\pi \leftarrow$ Hungarian$(C^{\gamma}, \Phi_\alpha, \Phi_\beta, fixedMap = \Phi_\psi)$

$\quad\quad f(x,y) \leftarrow$ T.P.S $(P, Q_T, \Phi_\pi, Regularized = true)$

$\quad\quad D_{be} \leftarrow D_{be} + \mathbf{WKW}^T$

$\quad\quad r_{affine} \leftarrow \sqrt{a_{1,x}^2 + a_{1,y}^2}$

$\quad\quad [U, D, V^T, \theta_\alpha, \theta_\beta] \leftarrow SVD\left(\begin{bmatrix} a_{x,x} & a_{x,y} \\ a_{y,x} & a_{y,y} \end{bmatrix}\right)$

$\quad\quad \omega_{affine} \leftarrow \omega_{affine} + |\theta_\alpha + \theta_\beta|$

$\quad\quad \tau_{affine} \leftarrow \log\left(\frac{D_{1,1}}{D_{2,2}}\right)$

$\quad\quad$ **if** $D_{be} > E_{max}$ **or** $r_{affine} \geq \delta_{max}$ **or** $\omega_{affine} \geq \theta_{max}$ **or** $\tau_{affine} \geq \tau_{max}$ **then**

$\quad\quad\quad$ **return** $\Phi_\psi$

$\quad\quad$ **end if**

$\quad\quad$ **for all** minutiae $p_i \in P$ **do**

$\quad\quad\quad [x, y] \leftarrow [p_i(x), p_i(y)]$

$\quad\quad\quad [p_i(x), p_i(y)] \leftarrow [f_x(x,y), f_y(x,y)]$

$\quad\quad$ **end for**

$\quad$ **end for**

$\quad$ **return** return $\Phi_\pi$

axes of $V^T$ with nonnegative diagonal elements in decreasing order, the evaluation of rotation

$$\omega_{affine} = |\theta_\alpha + \theta_\beta| < \omega_{max}, \tag{63}$$

and shear

$$\tau_{affine} = \log\left(\frac{D_{1,1}}{D_{2,2}}\right) < \tau_{max}, \tag{64}$$

is performed for empirically set values $\omega_{max}$ and $\tau_{max}$. If the above affine transform criteria of equations 61, 63, and 64 are not met, no extra minutiae pairs are produced. Unlike the previous method, this helps uphold spatial consistency by not creating un-natural pairs (see Figure 8 (bottom)). Essentially, the candidate affine transform is used as the ground truth registration over the T.P.S affine component.

A final integrity check of the validity of the additional minutiae pairs produced from the non-affine transform is the measured bending energy, previously defined in equation 45. If the non-affine transform produces a bending energy distance $D_{be} > E_{max}$, then all additional minutiae pairs are also rejected, in order to avoid un-natural warping to occur. The non-affine transform component is detailed in algorithm 3.

### 3.2.2 Matching algorithm

Once the minutiae pairs have been established, pruning is performed to remove unnatural pairings. However, if we closely analyse the orientation-based descriptor used for pruning, we can see that a fundamental flaw arises with partial fingerprint coverage, specifically for minutiae pairs near fingerprint image edges. In such a case, the typical formula for distance calculation cannot count orientation samples that lie outside the region of interest, and therefore, unnecessarily reduces the orientation distance measure (see Figure 11). Moreover, regions that have high noise also cannot have their orientation reliably estimated due to information to be missing (if regions with high noise are masked), and likewise, reduces the orientation-based descriptor common region coverage.

A proposed modification to the orientation-based descriptor is applied so that the amount of common region coverage that each descriptor has is reflected in the similarity score. This is achieved by a simple Gaussian weighting of equation 22 with

$$S^*(m_{A_i}, m_{B_j}) = S(m_{A_i}, m_{B_j}) \times \exp(-\max(0, \Delta_{cutoff} - \Delta_{g\_count}).\mu_s) \tag{65}$$

where $\Delta_{cutoff}$ is the cutoff point where all good sample totals below this value are weighed, $\Delta_{g\_count}$ is the total number of good samples (i.e. where a good sample is defined to be in a coherent fingerprint region), and $\mu_s$ is a tunable parameter. However, for a more exhaustive approach, one could empirically review the estimated distribution of orientation-based similarity scores for true and false cases, with specific attention towards the effect of coverage completeness on the accuracy of the similarity measure.

Equation 65 relies on the intersection set of valid samples for each minutiae, defined as

$$I(A_i, B_j) = \{s \mid s \in \{L, K_c\} \text{ and } valid(A(s_x, s_y))\} \cap \{t \mid t \in \{L, K_c\} \text{ and } valid(B(t_x, t_y))\} \tag{66}$$

where $L$ is the sample position set and $K_c$ is the concentric circle set. Thus, we can also define a variant of the function $S^*(m_{A_i}, m_{B_j}, I)$ where a predefined sample index set , $I$, is given to indicate which samples are only to be used for the similarity calculation, ignoring $i \notin I$ even if corresponding orientation samples are legitimately defined for both fingerprints (note: this variant is used later for similarity scoring in the matching algorithm).

Fig. 11. An example where the orientation-based descriptor for corresponding minutiae in two different impressions of the same fingerprint have no coverage in a substantial portion of the orientation sample.

After the filtered minutiae pair set is produced, we can now assess the similarity of the pairs. Minutiae $\delta$-neighbourhood structure families, where particular spatial and minutiae information of the $\delta$ closest minutiae to a reference minutiae are extracted as features, have been used before in minutiae based matching algorithms (such as Chikkerur & Govindaraju (2006) and Kwon et al. (2006)) for both alignment and similarity measure (see Figure 12). The $\delta$-neighbourhood structure proposed has the following fields:

- Distance $d_i$: the distance a neighbourhood minutia is away from the reference minutia.

- Angle $\angle_i$: the angle a neighbourhood minutia is from the reference minutia direction.

- Orientation $\theta_i$: the orientation difference between a neighbourhood minutia direction and the reference minutia direction.

- Texture $\Gamma_i$: the orientation-based descriptor sample set for a neighbourhood minutia used to measure the region orientation similarity with the reference minutia for a given sample index.

giving us the sorted set structure

$$n\delta(m_{A_i}) = \{\{d_{A_i(1)}, \angle_{A_i(1)}, \theta_{A_i(1)}, \{\Gamma_{A_i(1)}\}\}, \ldots, \{d_{A_i(\delta)}, \angle_{A_i(\delta)}, \theta_{A_i(\delta)}, \{\Gamma_{A_i(\delta)}\}\}\} \tag{67}$$

being sorted by the distance field in ascending order. The first two fields are considered as the polar co-ordinates of the neighbourhood minutiae with the referencing minutia as the origin, and along with the third field, they are commonly used in local neighbourhood minutiae based structures ( Kwon et al. (2006) and Jiang & Yau (2000)). Using the modified weighted orientation-based descriptor, the texture field provides additional information on how each local orientation information surrounding the $\delta$-neighbourhood minutiae set vary from that of the reference minutia, with the measure rated using equation 65. These structures can now be used to further assess and score the candidate minutiae pairs.
When comparing the $\delta$-neighbourhood elements of a candidate minutiae pair, the first three fields are straight forward to compute the differences with

Fig. 12. The $\delta$-neighbourhood ($\delta = 4$) for a given minutia (ignoring false crease minutiae).

$$r_{diff(m_{A_i}, m_{B_j})}(x, y) = |d_{A_i(x)} - d_{B_j(y)}|, \tag{68}$$

$$\angle_{diff(m_{A_i}, m_{B_j})}(x, y) = \min(|\angle_{A_i(x)} - \angle_{B_j(y)}|, 2\pi - |\angle_{A_i(x)} - \angle_{B_j(y)}|), \tag{69}$$

$$\theta_{diff(m_{A_i}, m_{B_j})}(x, y) = \min(|\theta_{A_i(x)} - \theta_{B_j(y)}|, \pi - |\theta_{A_i(x)} - \theta_{B_j(y)}|), \tag{70}$$

However, to make an accurate comparison of a minutiae pair's $\delta$-neighbourhood orientation similarity information, we must find the intersection of sample positions that are within valid and coherent regions for *both* images (see Figure 13). Thus, the fourth field is composed of the set of all orientation samples and is used to dynamically calculate the orientation difference of neighbouring minutiae for a minutiae pair's respective $\delta$-neighbourhoods, using

$$\Gamma_{diff(m_{A_i}, m_{B_j})}(x, y) = |\Gamma_{A_i(x)} - \Gamma_{B_j(y)}| \tag{71}$$

where $\Gamma_{A_i(x)} = S^*(m_{A_i}, m_{A_i(x)}, I_o)$, $\Gamma_{B_j(y)} = S^*(m_{B_j}, m_{B_j(y)}, I_o)$, and overlap index set $I_o = I(A_i, A_i(x)) \cap I(B_j, B_j(y))$.

We now require a systematic method for scoring a minutiae pair given their corresponding $\delta$-neighbourhoods. Since there is no guarantee that each $\delta$-neighbourhood has the same minutiae set, optimal mapping methods, such as the Hungarian algorithm, may not be desirable. Instead, we use a novel greedy algorithm, where we iterate through one neighbourhood and find the best match from another, provided that they meet pre-defined affine constraints of equations 61-63. Removal of the matching elements from the $\delta$-neighbourhood lists ensures one-to-one mappings of $\delta$-neighbourhood minutiae.

A $\delta$-neighbourhood similarity score is tallied for the candidate minutiae pair $(m_{A_i}, m_{B_j})$ for each matched $\delta$-neighbourhood minutiae pair having respective neighbourhood sorted set indexes $(x_s, y_t)$, found to also match, using

$$sim_\delta(m_{A_i}, m_{B_j}) = \sum_{s,t} (\alpha(x_s, y_t) + \gamma\beta(x_s, y_t)) \tag{72}$$

where

$$\alpha(x, y) = \exp(-r_{diff(m_{A_i}, m_{B_j})}(x, y) - \theta_{diff(m_{A_i}, m_{B_j})}(x, y) - \angle_{diff(m_{A_i}, m_{B_j})}(x, y)) \tag{73}$$
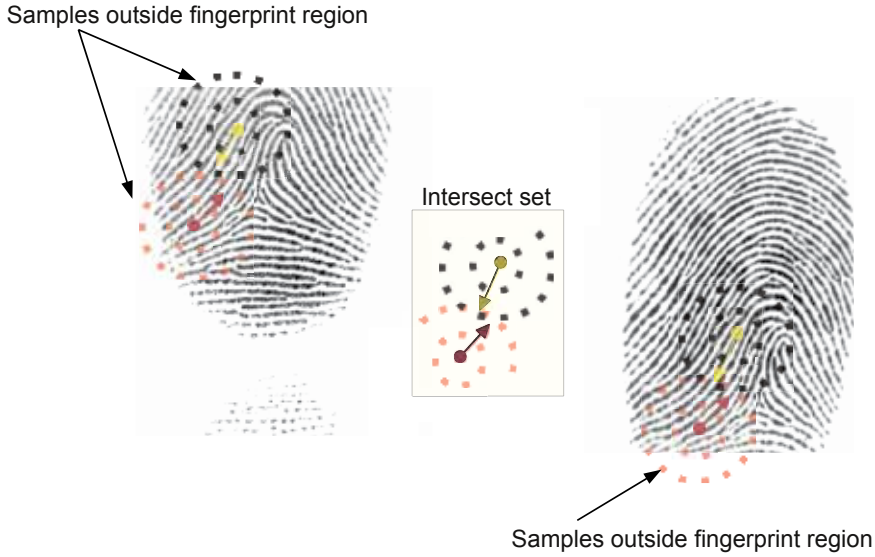
Fig. 13. An example where the orientation-based descriptor for corresponding minutiae in two different impressions of the same fingerprint have no coverage in a substantial portion of the orientation sample.

and

$$\beta(x,y) = \exp(-\Gamma_{diff(m_{A_i}, m_{B_j})}(x,y)) \tag{74}$$

with a tunable parameter $\gamma$ defined in $[0,1]$.

After all candidate minutiae pair $\delta$-neighbourhoods have been scored, we can now find a fingerprint matching similarity score as

$$sim(A,B) = \frac{n_M \left( \sum_{(i,j)} sim_\delta(m_{A_i}, m_{B_j}) \right) \cdot \left( \sqrt{S_{max}} \right)}{n_A . n_B} - \nu D_{sc}^{**} \tag{75}$$

where

$$S_{max} = \arg\max_{(i,j)} S^*(m_{A_i}, m_{B_j}, I(A_i, B_j)) \tag{76}$$

with $S^*(m_{A_i}, m_{B_j}, I(A_i, B_j))$ defined in equation 65 as the orientation-based descriptor similarity measure, $D_{sc}^{**}$ is the modified shape context distance in equation 50, $n_M$ is the number of matching filtered minutiae pairs, $n_A$ and $n_B$ are the number of minutiae in the overlap region from fingerprint $A$ and $B$, respectively, $i$ and $j$ are the index of the filtered minutiae pair elements in fingerprint A and B, respectively, and $\nu$ a tunable parameter in $[0,1]$. Additionally, we can add the type similarity of each pair to equation 71 by adding a small constant, $\zeta_t$, when the minutiae types agree. The matching method is summarised in algorithm 4.

---

**Algorithm 4** Proposed matching algorithm

---

**Require:** $P, Q_T$, candidate minutiae pair index set $\Phi_\pi$, and $n\delta_A = \left\{ n\delta(m_{A_i}) \mid i \in \Phi_\pi(1) \right\}$ and
  $n\delta_B = \left\{ n\delta(m_{B_j}) \mid j \in \Phi_\pi(2) \right\}$ as the neighbourhood sets
  $sim_\delta \leftarrow Nil$
  **for all** minutiae pair indexes $(i,j) \in \Phi_\pi$ **do**
    $sim_\delta(m_{A_i}, m_{B_j}) \leftarrow 0$
    $O_{score} \leftarrow S(m_{A_i}, m_{B_j})$
    **if** $O_{score} < O_{min}$ **then**
      remove $(i,j) \in \Phi_\pi$
      continue
    **end if**
    **for all** $s \in n\delta_{A_i}$ **do**
      **for all** $t \in n\delta_{B_j}$ **do**
        $r_{diff(m_{A_i}, m_{B_j})}(s,t) \leftarrow |d_{A_i(s)} - d_{B_j(t)}|$
        $\angle_{diff(m_{A_i}, m_{B_j})}(s,t) \leftarrow \min(|\angle_{A_i(s)} - \angle_{B_j(t)}|, 2\pi - |\angle_{A_i(s)} - \angle_{B_j(t)}|)$
        $\theta_{diff(m_{A_i}, m_{B_j})}(s,t) \leftarrow \min(|\theta_{A_i(s)} - \theta_{B_j(t)}|, \pi - |\theta_{A_i(s)} - \theta_{B_j(t)}|)$
        **if** $r_{diff(m_{A_i}, m_{B_j})}(s,t) > r_{max}$ **or** $\angle_{diff(m_{A_i}, m_{B_j})}(s,t) > \angle_{max}$ **or**
          $\theta_{diff(m_{A_i}, m_{B_j})}(s,t) > \theta_{max}$ **then**
            continue
        **end if**
        remove $s \in n\delta_{A_i}$
        remove $t \in n\delta_{B_j}$
        $\alpha(s,t) \leftarrow \exp(-r_{diff(m_{A_i}, m_{B_j})}(s,t) - \theta_{diff(m_{A_i}, m_{B_j})}(s,t)) - \angle_{diff(m_{A_i}, m_{B_j})}(s,t))$
        $\beta(s,t) \leftarrow \exp(-\Gamma_{diff(m_{A_i}, m_{B_j})}(s,t))$
        $sim_\delta(m_{A_i}, m_{B_j}) \leftarrow sim_\delta(m_{A_i}, m_{B_j}) + \alpha(s,t) + \gamma\beta(s,t)$
      **end for**
    **end for**
  **end for**
  {calculate shape context from the minutiae sets $P$ (which has been non-affinely transformed)
  and $Q_T$ (which has been affinely transformed)}
  $C^{**} \leftarrow TPS_{cost}(P, Q_T)$
  $D_{sc}^{**}(P, Q) \leftarrow \frac{1}{n} \sum_{p_i \in P | D_o(p_i, q_{\pi(i)}) < \delta} C(p_i, q_{\pi(i)}) + \Lambda D_o(p_i, q_{\pi(i)}) + \beta D_{be}$
  $S_{max} \leftarrow \arg\max_{(i,j)} S^*(m_{A_i}, m_{B_j}, I(A_i, B_j))$
  $sim_{score} \leftarrow \frac{n_M \left( \sum_{(i,j)} sim_\delta(m_{A_i}, m_{B_j}) \right) \cdot \left( \sqrt{S_{max}} \right)}{n_A . n_B} - \nu D_{sc}^{**}$
  **return** $sim_{score}$

---

## 4. Experimental results

The experiment was performed on two databases, the FVC2002 database Db1 set A ( Maio et al. (2002)) which contains 800 fingerprint images with 100 fingers having 8 impressions each, and fingerprint database from the University of Bologna ( Bologna (2000)), consisting of 168 fingerprint images formed by 21 fingers with 8 impressions each. The parameters of the algorithm (see Table 1) were tuned with the FVC2002 database Db1 set B, which contains 80 fingerprints (10 fingers with 8 impressions each). The program was written in Matlab and run on a 1.66GHz Linux PC with 2Gb memory. The FVC2002 protocol ( Maio et al. (2002)) was used in experimentation, comprising of $n \times 8 \times 7/2 = 2800$ genuine and $n \times (n-1)/2 = 4950$ imposter attempts for the FVC2002 database, and $n \times 8 \times 7/2 = 588$ genuine and $n \times (n-1)/2 = 210$ imposter attempts for the smaller database.

Image enhancement for the FVC2002 database was performed via the STFT method ( Chikkerur et al. (2004)), while the binarization/thinning based minutiae extraction method and smoothed orientation image creation proposed in Hong et al. (1998) were used for feature extraction, along with the core point detection algorithm based on the method described in Julasayvake & Choomchuay (2007). Spurious minutiae had very crude filtering applied, with only short spurs and minutiae near segmented border regions removed. All other false minutia structures remained in the feature set. In addition, the thinning algorithm is known to produce a higher number of spurious minutia in comparison to ridge following methods. Moreover, the STFT method was noted in Jirachaweng et al. (2009) to produce a high rate of spurious minutiae, while the extraction method encountered a substantial amount of minutia type interchange (approx. 30%). Thus, we can expect our feature set to have at least a moderate amount of noise.

Table 2 summarises the performance of our algorithm against numerous well known algorithms on the FVC2002 database, as does Table 3 with the University of Bologna database. One should note that the parameters were not tuned for this second database, but still managed to perform quite well. Figures 14 and 15 show the FMR vs. FNMR graphs. The proposed method managed to finish in a top 8 position for the FVC2002 database. Figure 16 illustrates the genuine and imposter distributions of the similarity score for both databases.
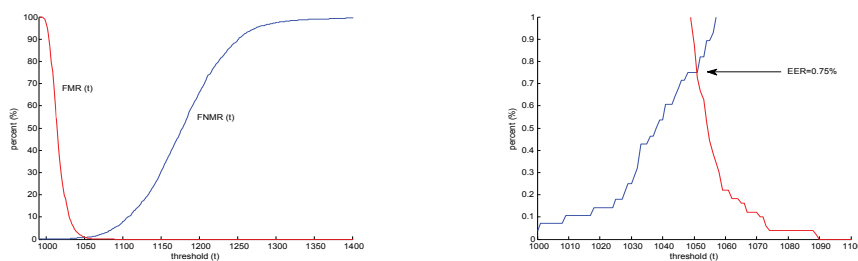


Fig. 14. **left:** Proposed method FNMR and FMR vs matching threshold $t$ on FVC2002 Db1 Set A **right:** Close up illustrating the EER (0.75%).

## 5. Conclusions

The proposed fingerprint matching method using a hybrid shape and orientation descriptor outperforms many well-known methods on the FVC2002 database (in the top 8th place) in the

| Parameter Name | Hybrid Component | Relevance | Value |
|---|---|---|---|
| iterations | shape context | T.P.S | 5 |
| annealing rate | shape context | T.P.S | 0.35 |
| regularization | shape context | T.P.S | 0.8 |
| inner/outer radii | shape context | log-polar | 1/64, 2 |
| radii/theta bins | shape context | log-polar | 8, 10 |
| $\sigma^2$ | shape context | log-polar | 4.5 |
| circle radii $r_l$ | orientation | minutia | 7, 14, 21, 28, 35, 42, 49 |
| circle samples | orientation | minutia | 12, 16, 22, 28, 32, 36, 36 |
| $\nu$ | scoring | shape context | 0.3 |
| $O_{min}$ | matching | orientation | 0.25 |
| $\delta_{max}$ | registration | T.P.S | 0.1 |
| $\delta_S$ | registration | orientation | 0.25 |
| $\delta_D$ | registration | core | 30 pixels |
| $\delta_{sc}$ | registration | shape context | 4.9 |
| $r_{max}$ | registration | T.P.S | 0.1 |
| $\omega_{max}$ | registration | T.P.S | $\pi/6$ |
| $\tau_{max}$ | registration | T.P.S | 0.3 |
| $E_{max}$ | registration | T.P.S | 12 |
| $\Delta_{cutoff}$ | matching | filtering | 70 |
| $\angle_{max}$ | matching | T.P.S | $\pi/8$ |
| $r_{max}$ | matching | T.P.S | 0.05 |
| $\theta_{max}$ | matching | T.P.S | $\pi/8$ |
| $\delta$ | matching | neighbourhood | 4 |

Table 1. Parameters setup for experimentation
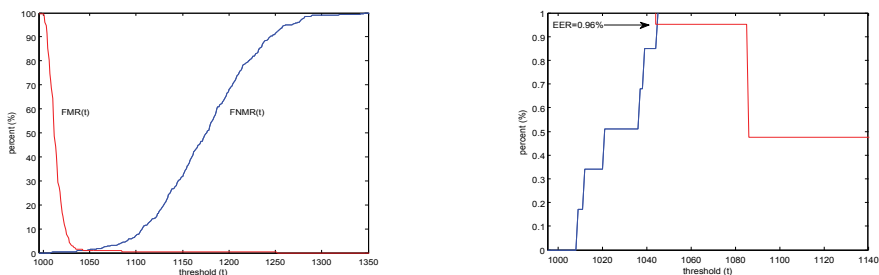


Fig. 15. **left:** Proposed method FNMR and FMR vs matching threshold $t$ on University of Bologna database **right:** Close up illustrating the EER (0.96%).

FVC2002 competition, considering that the feature set was not in pristine condition due to the chosen extraction and filtering methods, highlighting the overall robustness of the proposed algorithm.

In addition, we improved the performance of the algorithm substantially over the enhanced shape context on both public datasets, despite using parameters only tuned for the FVC2002 database. Finally, all known competing matching algorithms tested on the University of Bologna database were beaten by the proposed fingerprint matching method.
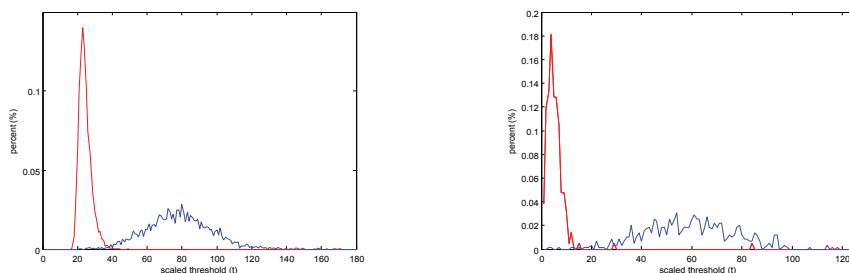
Fig. 16. Genuine (blue) and imposter (red) distributions for the **left:** FVC2002 Db1 set A database, and **right:** University of Bologna database.

| Matching Algorithm | EER (%) |
|---|---|
| CBFS  Chikkerur & Govindaraju (2006) | 1.50 |
| TPS based  Kwon et al. (2006) | 0.92 |
| Meshgrid based  Kwon et al. (2007) | 0.82 |
| Hybrid Spiral based  Shi & Govindaraju (2009) | 1.98 |
| PA08  Maio et al. (2002) ($8^{th}$ place) | 0.98 |
| PB35  Maio et al. (2002) ($5^{th}$ place) | 0.61 |
| PA15  Maio et al. (2002) ($1^{st}$ place) | 0.1 |
| **Proposed method** | **0.75** |

Table 2. Performance comparison of matching algorithms on FVC2002 Db1 Set A

| Matching Algorithm | EER (%) |
|---|---|
| WGHT/Orientation-based  Tico (2001)  Tico et al. (2002) | 1.07-1.97 |
| Mutual Information  Liu et al. (2006) | 1.5 |
| Delaunay Triangulation  Wang & Gavrilova (2006) | 5.1 |
| Enhanced Shape Context  Kwan et al. (2006) | 12.79 |
| **Proposed method** | **0.96** |

Table 3. Performance comparison of matching algorithms on University of Bologna database

The matlab source code for the proposed fingerprint matching algorithm can be found at the Matlab Central ( Abraham (2010)) website.

## 6. References

Abraham, J. (2010). Matlab code for fingerprint matching algorithm.
    URL: *http://www.mathworks.com/matlabcentral/fileexchange/29280-fingerprint-matching
    -algorithm-using-shape-context-and-orientation-descriptors*
Bazen, A. M. & Gerez, S. H. (2003).  Fingerprint matching by thin-plate spline modelling of
    elastic deformations, *Pattern Recognition* 36(8): 1859–1867.
Belongie, S., Malik, J. & Puzicha, J. (2000).  Shape context: A new descriptor for shape
    matching and object recognition, *Neural Information Processing Systems Conference
    (NIPS)*, pp. 831–837.

Belongie, S., Malik, J. & Puzicha, J. (2002). Shape matching and object recognition using shape contexts, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(4): 509–522.

Benhammadi, F., Amirouche, M., Hentous, H., Bey Beghdad, K. & Aissani, M. (2007). Fingerprint matching from minutiae texture maps, *PR* 40(1): 189–197.

Bologna (2000). Biometric system laboratory: www2.csr.unibo.it/research/biolab, 168 fingerprint database (21 fingers 8 impressions).

Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations, *IEEE Trans. Pattern Anal. Mach. Intell.* 11(6): 567–585.

Chikkerur, S. & Govindaraju, V. (2006). K-plet and CBFS: A graph based fingerprint representation and matching algorithm, *International Conference on Biometrics (accepted)*.

Chikkerur, S. & Ratha, N. (2005). Impact of singular point detection on fingerprint matching performance, *Automatic Identification Advanced Technologies, IEEE Workshop on* 1: 207–212.

Chikkerur, S., Wu, C. & Govindaraju, V. (2004). A systematic approach for feature extraction in fingerprint images, *Biometric Authentication* LNCS 3072: 344–350.

Farina, A., Kovacs Vajna, Z. & Leone, A. (1999). Fingerprint minutiae extraction from skeletonized binary images, *PR* 32(5): 877–889.

Gonzalez, R. C. & Woods, R. E. (2007). *Digital Image Processing (3rd Edition)*, Prentice Hall.

Hatano, T., Adachi, T., Shigematsu, S., Morimura, H., Onishi, S., Okazaki, Y. & Kyuragi, H. (2002). A fingerprint verification algorithm using the differential matching rate, *Pattern Recognition, International Conference on* 3: 30799.

Hong, L., Wan, Y. & Jain, A. (1998). Fingerprint image enhancement: Algorithm and performance evaluation, *"IEEE Transactions on Pattern Analysis and Machine Intelligence"* 20(8): 777–789.

Jiang, X. & Yau, W.-Y. (2000). Fingerprint minutiae matching based on the local and global structures, *Pattern Recognition, International Conference on* 2: 6038.

Jirachaweng, S., Leelasawassuk, T. & Areekul, V. (2009). Performance and computational complexity comparison of block-based fingerprint enhancement, *ICB09*, pp. 656–665.

Jonker, R. & Volgenant, A. (1987). A shortest augmenting path algorithm for dense and sparse linear assignment problems, *Computing* 38(4): 325–340.

Julasayvake, A. & Choomchuay, S. (2007). An algorithm for fingerprint core point detection, *Proc. of International Symposium on Signal Processing and its Applications 2007 (ISSPA-2007)* 1(1): 1–4.

Kisel, A., Kochetkov, A. & Kranauskas, J. (2008). Fingerprint minutiae matching without global alignment using local structures, *Informatica* 19(1): 31–44.

Kwan, P. W., Gao, J. & Guo, Y. (2006). Fingerprint matching using enhanced shape context, *Proceedings of The 21st Image and Vision Computing New Zealand (IVCNZ 2006) Great Barrier Island, New Zealand* 1(1): 115–120.

Kwon, D., Yun, I. D., Kim, D. H. & Lee, S. U. (2006). Fingerprint matching method using minutiae clustering and warping, *Pattern Recognition, International Conference on* 4: 525–528.

Kwon, D., Yun, I. D. & Lee, S. U. (2007). A robust warping method for fingerprint matching, *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on* 0: 1–6.

Liang, X. & Asano, T. (2006). Fingerprint matching using minutia polygons, *ICPR06*, pp. I: 1046–1049.

Lindoso, A., Entrena, L., Liu Jimenez, J. & San Millan, E. (2007). Correlation-based fingerprint matching with orientation field alignment, *ICB07*, pp. 713–721.

Liu, L., Jiang, T., Yang, J. & Zhu, C. (2006). Fingerprint registration by maximization of mutual information, *IP* 15(5): 1100–1110.

Maio, D. & Maltoni, D. (1997). Direct gray-scale minutiae detection in fingerprints, *PAMI* 19(1): 27–40.

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. & Jain, A. K. (2002). "fvc2002: Second fingerprint verification competition", pp. 811–814.

Nanni, L. & Lumini, A. (2009). Descriptors for image-based fingerprint matchers, *Expert Systems with Applications* 36(10): 12414–12422. Cited By (since 1996): 4.
URL: *www.scopus.com*

Qi, J., Wang, Y., Shi, Z., Xu, K. & Zhao, X. (2004). Fingerprint matching integrating the global orientation field with minutia, *ICBA*, pp. 337–343.

Ratha, N. K. & Bolle, R. (2003). *Automatic Fingerprint Recognition Systems*, SpringerVerlag.

Reisman, J., Jain, A. & Ross, A. (2002). A hybrid fingerprint matcher, *ICPR02*, pp. III: 795–798.

Rutovitz, D. (1966). "pattern recognition", *"J. Royal Statist. Soc"* (129): 504–530.

Shi, Z. & Govindaraju, V. (2009). Robust fingerprint matching using spiral partitioning scheme, *ICB '09: Proceedings of the Third International Conference on Advances in Biometrics*, Springer-Verlag, Berlin, Heidelberg, pp. 647–655.

Stoney, D. (1988). Distribution of epidermal ridge minutiae, *Am. J. Physical Anthropology* 77: 367–376.

Tian, L., Chen, L. & Kamata, S.-i. (2007). Fingerprint matching using dual hilbert scans, *SITIS '07: Proceedings of the 2007 Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, IEEE Computer Society, Washington, DC, USA, pp. 593–600.

Tico, M. (2001). *On Design and Implementation of Fingerprint-Based Biometric Systems, PhD thesis*, PhD thesis, Tampere Univ. of Technology, Tampere, Finland.

Tico, M. & Kuosmanen, P. (2003). Fingerprint matching using an orientation-based minutia descriptor, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(8): 1009–1014.

Tico, M., Onnia, V. & Kuosmanen, P. (2002). Fingerprint image enhancement based on second directional derivative of the digital image, *EURASIP J. Appl. Signal Process.* 2002(1): 1135–1144.

Wahba, G. (1990). *Spline Models for Observational Data*, SIAM.

Wang, C. & Gavrilova, M. L. (2006). Delaunay triangulation algorithm for fingerprint matching, *ISVD '06: Proceedings of the 3rd International Symposium on Voronoi Diagrams in Science and Engineering*, IEEE Computer Society, Washington, DC, USA, pp. 208–216.

Xiao, Q. & Raafat, H. (1991). Fingerprint image post-processing: A combined statistical and structural approach, *PR* 24(10): 985–992.

Yager, N. & Amin, A. (2005). Coarse fingerprint registration using orientation fields, *JASP* 2005(13): 2043–2053.

Yang, J. C. & Park, D.-S. (2008). A fingerprint verification algorithm using tessellated invariant moment features, *Neurocomputing* 71(10-12): 1939–1946.

Youssif, A. A., Chowdhury, M. U., Ray, S. & Nafaa, H. Y. (2007). Fingerprint recognition system using hybrid matching techniques, *"Computer and Information Science, ACIS International Conference on"* 0: 234–240.

Zhang, W. & Wang, Y. (2002). Core-based structure matching algorithm of fingerprint verification, *Pattern Recognition, International Conference on* 1: 10070.

Zhao, F. & Tang, X. (2007). Preprocessing and postprocessing for skeleton-based fingerprint minutiae extraction, *PR* 40(4): 1270–1281.

# Fingerprint Spoof Detection Using Near Infrared Optical Analysis

Shoude Chang[1], Kirill V. Larin[2], Youxin Mao[1],
Costel Flueraru[1] and Wahab Almuhtadi[3]

*[1]Institute for Microstructural Sciences, National Research Council Canada, Ottawa*
*[2]Department of Biomedical Engineering, University of Houston, Houston*
*[3]Algonquin College, Ottawa*
*[1,3]Canada*
*[2]USA*

## 1. Introduction

Fingerprints have been used for several centuries as a means of identifying individuals. Since every fingerprint is considered to be unique, fingerprint recognition is the most popular biometric identification method currently employed in such areas as law enforcement, financial transactions, access control, and information security. Fingerprints consist of ridges and furrows on the surface of a fingertip. The ridges are the raised portions of the fingerprint while the furrows are the spaces between the ridges. Recognition can be performed based on ridge ending and ridge bifurcation (Xiao & Raffat, 1990), tessellated invariant moment (Yang & Park, 2008), and image-based features (Nanni & Lumini, 2009). Since the ridges are created by nature, people may consider that stealing and duplicating a fingerprint is more difficult than stealing a password or token, but it turns out that it is not difficult to make an artifact to fool an automated fingerprint system. It has been reported that an automated fingerprint authentication system could be defeated either using "a combination of low cunning, cheap kitchen supplies and a digital camera" , or simply by creating false thumbprint images . These sensor-level attacks are called "spoofing" attacks in which an artifact containing a copy of the fingerprint traits of a legitimate enrolled user is used to fool a fingerprint system. The first step is to obtain the fingerprint of a legitimate user, which can be accomplished by lifting a latent print either with or without the co-operation of the fingerprint owner. Next, molding plastic and gelatin can be used to make "gummy fingers". Finally, the resulting fake fingers can be used to fool the fingerprint sensor and attack the security system.  The vulnerability to fake-finger attack has generated a wave of research concerned with adding "liveness detection" to improve system resistance to spoofing. Liveness detection is the ability to determine whether a biometric sample is being provided by a live human being rather than from a copy created using an artifact. The detection methods can be categorized into two groups: hardware-based and software-based. In hardware-based solutions, extra hardware must be integrated with biometric sensors to detect additional information such as heartbeat, temperature, and the tissue under the epidermis. For example, an extra sensor can be used to measure either blood flow or pulse

in the fingertip in order to identify a living finger, although there are several problems associated with this approach. Using heart rhythm as a biometric feature is unreliable, since a person's heartbeat can vary considerably and is affected by many different factors. Furthermore, the sensor can be easily spoofed by adding a pipe to the fake finger and pumping saltwater through the pipe to imitate blood flow. A second example of a hardware-based solution uses temperature as a liveness feature to distinguish a dummy finger from a real one. However, the temperature of the epidermis is about 8-10°C above the room temperature that is 18-20°C in an office environment. By using a silicone artificial fingerprint, the temperature only decreases by a maximum of 2°C (6-8°C), which is still in the working margins of the sensor. Another hardware-based detection method utilizes an optical sensor and multispectral imaging to capture the sub-surface of the skin to prevent spoofing. The fingerprint images are captured using different wavelengths of illumination to penetrate the skin to different depths, thereby obtaining multiple images of the surface and subsurface of the fingertip. Since these devices are more complex than conventional optical sensors, they are also more expensive.

On the other hand, software-based solutions focus on using the information captured from a standard fingerprint sensor, and liveness detection can be performed by simply modifying the algorithm to measure skin properties such as perspiration, elasticity, and deformation. The algorithms can be roughly divided into two groups based on whether they extract static or dynamic features: static approaches compare the features extracted from one or more fingerprint impressions, while dynamic methods analyze multiple frames of the same image captured over a certain time period. An early effort on software-based liveness detection used the perspiration pattern, which depends on a unique physiological feature of the skin – evaporation from the human body – to distinguish real fingers from artificial ones. Perspiration-based methods characterize and analyze the sweating pattern of live finger from two consecutive images captured over a period of a few seconds to detect the perspiration phenomenon. However, this approach is susceptible to a number of factors, such as the finger pressure applied, environmental moisture, and user cooperation. Another technique is based on the elastic deformation of the skin. When a finger touches a sensor, the elastic deformation will cause a distortion of the minutiae location. Because most of the materials used for making fake fingers are harder than human skin, the study showed that artificial fingers presented different deformations than the live ones. However, this approach performs poorly when the hardness of the fake finger is similar to that of live skin. The main advantages of software solutions are that no additional equipment needs to be integrated and there is less chance of revealing an individual's health status due to privacy concerns.

In order to take the advantages and avoid disadvantages of software-based and hardware-based methods, a combined software-hardware approach is presented to defeat fingerprint spoofing attack in this chapter. Two methods are presented based on analyzing different optical properties between human skin and artificial finger. The first method uses optical coherence tomography (OCT) technology and the second performs spectral analysis. Not only do both methods capture the ridge pattern appearing on the skin, but also measure the internal properties of the skin defined as internal biometrics to differentiate real and prosthetic fingers. This chapter is structured as follows: Section 2 introduces the layered structure of human skin and analyzes the skin's optical properties demonstrating its complexity, making it difficult to fake. Section 3 introduces OCT technology, presents a

concept of internal biometrics, and presents a solution to distinguish prosthetic and real fingers using internal biometrics measured with OCT. Section 4 investigates the differences of light reflection properties between prosthetic and real fingers, and presents an imaging system to explore the spectral features of prosthetic and real fingers. The experimental results, presented in Section 3 and Section 4, demonstrate the validity of OCT and spectral analysis as a software-hardware fingerprint anti-spoofing technique. Finally, a conclusion summarizes the main contributions of the work, discusses the technical challenges, and indicates the future research directions.

## 2. Optical properties of human skin

Since prosthetic and real fingers are often indistinguishable on the surface, it is necessary to study aspects of human skin, analyze its optical properties, and identify the features that make it difficult to replicate. In this section, we look at the multilayered skin structure of fingerprints. First, we introduce the major skin layers, and then analyze how they affect the skin's optical properties, such as scattering, absorption, and penetration depth, which will be used in performing near infrared (NIR) optical analysis to differentiate real and prosthetic fingers. For example, scattering and absorption are the two main physicochemical phenomena that occur with light inside the skin. A fake gelatin finger is a homogeneous medium that typically presents a significantly lower scattering profile than that of human skin.

The skin is the largest organ of the body and is composed of specialized epithelial and connective tissue cells. The skin has many important functions: it serves as a barrier to the environment as well as a channel for communication to the outside world; it protects the body from water loss and impact wounds; it uses specialized pigment cells to protect the body from ultraviolet radiation; and it helps to regulate body temperature and metabolism.

The skin is composed of several layers (Figure 1). The innermost layer contains subcutaneous fatty tissue with stores of adipose tissue. Above this is the dermis layer, which consists of connective tissue, blood vessels, nerve endings, hair follicles, and sweat and oil glands. The outermost layer of skin is called the epidermis.

The epidermis is mainly composed of keratinocytes, which differentiate into five layers: the Stratum Basale, the Stratum Spinosum, the Stratum Granulosum, the Stratum Lucidum, and the Stratum Corneum. The thickness of the epidermis is approximately 60-80 μm but varies greatly with age, gender, and location on the body. For example, the epidermis on the underside of the forearm is few cell-layers thick, but is an order of magnitude thicker on the sole of the foot.

The dermis is the next major skin layer just below the epidermis. The dermis is approximately 1-2 mm thick in humans and is divided into two layers: the papillary dermis and the reticular dermis. The dermis includes collagen (Type I collagen) and reticulin (Type III collagen), which provide tensile strength. Elastic fibers provide for the restoration of shape after a deformation. Fibroblasts (synthesize collagen, elastin, and reticulin), histiocytes, endothelial cells, perivascular macrophages and dendritic cells, mast cells, smooth muscle, and cells of peripheral nerves and their end-organ receptors are the cell lines which are found in the dermis.

Below the dermis is the hypodermis (subcutis). This layer contains adipose tissue and serves to attach the dermis to its underlying tissues. This fatty tissue also serves as a heat-isolator, protective layer and energy reservoir. Larger blood and lymphatic vessels criss-cross this layer.
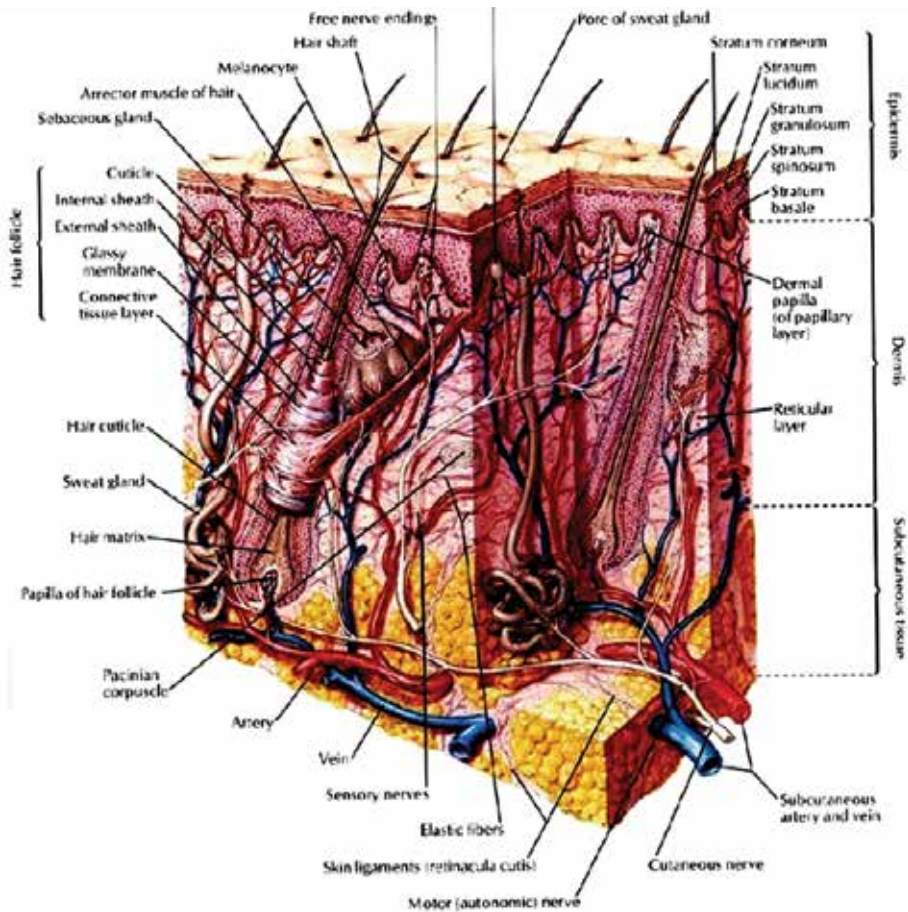
Fig. 1. Structure of the skin. Adapted from (Netter, 1997)

The skin is a highly inhomogeneous optical object and is almost impossible to mimic in a man-made phantom. Light interaction with the multilayer and multicomponent skin is a very complicated process. Non-uniform distribution of density and refractive index makes the skin a highly scattering media. Mean refractive index of background fluid can be calculated as weighted average of the refractive indices of interstitial fluid ($n_{ISF}$) and cytoplasm ($n_{cyt}$):

$$\overline{n}_{fluid} = \varphi_{cyt}n_{cyt} + (1 - \varphi_{cyt})n_{ISF} \, ,\tag{1}$$

where $\phi_{cyt}$ is the volume fraction of cytoplasm in tissues and is approximately equal to 0.6. Average refractive indices of the interstitial fluid and cytoplasm are 1.355 and 1.367, respectively, yielding a refractive index for the background fluid of 1.362. The refractive index of other skin constituents such as melanin ($n = 1.7$), collagen ($n$ of fully hydrated = 1.43), adipose tissue ($n = 1.46$), cellular nuclei ($n = 1.39$-1.41) and interstitial fluid ($n = 1.34$) cause the dermis and the epidermis regions to be highly scattering in the NIR spectral region as the optical turbidity of tissues is mainly caused by the refractive index mismatch between the intracellular and the extracellular components.

The absorption coefficient $\mu_a$ (cm$^{-1}$) and the reduced scattering coefficient $\mu'_s$ (cm$^{-1}$) are two optical parameters used to describe the absorption and scattering properties of tissue. The absorption coefficient of skin expresses how far light of a particular wavelength can penetrate into the skin before it is absorbed. The reduced scattering coefficient is related to the scattering coefficient, $\mu_s$ (cm$^{-1}$), and the anisotropy factor of scattering, g, based on the relationship $\mu'_s = \mu_s$ (1-g); it is used to describe the diffusion of photons in a random walk of step size of $1/\mu'_s$, where each step involves an isotropic scattering.

The penetration depth of light in skin is dependent on both the absorption and scattering coefficients, and also on wavelength. An estimation of the light penetration depth δ can be performed with the relation:

$$\delta = \frac{1}{\sqrt{3\mu_a(\mu_a + \mu'_s)}} \tag{2}$$

In the ultraviolet (UV) and infrared (IR) ($\lambda \geq 2$ μm) spectral regions, light is readily absorbed, which accounts for the small contribution of scattering and the inability of radiation to penetrate deep into skin (only through one or two cell layers) ; it is limited within the epidermis layer. For short-wave visible light, scattering and absorption both occur, with a penetration depth of 0.5 - 1.5 mm in human skin. In the wavelength range 0.6 - 1.4 μm, the penetration depth in skin reaches 1.5 - 3.5 mm, and in this case, scattering prevails over absorption. When the wavelength is close to 2 μm, the penetration depth appears relatively stable around 0.7 - 1.3 mm in skin.

## 3. Spoof detection using optical coherence tomography

In Section 2, we saw how the complexity of skin gives it optical properties that make it difficult to fake. In this section, we look at a new method of measuring those differences: optical coherence tomography (OCT). OCT allows us to see not only the surface of the skin but also some of the subsurface characteristics detailed in the last section. By extending the existing biometrics that are based on surface scan of external features, the OCT system can probe and extract the internal features of multilayered objects and tissues. The internal biometrics based technology is more robust against the tampering and counterfeiting comparing with conventional biometric systems.

Different approaches to OCT vary in both cost and effectiveness. We discuss OCT in general and the pros and cons of the various approaches. We then present the experimental results performed by Chang et al to demonstrate that OCT has great promise as a combine hardware and software approach to fingerprint analysis and liveness detection (Chang et al, 2006; Cheng et al, 2008).

### 3.1 Internal biometrics

The traditional biometric technologies that are used for security and the identification of individuals primarily deal with fingerprints, hand geometry and face images. These traditional technologies use external features of the human body and can thus be easily spoofed or tampered with by distorting, modifying or counterfeiting the apparent features. The extraction of internal body features that are unique to individuals is now becoming a new trend for biometrics, which is termed "Internal Biometrics" (Chang et al, 2008).

In addition to the well-known technologies for iris and retina recognition, other versatile technologies for internal biometrics are currently being developed. Vein scan technology can

automatically identify a person based on the patterns of blood vessels in the back of the hand. Vein patterns are distinctive between twins, and even between a person's left and right hands. Developed before birth, they are highly stable and robust, changing throughout one's life only in overall size.

Skin pattern recognition technology measures the characteristic of an individual's skin. The exact composition of all the skin elements is distinctive to each person. For example, skin layers differ in thickness, the interfaces between the layers have different undulations, pigmentation differs, collagen fibers and other proteins differ in density, and the capillary beds have distinct densities and locations beneath the skin.

Iris recognition has been used as a biometric application since 1987. Iris recognition is based on the visible characteristics of the human iris, including rings, furrows, freckles, and the iris corona. Iridian's iris-recognition technology converts these visible characteristics into a template that can be stored for future verification. An 11-mm diameter iris can have 266 unique spots—compared to 10 to 60 unique spots for traditional biometric technologies. Another eye-related internal biometric feature involves retinal scanning, which analyses the layer of blood vessels at the back of the eye (http://en.wikipedia.org/wiki/Retinal_scan).

Fingernail-bed identification   is based on the spatial distribution of the distinct grooves in the epidermal structure directly beneath the fingernail. This structure is mimicked in the ridges on the outer surface of the nail. When an interferometer is used to detect phase changes in back-scattered light shone on the fingernail, the distinct dimensions of the nail-bed can be reconstructed and a one-dimensional map can be generated.

Previous works have shown that the ear is a promising candidate for biometric identification. In a report (Yan and Bower, 2003), authors present a complete system for ear biometrics, including an automated segmentation of the ear in a profile view image and 3D-shape matching for recognition. Authors evaluated their system with the largest experimental study to date in ear biometrics, achieving a recognition rate of 97.6%. The algorithm they developed also shows good scalability of recognition rate with size of dataset size.

Fingerprints are the most commonly used type of biometric technology used for various applications, including law enforcement, financial transactions, access control, and information security. Fingerprint recognition has several benefits over other biometric identification techniques such as iris recognition, face recognition and hand-geometry verification methods. It has been revealed that fingerprint readers can be defeated either using cheap kitchen supplies, or by creating false thumbprint images. These attack methods are called "spoofing" because they attempt to fool a biometric system by presenting a fake fingerprint trait to the sensor. With less than $10 worth of household supplies, artificial fingerprint gummies can be made and easily spoof the fingerprint system. Figure 2, 3 shows



Fig. 2. Dummy fingerprint made by polymer.

Fig. 3. Dummy fingerprint made by kitchen powder.

human fingerprints made by polymer, and by kitchen powder, respectively, on which all the visible external features are delicately created; these could successfully spoof a traditional fingerprint scanning system.

The following technologies have been proposed and tested both in hardware and software to defeat spoofing attacks:

1. Analyzing skin details through very high-resolution sensors (1000 dpi) to capture details such as sweat pores or coarseness of the skin texture.
2. Analyzing dynamic properties of the finger, such as pulse oximetry, blood pulsation, skin elasticity and skin perspiration.
3. Analyzing static properties of the finger by adding hardware to capture information such as temperature, impedance or other electric measurements, and spectroscopy.
4. Using multi-spectral imaging technology to measure the fingerprint characteristics that are at and beneath the surface of the skin.

To detect and explore the internal features for internal biometrics, special tools which have the capability of penetrating the bio-sample are needed. Having features of μm-level resolution of cross-sectional image, non-contact probing, and relatively cheap cost, optical coherence tomography becomes the best candidate for internal biometric applications.

## 3.2 Principles of OCT

Optical coherence tomography is an emerging technology for high-resolution cross-sectional imaging of 3D structures. The first OCT system was reported by Huang et al in 1991. Since then, OCT technology has been attracting the attention of researchers throughout the world.

OCT relies on the interferometric measurement of coherent backscattering variations to detect internal interface structures of tested samples, such as biomedical tissues or internal scattered and layered materials. While similar to ultrasound B-mode imaging, OCT uses an infrared light source rather than ultrasound.

OCT has several advantages over other volume-sensing systems:

- Higher resolution: This feature enables greater visualization of details. Normally OCT has a cross-sectional resolution about 5-20 microns. For comparison, ultrasound has a resolution of 150 microns; high-resolution CT has 300 microns; and MRI has 1,000 microns.
- Non-invasive, non-contact measurement: This feature increases safety and ease of use and extends the possibility for *in vivo* applications, which is important for biological applications such as biometrics.

- Fiber-optics delivery: As optical fiber diameter is normally 125 microns, it allows OCT with a miniature optic fiber probe to be used for in situ applications, particularly for tiny lumen and intravascular applications.
- High speed: The new generation of OCT technology has no mechanical scanning procedures. This allows for high-resolution 3D sensing by the full-field OCT system.
- Potential for obtaining additional information from the testing sample: Many optical properties of samples could be explored by functional OCT. For examples, polarization contrast, Doppler Effect, and spectroscopic information.
- Use of non-harmful radiation. OCT systems work with visual and infrared band, unlike traditional CT working with X-ray and ultrasound relying on mechanical vibration.

In the past decade, OCT systems have been developed mainly for medical and biomedical applications, especially for the diagnostics of ophthalmology, dermatology, dentistry (Smolka, 2008) and cardiology. To explore the capabilities of OCT system for probing the internal features of an object, references (Chang et al, 2006) reported the applications for multiple-layer information retrieval and internal biometrics (Cheng and Larin, 2006; Chang et al, 2008). In addition, because OCT has the voxel resolution of micrometer size, it has potential applications in material investigation (Wiesauera et al, 2005; Bashkansky et al, 2001; Chinn & Swanson, 1996; Dunkers et al, 1999) and artwork diagnostics. Reference (Targowski et al, 2006) describes OCT diagnostics used for museum objects, involving stratigraphic applications (Szkulmowskaki et al, 2007); varnish layer analysis (Liang et al, 2005; Rie, 1987); structural analysis and profilometric applications (Spring et al, 2008; Targowski et al, 2004 ; Yang et al, 2004; Targowski et al, 2006). In Reference (Szkulmowska et al, 2005), the use of different OCT systems for oil painting layer examination, varnish thickness determination, and environmental influence on paintings on canvas are described.

To explore the capabilities of OCT systems for probing the internal features of an object, authors have performed research in applying OCT technology for information encoding and retrieving with a multiple-layer information carrier. Since OCT has a resolution on the scale of microns and is able to peel cross-sectional images from the inside of an object, it has potential applications in documents security and object identification.

In direct imaging using an ordinary camera, all of the layers reflected from the surface of an object will be fused together in the resulting image. However, in optical coherence tomography imaging, a coherence gate generated by an interferometer and broadband light source could be used to extract cross-sectional images at different depths. The depth resolution of an OCT system is determined by the bandwidth of the light source, normally, the bandwidth is around 100 nm, and the depth resolution in air is about 7 μm.

### 3.2.1 Time-domain OCT

OCT technology originates from low coherence interferometry (LCI) (a nonscanning /imaging version of OCT) where axial (depth) ranging is provided by linearly scanned low-coherence interferometry. This method of signal acquisition is referred to as time-domain OCT (TD-OCT). TD-OCT system is typically based on a Michelson interferometer. There are two main configurations: free space and fiber-based setups.

In TD-OCT systems, a broadband light source is used in a Michelson-type interferometer. A mechanical scanning device is introduced to select different layers at different depths by moving a reference mirror (Figure 4 shows the basic concept). With a broadband light source, the motion of the mirror produces moving interference fringes, called a coherence

gate, which scan through the sample. For imaging with an ordinary camera, all of the reflected/scattered light from different layers, L1, L2…Ln, are collected together and form a fused image. However, in OCT imaging, only the layer whose optical length is the same as that in the reference arm gets modulated by the interfering fringes, i.e., framed by the coherence gate. Using a specially designed algorithm, the image of this layer can be extracted from the others. Sources with broader bandwidths have a narrower gate, and achieved finer resolutions in the cross-sectional images that are extracted.
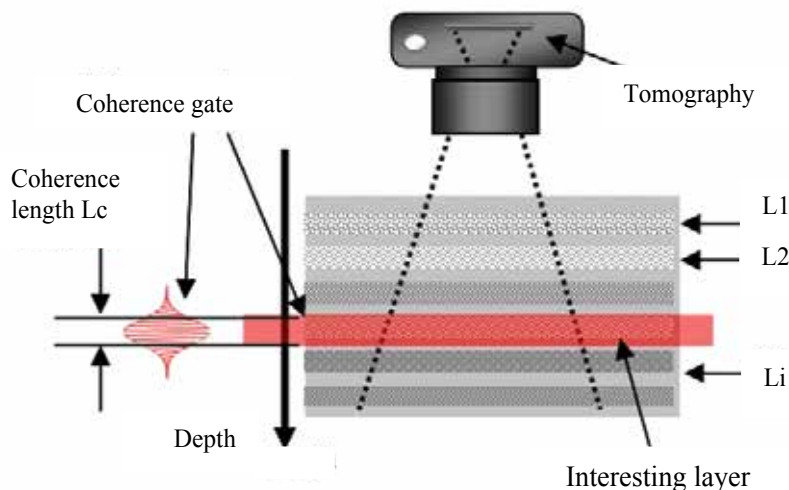


Fig. 4. Coherence gate used for separating layers

### 3.2.2 Full-field OCT

Most OCT systems are fiber-optic interferometers, a technology based on point-scanning. To get an enface image, i.e., an image consists of many A-scans, the 2D scanning is a must. Depth scanning is achieved by the longitudinal translation of a reference mirror for TD-OCT. Such a 3-axis scanning procedure makes the system slow and cumbersome. Parallel detection schemes have been investigated to increase the acquisition speed and eliminate the need for lateral scanning. Parallel OCT systems illuminate the entire 2D target and collects light from all pixels simultaneously. These parallel OCT systems are often called full-field OCT systems (FF-OCT) (Dubois et al, 2002; Akiba et al, 2007). A few OCT systems working directly on 2D full-field images have been reported. Figure 5 shows a FF-OCT system using two cameras that can perform real-time video-rate OCT imaging.

For a typical OCT light source, a super-luminescent laser diode (SLD) with central wavelength 830nm and 15 nm bandwidth, the system resolution has a depth resolution of 20 μm. To further increase the resolution of an OCT system, the broadband light source is needed. A very good candidate for the broadband light source is a white light source such as a halogen lamp. The combination of white light interferometric techniques with modern electronics and software can yield powerful measurement tools. White-light interferometry methods have already been established for the measurement of topographical features of sample surfaces, and these can be further modified to extract the cross-sectional image of an object (Chang et al, 2008). In the presented white light TD-OCT system, we use a halogen light source with a central wavelength of 700 nm and bandwidth of 200 nm. The depth resolution is 0.9 μm. The image grabbing rate can achieve 30 frames/second, with a

resolution of 1024x1024 pixels at 12-bit gray levels. The mechanical depth scanning accuracy is 37 nm. For such a high-accuracy system, alignment and fine-tuning is critical. A technology, combining automatic vision and motion control, was developed to perform this task. The imaging area is designed for 25 mm by 25 mm, which is good for most of the fingerprints. As the halogen lamp has very high power, the working area can be extended even larger. This system was built for fingerprint and document security, as well as 3D sensing.
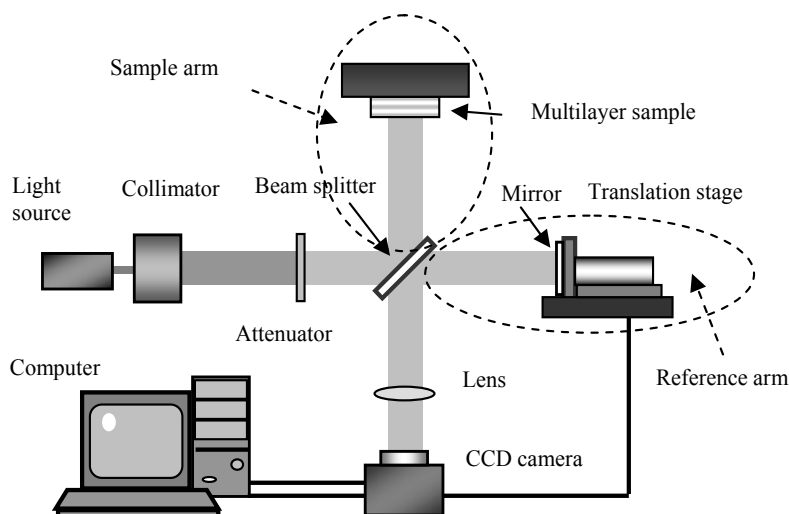


Fig. 5. FF-OCT system

### 3.3 Dummy fingerprint detection using OCT systems
### 3.3.1 Detection performed by fiber-based OCT system

#### 3.3.1.1 Description of TD-OCT system and preparation of artificial fingerprints

TD-OCT systems are useful for the noninvasive identification of artificial materials that can be used to bypass fingerprint biometric devices. High in-depth and lateral resolution versions of this technique, along with real-time image acquisition, allow for the identification of false fingerprints by analyzing the sample's optical properties to detect any extra layers of artificial materials placed on a finger.

Figure 6 shows a diagram of a TD-OCT system used in these studies. A low-coherence SLD with a wavelength of 1300±15 nm and an output power of 10 mW was used as the optical source in this system. Light in the sample arm of the interferometer was directed into the tissue sample using a single-mode optical fiber and an endoscopic probe.

The endoscopic probe allowed for the lateral scanning of sample surfaces. Light scattered from the sample and light reflected from the reference arm mirror form an interferogram that is detected by a photodetector. In-depth scanning was produced electronically by piezoelectric modulation of the fiber length, and 2D images were obtained by scanning over the sample surface in both the lateral direction (X-axis) and in-depth (Z-axis). Operation of the TD-OCT system was fully computer-controlled. The resulting images were 450 by 450 pixels with a full-image acquisition rate of approximately 3 seconds. In-depth scanning was up to 2.2 mm (in air), while lateral scanning was over 2.4 mm. Figure 7(a) shows typical

OCT image of the skin of a finger. Three layers of human skin (the stratum corneum, epidermis and dermis) are clearly visible. The 2D images were then averaged in the lateral dimension over ≈ 1 mm (sufficient for speckle-noise suppression) to obtain a single curve. The output OCT signal represents the 1D distribution of light in depth (plotted on a logarithmic scale), as shown in Figure 7(b). The resolution obtained by the system was estimated at about 25 µm in air, corresponding to about 19 µm in tissues (assuming refractive index of 1.4).
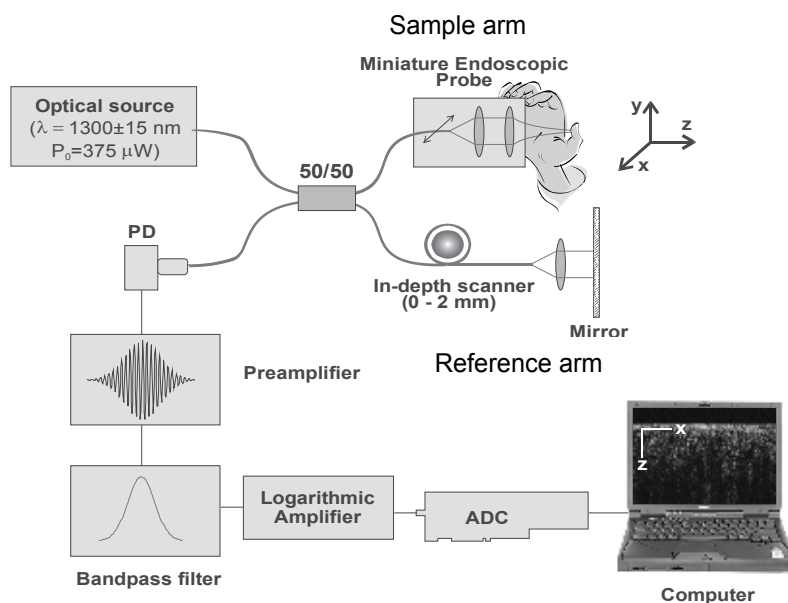


Fig. 6. Schematic diagram of the TD-OCT system (PD = photo detector, ADC = analog-to-digital converter).

This fiber-based TD-OCT system was used to study different artificial materials that might be used in preparation of artificial fingerprints. Artificial fingerprint dummies were made using a plasticine (Dixon Ticonderoga Company, Mexico), a household cement (ITW Devcon Corporation, MA), and a liquid silicone rubber (GE Silicones, General Electric Company, New York). We used general household materials that could be found in any supermarket. The following procedures were followed to make an artificial fingerprint dummy (male mold) from the plasticine (female mold).

The plasticine was cut and kneaded into thick pieces for the preparation of a female mold. To obtain the best imprint of the original fingerprint patterns, the finger was carefully washed with soap to get rid of any dust and oil. The finger was then pressed firmly into the plasticine to leave a fingerprint pattern (female mold, Figure 8(a)). To prepare the male mold, we poured glue or liquid silicone rubber into the female mold. After natural solidification, the dummy was removed and its fingerprint surface was carefully wiped to remove plasticine pieces (internal impurities such as air bubbles or tiny pieces of plasticine were still present, althou the OCT images were obtained from regions that were free from structural defects).At this point, the artificial fingerprint dummy (male mold) was ready for experimentations (as shown in Figure 8(b)).
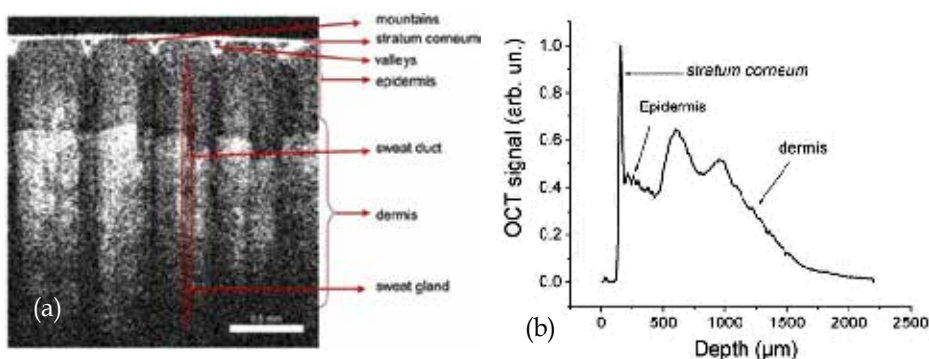
Fig. 7. A typical OCT image obtained from the skin of a thumb (a) and the corresponding OCT signal (b) showing the depths of the major skin layers.
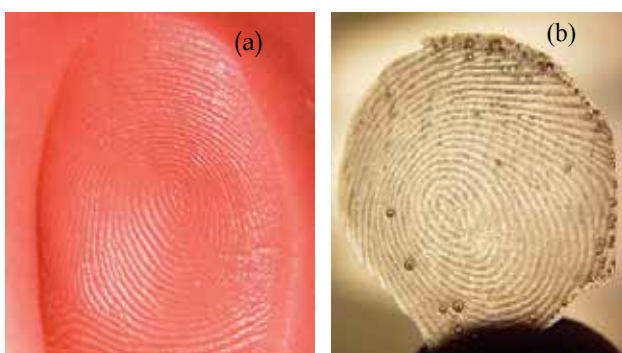


Fig. 8. Dummy fingerprint (a) Plasticene fingerprint (female mold) (b) Artificial fingerprint dummy (male mold)

The plasticine was cut and kneaded into thick pieces for the preparation of a female mold. To obtain the best imprint of the original fingerprint patterns, the finger was carefully washed with soap to get rid of any dust and oil. The finger was then pressed firmly into the plasticine to leave a fingerprint pattern (female mold, Figure 8(a)). To prepare the male mold, we poured glue or liquid silicone rubber into the female mold. After natural solidification, the dummy was removed and its fingerprint surface was carefully wiped to remove plasticine pieces (internal impurities such as air bubbles or tiny pieces of plasticine were still present, although the OCT images were obtained from regions that were free from structural defects). At this point, the artificial fingerprint dummy (male mold) was ready for experimentations (as shown in Figure 8(b)).

### 3.3.1.2 In-depth analysis of structural characteristics of TD-OCT images

Several materials that can be used to make artificial fingerprints have been studied, including gelatin, silicon, wax, and agar. Figure 9 illustrates an OCT image and the corresponding signal curve typical of a gelatin layer (25% concentration) placed over a finger. The artificial gelatin layer and the human skin layers beneath it can be clearly detected in both Figures 9(a) and 9(b). The gelatin layer is a homogeneous medium and has a significantly lower scattering profile than that of the skin, as illustrated by the OCT signal

curve between 0 and 0.2 mm. The characteristic layers of the human skin under the gelatin layer can be identified as in Figure 7(b).

Typical results obtained from a 0.2 mm thick layer of 10%-concentration agar over a thumb are shown on Figure 10. The resulting OCT image is shown in Figure 10(a), and the corresponding signal in Figure 10(b) shows the thickness of the agar and its scattering profile. Similarly, Figure 11(a) shows the OCT image of a thumb with an outer layer of silicone of 0.08 mm average thickness. The corresponding 1D graph in Figure 11(b) shows that the layer of silicon is easily distinguishable. Finally, an OCT image and the corresponding signal for a wax sample placed on a thumb are shown on Figures 12(a) and 12(b), respectively.
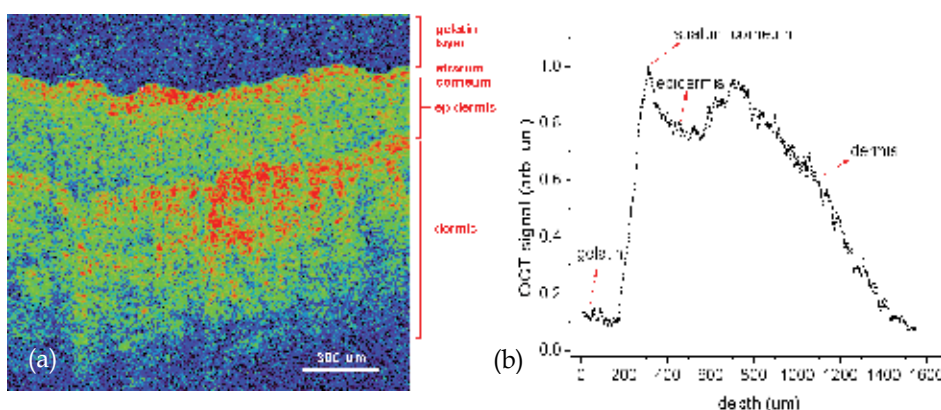


Fig. 9. OCT image (a) and the corresponding signal (b) of 25%- concentration gelatin placed over skin.

Typical results obtained from a 0.2 mm thick layer of 10%-concentration agar over a thumb are shown on Figure 10. The resulting OCT image is shown in Figure 10(a), and the corresponding signal in Figure 10(b) shows the thickness of the agar and its scattering profile. Similarly, Figure 11(a) shows the OCT image of a thumb with an outer layer of silicone of 0.08 mm average thickness. The corresponding 1D graph in Figure 11(b) shows that the layer of silicon is easily distinguishable. Finally, an OCT image and the corresponding signal for a wax sample placed on a thumb are shown on Figures 12(a) and 12(b), respectively.

To demonstrate the usefulness of OCT as a detection technique, a commercially available fingerprint reader device (Microsoft Fingerprint Reader, Model: 1033, Redmond, WA) was tested in some experiments. The fingerprint patterns of a volunteer's thumbs, forefingers, middle fingers and ring fingers from both the left and right hands were recorded and registered on a computer. The same fingers were used to prepare the artificial dummies. The dummies were then placed on another person's fingers and were scanned using both the fingerprint reader and the OCT system. Each dummy was tested 10-20 times using both systems and the corresponding FARs (False Accept Rates) were calculated. FAR represents the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user. When these artificial dummies were applied to the reader, the resulting FARs were from 80% to 100%. However, the artificial fingerprint dummies were always detected by the OCT system and the FARs were reduced to 0%.
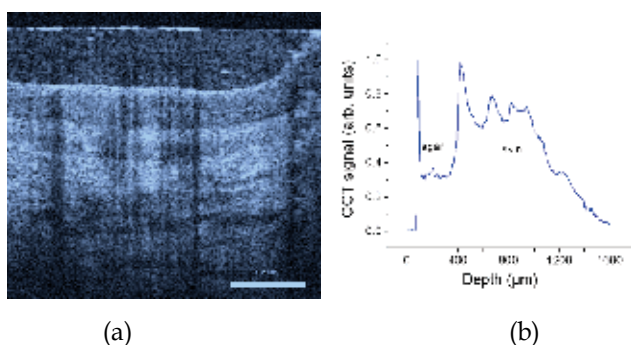
(a)                                                                                            (b)

Fig. 10. OCT image (a) and corresponding OCT signal (b) obtained from 10% agar of average thickness of 0.2 mm over a thumb.



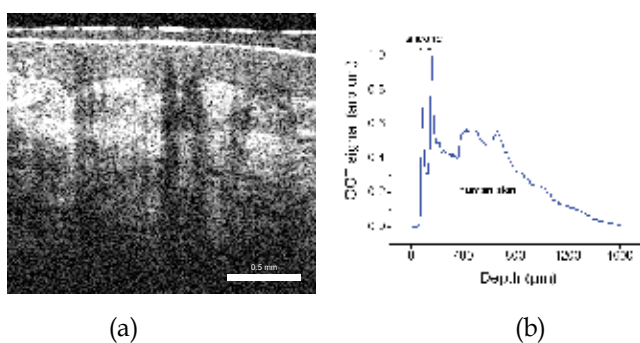(a)                                                                                            (b)

Fig. 11. OCT image (a) and corresponding OCT signal (b) obtained from silicon of average thickness of 0.08 mm over a thumb.



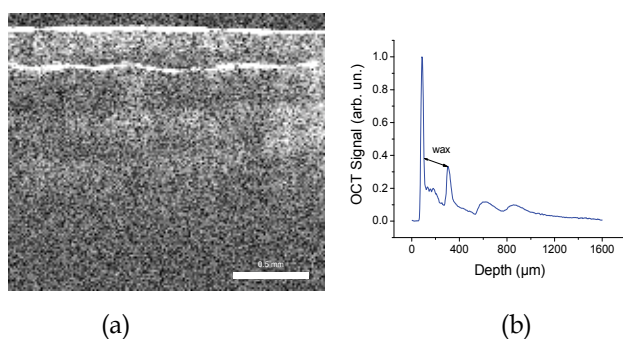(a)                                                                                            (b)

Fig. 12. OCT image (a) and corresponding OCT signal (b) obtained from wax over a thumb.

### 3.3.1.3 In-depth analysis of tissues optical properties

Another method for identifying artificial materials is based on an analysis of their optical properties. Previously, we and others demonstrated that by analyzing the exponential profile of light distribution in tissues, one can calculate its scattering coefficient. Similarly, an analysis of the optical properties of several artificial materials relative to that of human skin has revealed that it can be used for differentiation purposes. The OCT signals were

plotted on a logarithmic scale and the slopes of the OCT signal (OCTSS) were calculated at regions corresponding to artificial materials and human skin using a least-squares algorithm. Table 1 shows the calculated OCTSS values for the materials studied. From this table one can see that OCTSS (and, thus, the optical properties) of all studied artificial materials are significantly different from that of skin (except for samples made of wax). These results demonstrate that this method may help in identifying artificial materials present on human skin.

Results shown in Table 1 also demonstrate that the fingerprint dummies prepared using wax could have similar optical properties as for the dermis. Therefore, it might be difficult to differentiate wax-based artificial materials from the skin based solely on calculation of its optical properties. In such cases, combination of two or more methods might be required for more robust identification of artificial materials placed on real skin.

| Concentration / OCTSS | Gelatin | Agar | Skin (dermis) | Wax | Silicone |
|---|---|---|---|---|---|
| 100% | | | 1.0 | 1.25 | 0.08 |
| 10% | 0.07 | 0.15 | | | |
| 25% | 0.18 | 0.33 | | | |
| 33% | 0.20 | 0.38 | | | |

Table 1. OCT signal slopes calculated for the different materials (gelatin, agar, silicone, and wax) relative to that of human dermis.

### 3.3.1.4 Autocorrelation analysis of speckle variance in OCT images

Another method for robust identification of fingerprint dummies is based on a multidimensional autocorrelation analysis of OCT images. Autocorrelation refers to the cross-correlation of a signal with itself, and is a commonly used method in signal processing to analyze functions and series. Autocorrelation analysis is a useful technique in the search for repeating patterns, such as periodic signals that have been buried in noise, e.g. speckle noise. Speckles result from the coherent superposition (mutual interference) of light scattered from random scattering centers. In OCT imaging of scattering media, the speckle noise results from the coherent nature of laser radiation and the interferometric detection of the scattered light. Speckle noise substantially deteriorates resolution and accuracy of the OCT images and, therefore, several methods have been proposed to reduce its effect. However, speckle noise bears useful information about an object's properties and can be utilized in tissue classification and monitoring of different processes.

Recently, we obtained encouraging results in the application of autocorrelation analysis for distinguishing gelatin- and agar-based fingerprint dummies from skin. The method is described as follows. Two-dimensional OCT images are converted into relative intensity values and these are recorded in a square matrix (450×450 pixels). Each column in the matrix contains information about one independent Z-scan of the OCT system. A discrete autocorrelation method is applied to process data in all columns. We define the function $u(d)$, for a column intensity data in the image matrix, where $d$ is the depth ranging from 1 to 450 pixels (corresponding depths of to 0 - 1.6 mm in the sample with a refractive index of

1.4). Before autocorrelating, we remove the mean value of $u(d)$, as $x(d)=u(d)-\mu$, where $\mu = \frac{1}{N}\sum_{n=1}^{N} u(d)$ was the mean value of the function $u(d)$. The discrete autocorrelation function for $x(d)$ is

$$R_{xx}(r\Delta d) = \frac{1}{N-r}\sum_{n=1}^{N-r} x_n x_{n+1} = \frac{1}{N-r}\sum_{n=1}^{N-r} x(d)\,x(d+r\Delta d) \qquad (10)$$

in which $r$ is the depth number such that $r = 0,1,2,\ldots m$ ($m<N$), $m$ refers to the maximum depth of autocorrelation, $\Delta d =1$ - space interval unit (with value of 1 OCT pixel), $N$ is the total available depth for a given region of an OCT signal curve used in the autocorrelation analysis (either the artificial material or real skin). The value of $N$ differs for the autocorrelation functions in the artificial material and the human real finger regions; we use $m = N$-1, where $m=50$ for the artificial material region (the artificial fingerprint), and $m=100$ for the real finger region. Using this algorithm, we calculate the autocorrelation function in each column of the OCT signal matrix and then average to find the arithmetic mean value.

Figure 13 shows encouraging results obtained from the autocorrelation analysis of gelatin, agar, and real finger samples. The autocorrelation analysis was applied in the regions of OCT images corresponding to the artificial materials and human skin. Depth, shown in pixels (1 pixel is approximately equal to 3.5 µm), refers to the interval where the OCT signal compared with itself. As depth increases, the autocorrelation function values for gelatin and agar fall sharply to zero due to the homogeneous structure of the artificial materials. Since
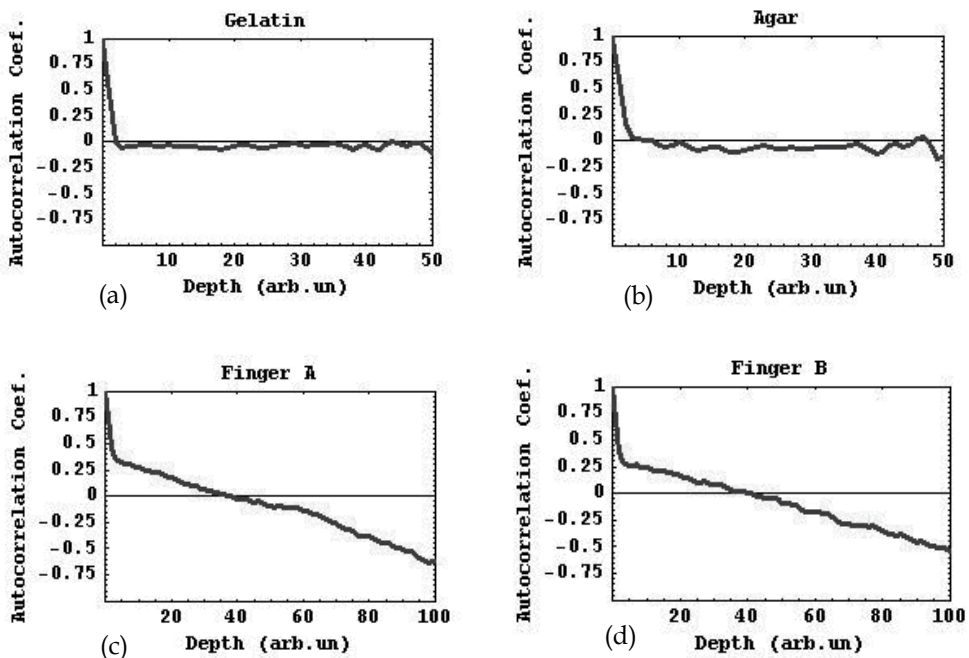


Fig. 13. Autocorrelation curves for artificial materials: (a) gelatin, (b) agar, (c, d) human fingers.

the fluctuations of the speckle intensity in the OCT images are random and homogeneous in these regions, the autocorrelation function values at each depth are around zero. Unlike the artificial material, human skin is a highly inhomogeneous tissue. When the autocorrelation analysis is applied to skin, the resulting autocorrelation function curves (Figures 13 (c) and (d)) decrease nearly monotonically with increasing depth and therefore differ significantly from the artificial materials. As a result, an autocorrelation analysis of OCT speckle-noise could potentially be used as a criterion for automatic and semi-automatic discrimination of artificial materials from real fingers.

### 3.3.2 Detection using full-field OCT systems

In section 3.2.2, the full-field OCT, FF-OCT, has been introduced. Because it has the capability of grabbing A-scans in parallel, FF-OCT system has much higher processing speed. In this sub session, we will describe detection of dummy fingerprint using FF-OCT systems.

As dummy fingerprints are normally made using translucent materials, the full-field OCT becomes a powerful tool for quickly and effectively distinguishing real fingers from artificial ones. FF-OCT can detect both surfaces of a dummy — the fingerprint surface as well as the non-print surface — and can probe the internal structures within them. The features observed using an FF-OCT system will differ from those for a real finger, as demonstrated in the set of cross-sectional images shown in Figure 14. Two obvious surfaces exhibiting different features were discovered during the depth scanning. The outer surface shows a smooth 2D curve (Figures 14(a)–14(f)), a feature that does not exist in a real fingerprint. However, the inner surface shows segmented fingerprints at different layers (Figures 14(g)–14(l)).
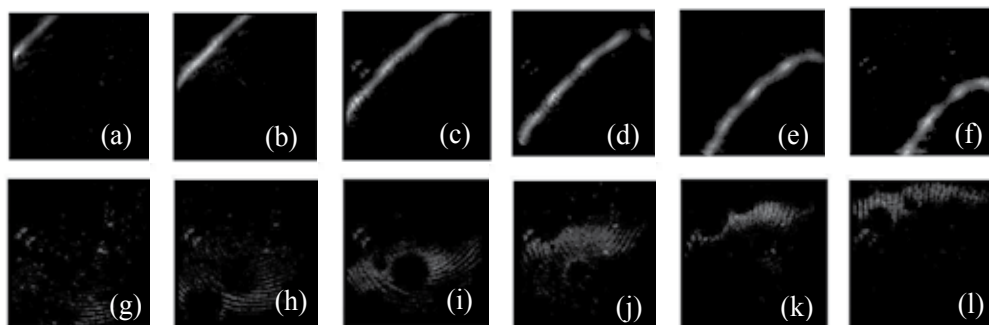


Fig. 14. OCT Images of a dummy fingerprint obtained by a FF-OCT system (a)-(f). OCT images extracted from the outer surfaces (layer distance: 50 μm). (g)-(l). OCT images extracted from the inner surfaces (layer distance: 20 μm).

The summation of Figures 14(a)–14(f) is shown in Figure 15(a), which reveals a bright area without any fingerprints in it. (The image sampling separation is set to 50 μm, so that the black fringes appear when these two images are overlapped). Figure 15(b) shows the summation of the segmented fingerprints from Figures 14(g)–14(l). Figure 15(c) gives the summation of all the tomographic images, wherein which the fingerprint image is completely destroyed The summation is very different from the image captured by a common 2D camera used in a fingerprint recognition system, as shown in Figure 15(d).

Figure 16 provides another set of rotated 3D volume data. The red parts show the internal structure of the dummy, which is totally different from the internal tissues in a real finger. The presence of the two surfaces and their distinct patterns reveal that the scanned object is a fingerprint dummy.



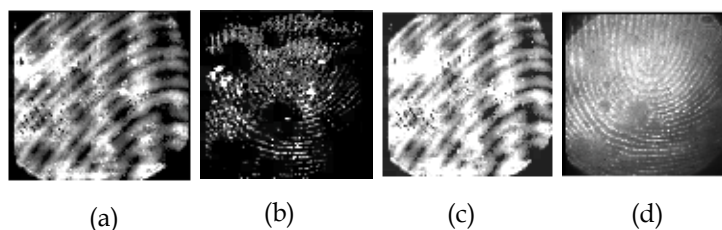(a)                    (b)                    (c)                    (d)

Fig. 15. Images of a dummy fingerprint. (a) OCT image of the outer surface of dummy. (b) OCT image of the inner surface of dummy. (c) Summation of above images (a) and (b). (d) Direct imaging of the dummy by a camera
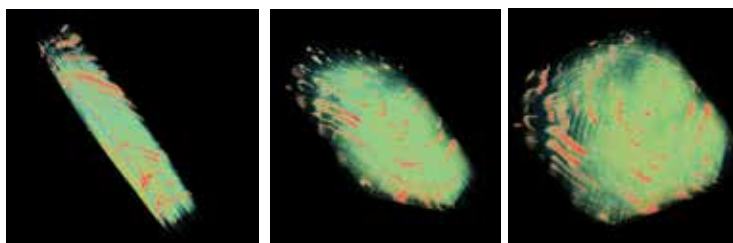


Fig. 16. Three rotated images of the 3D volume data of a dummy fingerprint.

In a traditional fingerprint reader, the finger must be pressed on a transparent flat surface in order to produce a 2D fingerprint pattern. There are some problems associated with these types of devices. Firstly, any motion of the finger may blur the imprinted image. Secondly, obtaining a clear image requires that the imprinting surface be cleaned for every new user, which adds complexity to the mechanical implementation. Another important issue lies in the fact that the 2D (flat) fingerprint pattern losses 3D-profile features that also provides information that can be used to uniquely identify an individual. The OCT-based fingerprint recognition system integrates all of the 2D morphologic features, along with the 3D profile and internal structure, which increases the ability of the system to robustly discriminate artificial fingerprints from real ones.

## 4. Spoof detection using spectral analysis

As previously described, fingerprint readers can be defeated using artificial (or prosthetic) fingers that can be created from cheap kitchen supplies or polymeric materials. These methods are commonly called "spoofing" because they are attempts to spoof the credentials of a valid user by presenting a fake fingerprint trait to the biometric sensor. Fingerprint spoofing uses simple techniques that can be quite effective (see Figure 17), so spoof detection is becoming increasingly important.

To address the threat of spoofing, we investigate the possibility of detecting artificial fingers by using spectrum analysis. In this section, we will first study the characteristics exhibited when human and fake fingers are exposed to different wavelengths of light. Based on the spectral images, we develop an algorithm to process the captured image, calculate the average image energy, extract the spectral features, and then distinguish the artificial fingers from the real ones.
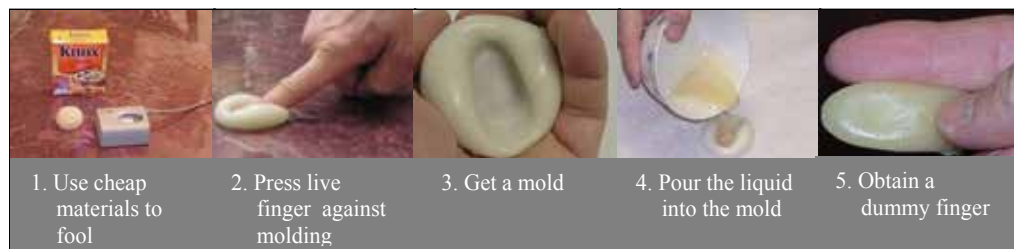


| 1. Use cheap materials to fool | 2. Press live finger against molding | 3. Get a mold | 4. Pour the liquid into the mold | 5. Obtain a dummy finger |

Fig. 17. Steps to make a dummy finger using a cheap kitchen powder

## 4.1 Differences between real finger and fake finger

A living human finger has rich blood vessels, sweat glands, and soft tissues under the skin. A cross-sectional image of a real finger extracted by an OCT system is shown in Figure 18. This image shows that the underlying morphology has layered tissues with some sweat glands. This complicated structure causes a strong scattering when light is shone on the surface. Because human tissue mostly consists of water, the absorption of the incident light is quite strong and varies according to the wavelength applied.
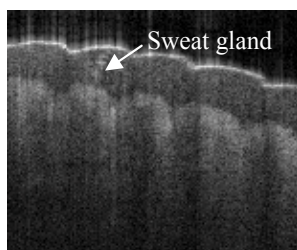


Fig. 18. OCT image of a real finger

The dummy or prosthetic finger is made by casting a real human finger with silicon or polymer material, and then painting the skin color, hair, even a nail on its surface. (Figure 19).



Fig. 19. The real finger (bottom) and the prosthetic finger (top) made from it.
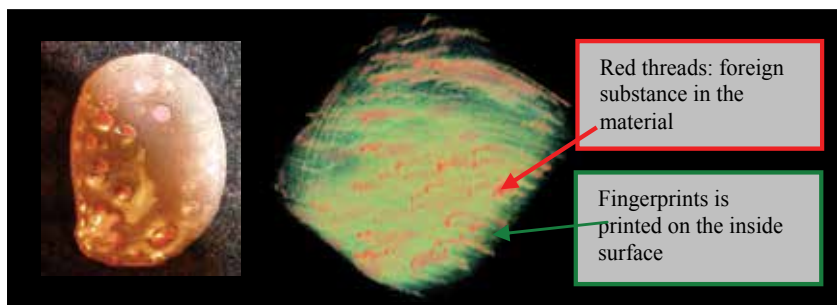
Fig. 20. A dummy fingerprint and its 3D OCT volume structure

Since no internal bio-structure information exists within the prosthetic finger, there must be a significant difference in the optical properties between it and a real finger. Figure 20 illustrates the 3D volume structure (right) of a dummy fingerprint casting from a person (left). This 3D volume data was obtained by the FF-OCT system described in the previous section. None of the red threads inside the volume exist in the real human finger. Comparing to the layered structure shown in Figure 18, the optical properties of the reflected light from dummy finger should be quite different.

### 4.2 Finger image spectrum analysis

A fingerprint imaging system was constructed to explore the spectral features of the real and fake fingers, and is shown schematically in Figure 21. The imaging system has the following specifications:

- Light source: Broadband white light with a mercury arc lamp. An attenuator is used to change the intensity of the light.
- Filters: Nine wavelength filters from near ultraviolet through visible light to near infrared (400nm, 450nm, 500nm, 550nm, 600nm, 650nm, 700nm, 800nm, 850nm).
- Camera: Dalsa 1M15 CCD, which covers the visual band (with the original lens) and the near infrared band (with the original lens removed).
- Imaging: To avoid the sensitivity to oil or dirt contaminates, there was no glass used in front of the finger during the fingerprint acquisition.
- A computer was used for image acquisition and processing.

Each test's finger was placed on the finger holder, and nine images were grabbed for each person or sample. The output intensity of the light source, and the attenuator in front of the camera, was adjusted in order to get the best image quality. Six human fingers and one prosthetic were tested in this project. They are three Asian adults (two male and one female); two white adults (one male and one female); one black adult male; and one prosthetic finger (shown in Figure 22) reproduced from one of the Asian male subjects.

The procedures involved in the spectrum analysis are illustrated in the flowchart of Figure 23. The input is a 1024x1024 pixel image with 16-bit grey scale. A nonlinear median filter is used to preserve edges while removing noise. A normalization process is carried out to reduce the effects of differences in illumination, skin reflection and ambient light. All the values are normalized between 0 and 255. A banded image is created to collect the energy reflected from the central area of the finger (Figure 24). In the experiments, band parameters are X direction 300 pixels and Y direction 300 pixels centered in the image. Average energy (AE) is calculated by averaging the pixel values over the banded image area. The Canny

edge detection algorithm is implemented to detect the edges with a scaling factor of 25%. The detected edges are used to generate a mask. Edge energy (EE) is measured by averaging the image values under the mask. A decision function is constructed based-on the average of image energies AE and EE. The output status is "true" if the input is detected as a real finger and "false" otherwise.
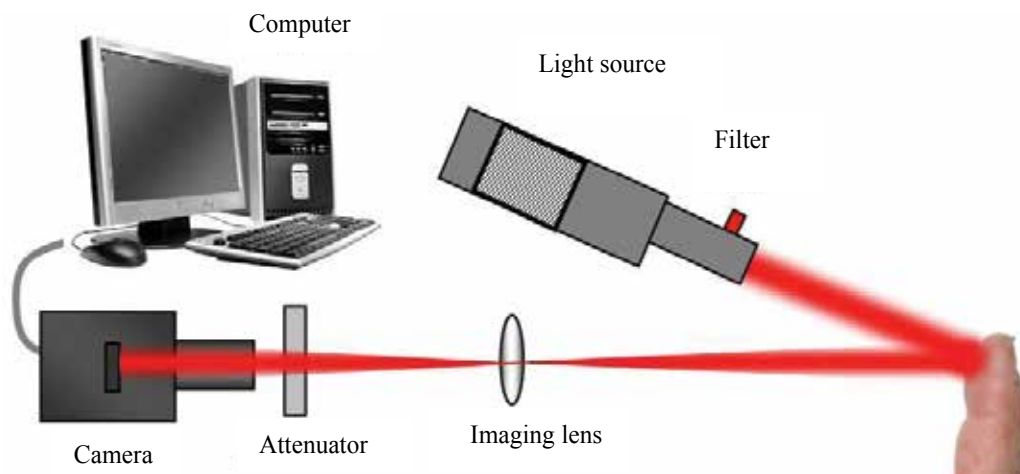


Fig. 21. Fingerprint imaging system configuration.



Fig. 22. Fake finger used in experiments.

Figure 25 shows finger images and the corresponding edge images for a real finger and its equivalent artificial finger. From those edge images, we observed that: 1) At short wavelengths, particularly at 400 nm (near ultraviolet), both the real and artificial fingers show more edge details than those extracted using longer wavelengths. 2) At longer wavelengths, particularly at 850 nm (near infrared), the artificial finger image becomes blurred. Nevertheless, we observe that the artificial finger is much brighter at longer wavelengths, especially at 850nm, even though the images were taken under the same conditions and were normalized using the same algorithm.

These observations can be explained by differences in the optical properties between human skin and materials, normally polymers, used in the fabricated artificial fingers. The *in vivo*

absorption coefficient of human skin is about 70% of the absorption coefficient of water, and the isotropic scattering coefficient ranges from 3 to 16 cm$^{-1}$, much stronger than that of polymers. Because human tissue absorbs more light than polymers at wavelength greater than 700 nm, there is more back-reflected and scattered light collected by the camera for the artificial finger; this means that the artificial finger image contains more energy, or optical power. At the same time, since the back-reflecting and scattering light from the deeper penetration lacks the additional structural information found in human skin, it overwhelms the ridge information reflected from the surface. This lack of internal structure is the reason why the fake finger images appear more blurred. The blurring effect is measured by the average energy of the edge image of a finger: the fewer edges detected in an edge image, the lower the average edge energy the image contains.
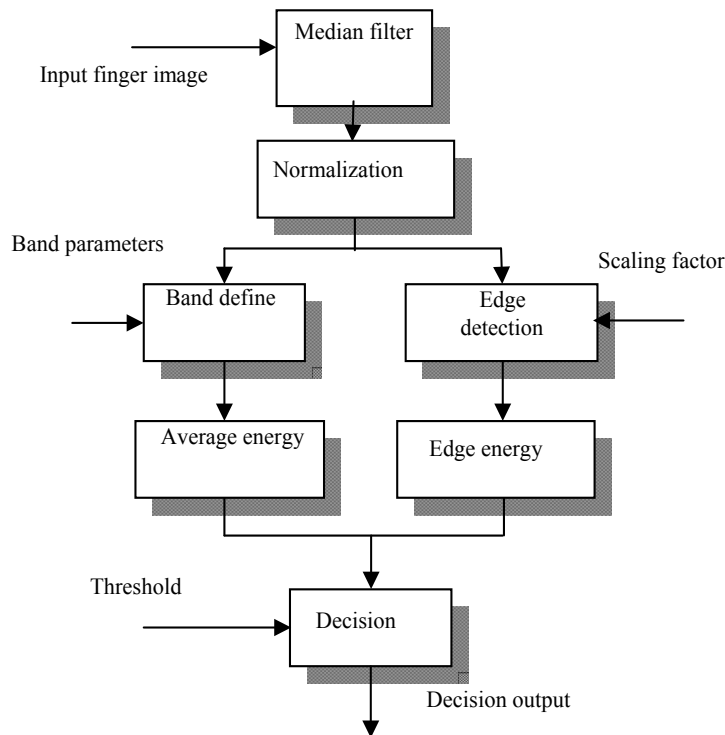


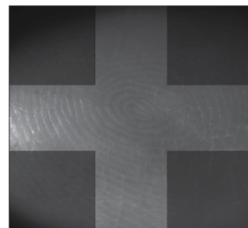Fig. 23. Image processing flowchart



Fig. 24. Banded image of a real finger at 550nm wavelength with a band width of 300 pixels.
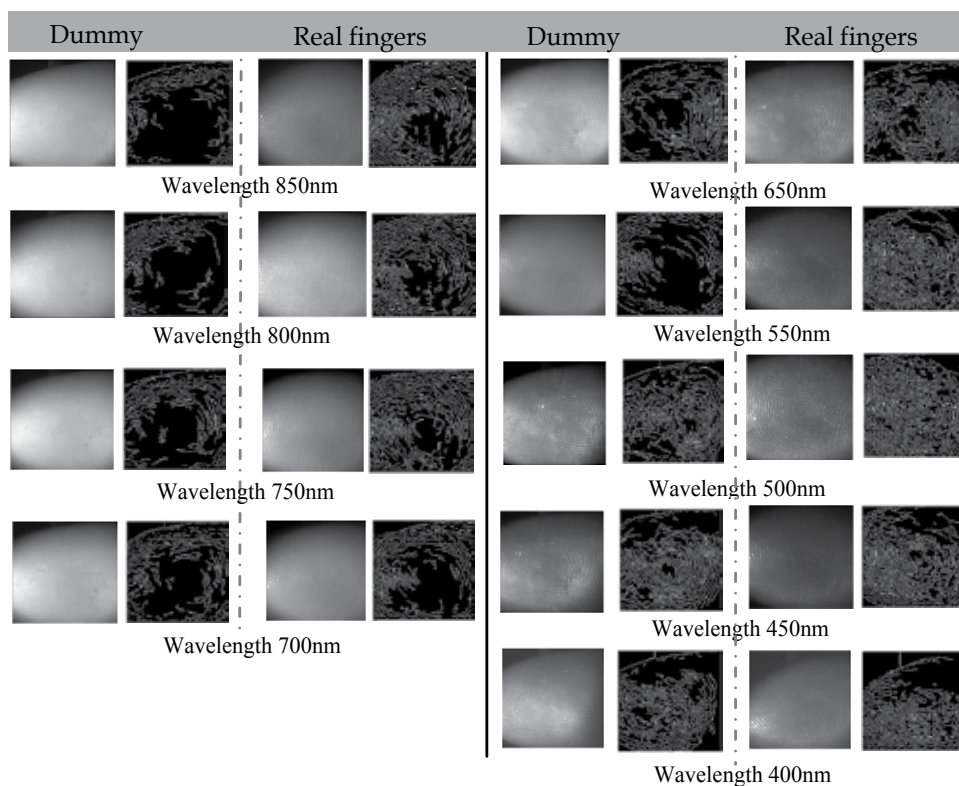
Fig. 25. Finger and finger edge images from dummy and real fingers.

Given the analysis above, the image energy and edge information at NIR wavelengths, (e.g. 850nm), can be used for discriminating between fake and real fingers. Considering that the energy is mostly distributed in the central area of the finger, it is advantageous to setup a banded area to collect the energy as shown in Figure 24.

Table 2 lists the values of average energy (AE) extracted from six fingers at nine wavelengths. Table 3 shows all the values of edge energy (EE) extracted from the same six

| Fake | 3.27 | 3.28 | 3.53 | 3.42 | 4.69 | 4.76 | 4.95 | 4.76 | 4.71 |
|------|------|------|------|------|------|------|------|------|------|
| RM1 | 2.47 | 2.19 | 3.19 | 2.56 | 3.59 | 3.76 | 4.00 | 4.59 | 3.32 |
| RM2 | 2.65 | 2.98 | 2.37 | 2.33 | 3.70 | 3.27 | 3.67 | 1.99 | 3.95 |
| RM3 | 3.81 | 3.14 | 3.89 | 3.47 | 4.14 | 4.41 | 3.95 | 3.95 | 4.24 |
| RM4 | 2.35 | 2.83 | 4.01 | 3.74 | 5.03 | 4.58 | 4.71 | 4.78 | 3.76 |
| RF1 | 4.30 | 4.11 | 4.36 | 4.25 | 4.77 | 4.83 | 4.74 | 3.44 | 3.86 |
| RF2 | 3.08 | 2.65 | 2.64 | 2.59 | 2.78 | 3.04 | 2.82 | 3.14 | 3.22 |
| $\lambda$(nm) | 400 | 450 | 500 | 550 | 650 | 700 | 750 | 800 | 850 |

- Fake: fake finger, modeled from real male finger, RM1.  • RM1: real male finger #1. Asian male.
- RM2: real male finger #2. Asian male.  • RM3: real male finger #3. White male.
- RM4: real male finger #4. Black male. • RF1: real female finger #1. Asian female.
- RF2: real female finger #2. White female.

Table 2. Average energies (AE values x10$^4$ ) from finger images.

fingers at nine wavelengths. These data show that at the longer wavelengths, particularly at 850 nm, the energy in the fake-finger image is higher than those of real fingers. However, at the shorter wavelengths, there is no clear difference between fake and real ones.

| Fake | 10.00 | 12.71 | 12.39 | 12.68 | 9.61 | 7.05 | 6.96 | 5.81 | 5.19 |
|------|-------|-------|-------|-------|------|------|------|------|------|
| RM1 | 10.50 | 13.31 | 17.02 | 14.83 | 7.24 | 8.43 | 11.74 | 11.53 | 12.11 |
| RM2 | 9.41 | 13.43 | 13.40 | 13.88 | 8.93 | 9.52 | 9.42 | 7.66 | 10.90 |
| RM3 | 13.75 | 13.43 | 14.90 | 14.92 | 9.61 | 9.33 | 9.16 | 10.04 | 10.74 |
| RM4 | 9.51 | 14.62 | 14.84 | 14.57 | 13.47 | 9.26 | 8.84 | 11.50 | 12.20 |
| RF1 | 12.35 | 11.90 | 11.99 | 11.83 | 9.45 | 11.08 | 10.93 | 10.48 | 10.60 |
| RF2 | 14.49 | 12.80 | 12.64 | 12.57 | 7.30 | 7.76 | 5.95 | 7.13 | 8.08 |
| $\lambda$(nm) | 400 | 450 | 500 | 550 | 650 | 700 | 750 | 800 | 850 |

Table 3. Edge energies (EE values x10$^{-2}$) from finger edge images.

### 4.3 Distinguishing artificial fingers from real ones

From the analysis in last section, we see that the ability to discriminate real fingers from fake ones is proportional to the average energy of the finger image, particularly in the longer wavelength band. However, the ability to discriminate is inversely proportional to the average energy of the finger-edge image, again particularly in the longer wavelength band.
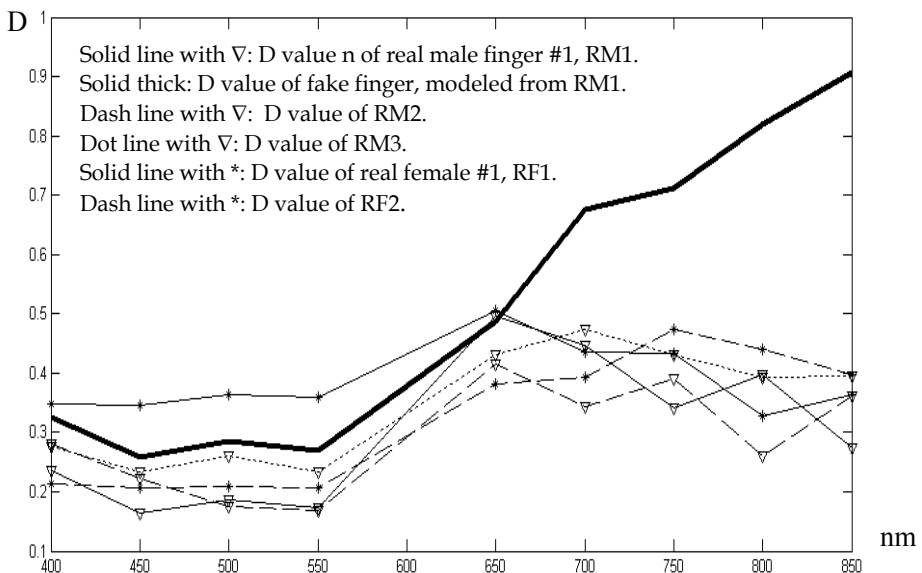


Fig. 26. Decision values for all finger images

Therefore, the decision function for distinguishing artificial fingers from real ones can be constructed using the following equations:

$$F = \begin{cases} 1 & \text{Fake, if } D >= T \\ 0 & \text{Real, if } D < T \end{cases} \tag{11}$$

where D is the discrimination value, given by:

$$D = \quad AE / EE \tag{12}$$

The D values calculated from Table 3 are plotted in Figure 26. From our experiments, we can clearly distinguish the artificial fingers from all the real fingers by setting a threshold value of T = 0.55 at a wavelength of 800nm or 850 nm.

Although the number of testing samples is relatively small, they represent different races. In fact, such a multiple wavelengths database obtained from same living fingers doe not exist at the time being. However, the experimental result clearly show a trend, which demonstrates that as the wavelengths become longer, the fake one gradually separate itself from other real ones, which agrees with the analysis described in earlier sessions.

## 5. Summary and conclusions

In this chapter, we have described several approaches to fingerprint anti-spoofing by means of internal biometrics. Two different methods using NIR optical imaging technology to detect a fake finger are introduced and discussed: 1) optical coherence tomography and, 2) fingerprint NIR image analysis.

Optical coherence tomography provides a powerful new tool for applications in security and document identification. By extending the existing biometric techniques that are based on surface scans of external features, the OCT system can probe and extract the internal features of multilayered objects and tissues. This will be more robust against tampering and counterfeiting. We have demonstrated that OCT techniques can be successfully applied for detecting artificial materials that are commonly used to make fraudulent fingerprints. Overall, the results demonstrate that: 1) Current commercial fingerprint systems have security vulnerabilities and can easily be spoofed by fingerprint dummies; 2) High-resolution 2D OCT images and the corresponding signal curves can clearly reveal the artificial materials; 3) OCT is capable of providing high-resolution 3D images for security identification reference.

Our future research will be focused on increasing the OCT 3D-image acquisition rate, and on applying the current pattern recognition method in processing OCT images (2D and 3D). This will allow systems to better distinguish the artificial fingerprint layer in order to resist spoofing attacks, thereby enhancing the security provided by these applications. However, the cost of such an OCT system is relatively high: the primary component is the broadband IR light source, which can cost thousands of dollars.

In this chapter, we have also described a relatively cheap and practical approach for distinguishing dummy fingers from real ones. The proposed method is based on using the observed spectral features of the sample fingers; experimental testing has shown that near-infrared light can be used to successfully detect differences in the optical properties between real fingers and an artificial one. The artificial finger exhibits back-reflected and scattered light from the deeper structures at NIR wavelengths, whereas real *in vivo* human tissue can absorb more light. The strong back-scattered light from fake finger washes out the surface structure information, thereby blurring the fingerprint ridges.

A simple algorithm based on overall energy and edge energy can be used to identify real fingers from fake ones. The ability to discriminate may improve when using longer wavelengths (e.g., >800 nm), although the use of IR camera will be very expensive.  As most of the cameras using the visual band provide some sensing capability at 850 nm, these can be used to create a cost-efficient fingerprint recognition system. Although we used a mercury arc lamp as the light source in the proposed setup, a very cheap but powerful halogen bulb can also serve the same purpose.

This novel method for discriminating between real and dummy fingers is simple, low-cost, and effective. Such a system can be fabricated as a stand-alone detection device, or it can be easily integrated into an existing fingerprint recognition system for the purpose of pre-screening fingers to defeat spoofing.

## 6. References

Akiba M, & Chan KP. (2007). *In vivo* video-rate cellular-level full-field optical coherence tomography. *J Biomed Opt* 12:064024.

Bashkansky M, Lewis D, Pujari V, Reintjes J, & Yu HY. (2001). Subsurface detection and characterization of Hertzian cracks in Si3N4 balls using optical coherence tomography. *Ndt & E International* 34:547-55.

Chang S, Cheng Y, Larin KV, Mao Y, Sherif S, & Flueraru C. (2008). Optical coherence tomography used for security and fingerprint-sensing applications. *IET Image Processing* 2:48-58.

Chang S, Liu XP, Cai XY, & Grover CP. (2005). Full-field optical coherence tomography and its application to multiple-layer 2D information retrieving. *Optics Communications* 246:579-85.

Chang S, Cai XY, & Flueraru C. (2006a). Image enhancement for multilayer information retrieval by using full-field optical coherence tomography. *Applied Optics* 45:5967-75.

Chang S, Mao YX, Sherif S, & Flueraru C. (2006b). Full-field optical coherence tomography used for security and document identity - art. no. 64020Q. *Optics and Photonics for Counterterrorism and Crime Fighting II* 6402:Q4020-Q.

Cheng Y, & Larin KV. (2006). Artificial fingerprint recognition using optical coherence tomography with autocorrelation analysis. *Applied Optics* 45 (cover paper):9238-45.

Cheng Y, & Larin KV. (2007). *In vivo* Two- and Three-Dimensional Imaging of Artificial and Real Fingerprints with Optical Coherence Tomography. *Photonics Technology Letters* 19:1634-6.

Chinn SR, & Swanson EA. (1996). Multilayer optical storage by low-coherence reflectometry. *Optics Letters* 21:899-901.

Dubois A, Vabre L, Boccara AC, & Beaurepaire E. (2002). High-resolution full-field optical coherence tomography with a Linnik microscope. *Applied Optics* 41:805-12.

Dunkers JP, Parnas RS, Zimba CG, Peterson RC, Flynn KM, Fujimoto JG, & Bouma BE. (1999). Optical coherence tomography of glass reinforced polymer composites. *Composites Part a-Applied Science and Manufacturing* 30:139-45.

Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, & Fujimoto JG. (1991). Optical Coherence Tomography. *Science* 254:1178-81.

Liang H, Cid M, Cucu R, Dobre G, Podoleanu A, Pedro J, & Saunders D. (2005). En-face optical coherence tomography - a novel application of non-invasive imaging to art conservation. *Optics Express* 13:6133-44.

Manapuram RK, Ghosn M, & Larin KV. (2006). Identification of Artificial Fingerprints Using Optical Coherence Tomography Technique *Asian Journal of Physics* 15:15-27.

Nanni L, & Lumini A. (2009) Descriptors for image-based fingerprint matchers, *Expert Systems with Applications*, 36(10):12414-22.

Netter FH. (1997). *Atlas of Human Anatomy*. Novartis, East Hanover

Nixon KA, Aimale V, & Rowe RK. (2008). Spoof Detection Schemes. In *Handbook of Biometrics*: Springer Science + Business Media, L.L.C.

Rie ER. (1987). The influence of varnishes on the appearance of paintings. *Studies on Conservation* 32:1–13.

Smolka G. (2008). Optical Coherence Tomography: Technology, Markets, and Applications 2008-2012. *BioOptics World*

Spring m, Liang H, Peric B, Saunders D, & Podoleanu A. (2008). Optical coherence tomography—a tool for high resolution non-invasive 3D-imaging of the subsurface structure of paintings. In *ICOM Committee For Conservation Newsletter*, pp. 633-40

Szkulmowska A, Gora M, & Targowska M. (2005) of Conference. The applicability of optical coherence tomography at 1.55 um to the examination of oil paintings. *6th International Congress on Lasers in the Conservation of Artworks (LACONA VI '05)*. Vienna, Austria

Szkulmowskaki A, Gora M, Targowska M, Rouba B, Stifter D, & Targow EBP. (2007). *Lasers in the Conservation of Artworks*. New York: Springer Science + Business Media

Targowski P, Gora M, & Wojtkowski M. (2006). Coherence Tomography for Artwork Diagnostics. *Laser Chemistry*

Targowski P, Rouba B, Wojtkowski M, & Kowalczyk A. (2004). The application of optical coherence tomography to nondestructive examination of museum objects. *Studies on Conservation* 49:107-14.

Tuchin VV. (2000). *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*. Bellingham, WA: SPIE

Tuchin VV. (2007). *Tissue optics : light scattering methods and instruments for medical diagnosis*. Bellingham, Washington: SPIE Press

Tuchin VV, Maksimova IL, Zimnyakov DA, Kon IL, Mavlyutov AH, & Mishin AA. (1997). Light propagation in tissues with controlled optical properties. *J Biomed Opt* 2:401-17.

Tuchin VV, Xu XQ, & Wang RK. (2002). Dynamic optical coherence tomography in studies of optical clearing, sedimentation, and aggregation of immersed blood. *Applied Optics* 41:258.

Wiesauer K, Pircher M, Gotzinger E, Bauer S, Engelke R, Ahrens G, Grutzner G, Hitzenberger CK, & Stifter D. (2005). En-face scanning optical coherence tomography with ultra-high resolution for material investigation. *Optics Express* 13:1015-24.

Xiao Q, & Raffat H. (1990). Combining statistical and structural information for fingerprint image processing, classification and identification. In *Pattern recognition: architectures, algorithms and applications*, ed. R Plamondon, H Cheng. Singapore: World Scientific Publishing Co.

Yan P. & Bowyer, K. W. Biometric Recognition Using Three-Dimensional Ear Shape. http://www.cse.nd.edu/Reports/2006/TR-2006-01.pdf

Yang H, Xie, S., Li, H., and Lu, Z. (2007). Determination of human skin optical properties in vivo from reflectance spectroscopic measurements. *Chinese Optics Letters* 5:181-3.

Yang JC, & Park DS. (2008). A fingerprint verification algorithm using tessellated invariant moment features, *Neurocomputing* 71:1939–46.

Zimnyakov DA, Agafonov DN, Sviridov AP, Omel'chenko AI, Kuznetsova LV, & Bagratashvili VN. (2002). Speckle-contrast monitoring of tissue thermal modification. *Appl Opt* 41:5989-96.

Zimnyakov DA, Ryabukho VP, & Larin KV. (1994). Microlens Effect Due to the Diffraction of Focused Beams on Large-Scale Phase Screens. *Letters to Journal of Theoretical Physics* 20:14.

# Optical Spatial-Frequency Correlation System for Fingerprint Recognition

Hiroyuki Yoshimura
*Graduate School of Engineering, Chiba University*
*Japan*

## 1. Introduction

Individual recognition systems with high-speed and high-accuracy have been recently demanded in the automatic logging into a PC, the immigration at the airport, the access control and diligence & indolence management in an office, and so on. Biometric authentication is now being regarded as the most valid method because of the receptivity, individuality and invariability of biometric identifiers. Various types of the individual recognition systems based on biometrics have been studied and partially realized. Fingerprints, faces, hand geometry, irises, vein patterns, gait, signatures, etc., are known as biometric identifiers. In particular, the fingerprint recognition system has been widely used because of its high reliability and reasonable price (Maltoni et al., 2003a; Jain et al., 2010). The fingerprint recognition methods can be classified into the following three types: (i) the minutiae-based, (ii) the frequency-based and (iii) the image-based methods.

The minutiae-based method is mainly being used in the practical fingerprint recognition system. For example, a memetic fingerprint matching algorithm has been recently proposed to identify the optimal or near optimal global matching between two minutiae sets (Sheng et al., 2007). A matching technique for fingerprint recognition using the Minutia Cylinder-Code (MCC), which is based on 3D data structures built from minutiae distances and angles, has also been made a proposal (Cappelli et al., 2010) to exclude the drawbacks in the fingerprint authentication using local minutiae structures. Moreover, a fingerprint verification using spectral minutiae representations has been suggested to overcome translation, rotation and scaling which are the drawbacks of minutiae-based algorithms (Xu et al., 2009a, 2009b).

The frequency-based method, such as the frequency analysis method, is also being used in the practical fingerprint recognition system in order to secretly hide the original fingerprint information (Takeuchi et al., 2007). Recently, a fingerprint recognition method based on mel-frequency cepstral coefficients and polynomial shape coefficients has been proposed because of the robustness to noise and the insensivity to translation (Hashad et al., 2010).

The image-based method has been being studied to improve the accuracy in the fingerprint recognition. For example, an enhanced image-based algorithm for fingerprint verification based on invariant moment features has been recently proposed to improve matching accuracy and processing speed (Yang & Park, 2008a, 2008b). A novel image-based fingerprint matcher based on the minutiae alignment has also been made a proposal to improve the verification performance (Nanni & Lumini, 2009).

The correlation-based method, which can be classified into the image-based method, has also been being studied on the background that the improvement of accuracy in the fingerprint recognition system is demanded, though there are demerits of suffering from displacement and rotation of a fingerprint in the authentication process. For example, recently, a correlation-based fingerprint matching with orientation field alignment has been proposed to reduce the processing time (Lindoso et al., 2007). Previously, the joint transform correlator (Goodman, 1996) was applied to the fingerprint recognition system and the fingerprint recognition optical system based on the joint transform correlator was produced experimentally (Kobayashi & Toyoda, 1999). However, there were demerits that the optical system needed the reproduction of intensity distribution by use of a CCD camera, a liquid crystal spatial light modulator and a PC so that the speed of the authentication was strongly dependent on the speed of these electronic devices and optical components. Therefore, the merit of light was not fully taken advantage of in the optical system. In addition, the size of the optical system became large because the optical system was complicated.

In this chapter, we describe our proposed optical information processing system for biometric authentication using the spatial-frequency correlation of subject's and enrolled biometric identifiers. We call it the optical spatial-frequency correlation (OSC) system for the biometric authentication (Yoshimura & Takeishi, 2009). The merit is that high-speed authentication would be possible because of all optical system. In addition, our OSC system is very simple so that it could be composed in small size. Our OSC system could be classified into a combination of the correlation-based and the frequency-based methods.

First, we introduce the idea of the OSC system especially for the fingerprint recognition. Next, we analyze the basic properties of the OSC system by use of a modeled fingerprint image of which the grayscale in a transverse line is the 1D finite rectangular wave with a period of 0.5mm and the whole width of the fingertip of 15mm. Concretely, the effects of (i) transformation of the subject's fingerprint, such as variation of positions of ridges, and (ii) random noise, such as sweat, sebum and dust, etc., superimposed on the subject's fingerprint on the fingerprint recognition in the OSC system are analyzed. Furthermore, we investigate the recognition accuracy of the OSC system by use of real fingerprint images on the basis of the false acceptance rate (FAR), the false rejection rate (FRR) and the minimum error rate (MER). Finally, we conclude our chapter.

The following sections consist of 2) The OSC system; 3) Basic properties and recognition accuracy of the OSC system; 4) Conclusions.

## 2. The OSC system

The spatial-frequency correlation function (SCF) between the subject's and enrolled fingerprints can be obtained by the optical system shown in Fig. 1. In the figure, $f$ stands for the focal length of the lens. $P_1$ denotes the input plane with the coordinate system of $x_1$ and $y_1$ and $P_2$ does the output plane with the coordinate system of $x_2$ and $y_2$. The subject's fingerprint image $g(x_1,y_1)$ and the enrolled fingerprint image $h(x_1,y_1)$ are superimposed, placed in the $P_1$, and illuminated by the plane wave radiated from a laser. Then, the optical field $U_1(x_1,y_1)$ in the $P_1$ is given by

$$U_1(x_1,y_1)=g(x_1,y_1)h(x_1,y_1)=g(x_1,y_1)h^*(x_1,y_1), \tag{1}$$

where $^*$ stands for the complex conjugate. In Eq. (1), $h=h^*$, because we consider $h$ as a real function such as a fingerprint image. The optical field $U_2(x_2,y_2)$ in $P_2$ is obtained by the Fourier transform of $U_1(x_1,y_1)$ and given by

$$U_2(x_2,y_2) = \frac{1}{\lambda f} G\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \otimes H^*\left(-\frac{x_2}{\lambda f}, -\frac{y_2}{\lambda f}\right), \qquad (2)$$

where $\otimes$ and $\lambda$ stand for the convolution and the wavelength of a laser light, respectively. $G$ and $H$ denote the Fourier transforms of $g$ and $h$, respectively. We can find that Eq. (2) expresses the SCF between $g$ and $h$. In the following section, we analyze the basic properties of our OSC system and investigate whether our proposed system is valid for the fingerprint recognition or not. In the investigation, the intensity distribution of the SCF is used because only the intensity distribution in the output plane $P_2$ could be obtained by the optical detector like a CCD camera.
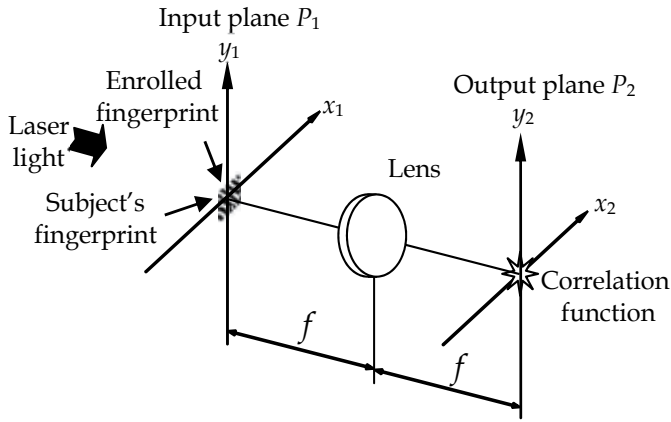


Fig. 1. The OSC system.

## 3. Basic properties and recognition accuracy of the OSC system

In this section, first, the FAR, FRR and MER which are related to the accuracy of the fingerprint recognition system are introduced in subsection 3.1. Next, the basic properties of the OSC system are investigated using a modeled fingerprint image in subsection 3.2. Finally, the recognition accuracy of the OSC system is investigated using real fingerprint images in subsection 3.3.

### 3.1 FAR, FRR and MER

Fig. 2 illustrates the basic concept of the FAR and FRR. In the figure, the left-side red curve is the impostor distribution and the right-side blue curve is the genuine distribution. The longitudinal axis denotes the probability density funcion (PDF).

The FAR is the probability of accepting impostors erroneously. As shown in the figure, it corresponds to an area of the impostor distribution higher than the authentication threshold. On the other hand, the FRR is the probability of rejecting authentic person and corresponds to the area of the genuine distribution lower than the authentication threshold.

As an example, the authentication threshold is decided by a value satisfied with the condition that the FAR and FRR take the same value. It is called the MER. However, in general, the authentication threshold is shifted toward the right side in order to reduce the value of the FAR, though the value of the FRR increases. In our analysis, the horizontal axis in Fig. 2 corresponds to the peak value of the normalized intensity distribution of the SCF between the two fingerprint images. In the following figures, it is simply written as "peak value of SCF".



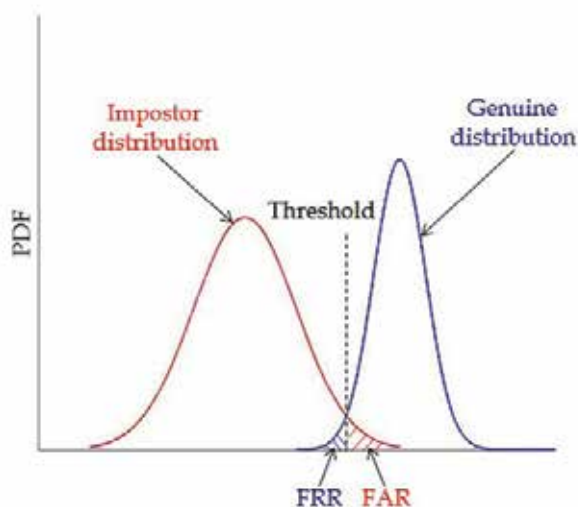Fig. 2. Basic concept of the FAR and FRR. The MER can be obtained under the condition that FAR=FRR.

### 3.2 Basic properties of the OSC system for a modeled fingerprint image

In this subsection, the basic properties of the OSC system are analyzed using a modeled fingerprint image. First, in subsection 3.2.1, the modeled fingerprint image is introduced and its spatial-frequency autocorrelation function, i.e., the spatial-frequency correlation of the genuine fingerprint of his or her own, is shown. Next, in sebsection 3.2.2, the SCF between the modeled fingerprint image and the modified one, i.e., the spatial-frequency correlation between the genuine and impostor fingerprints, is shown. Moreover, in subsection 3.2.3, the SCF between the modeled fingerprint images with and without random noise, is shown. Finally, in subsection 3.2.4, the recognition accuracy of the OSC system for the modeled fingerprint images is indicated.

### 3.2.1 Modeled fingerprint image and its spatial-frequency autocorrelation function

Fig. 3 illustrates an example of the fingerprint image used in the FVC2002 (Maltoni & Maio, 2002; Maltoni et al., 2003b). FVC2002 denotes the abbreviation for the Fingerprint Verification Contest held in 2002. This fingerprint image consists of the tiff form with 374 pixels in height and 388 pixels in width, and the black and white in the image was reversed. In general, the grayscale distributions which correspond to the waveforms of the cross-

sections of the fingerprint are different from each other in the transverse lines of the fingerprint image.

We regard the left side of Fig. 4 as the modeled fingerprint image in order to evaluate the basic properties of our proposed system. The normalized grayscale distribution in the transverse line of the modeled fingerprint image is expressed in terms of the 1D finite rectangular wave shown in the right side of Fig.4. The period of ridges is 0.5mm and the whole width of the fingertip is 15mm. The normalized grayscale distribution is intentionally composed of 2048 ($2^{11}$) pixels in order to obtain the correct results of the Fourier Transform. Concretely, the ridge and valley in the distribution are composed of 25 pixels, respectively. The interval of neighboring pixels is 0.01mm.

Now we consider the case that the normalized grayscale distributions of subject's and enrolled fingerprint images are the same as the 1D finite rectangular wave shown in the right side of Fig. 4. This case corresponds to the recognition of his or her own. Then, the spatial-frequency autocorrelation function between the subject's and enrolled fingerprint images can be obtained in the output plane $P_2$ in Fig. 1. We derived it numerically under the conditions that $\lambda = 0.6328 \times 10^{-3}$mm and $f$=100mm. The obtained intensity distribution of the spatial-frequency autocorrelation function was normalized by its maximum value. The normalized intensity distribution is shown in Fig. 5. It has a sharp peak at the center of the distribution and takes a value of 1. In general, the peak value denotes the degree of spatial-frequency correlation and takes a value with a range from 0 to 1. The large value means high spatial-frequency correlation and the small one does low spatial-frequency correlation. In addition, in this figure, the second maximum value is 0.404 located at $x_2=\pm0.127$mm which is related strongly to a period of the normalized grayscale distribution of the modeled fingerprint image, i.e., $d$ =0.5mm and obtained by $\pm \lambda f/d$.

In the following analyses, we evaluate the behavior of the peak value of the normalized intensity distribution of the SCF between the subject's and enrolled fingerprint images. Moreover, we investigate whether our proposed optical system is valid for the fingerprint recognition or not.



Fig. 3. Example of the fingerprint image used in the FVC2002. The black and white in the image was reversed.
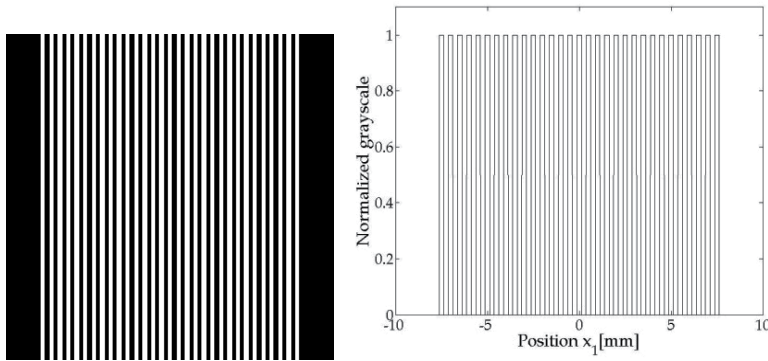
Fig. 4. Modeled fingerprint image (left) and the normalized grayscale distribution in a transverse line of the image (right). The period is 0.5mm and the whole width of the fingertip is 15mm.
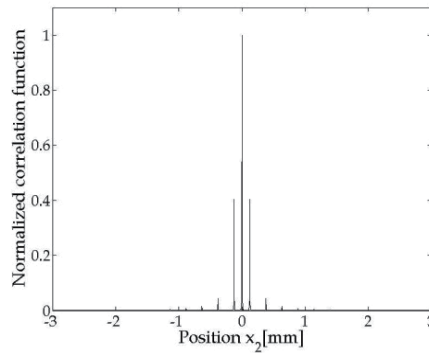


Fig. 5. Normalized intensity distribution of the spatial-frequency autocorrelation function of the 1D finite rectangular wave shown in the right side of Fig. 4. The second maximum value is 0.404 located at $x_2=\pm0.127$mm.

### 3.2.2 SCF between the modeled fingerprint image and the modified one

In the previous subsection, the normalized grayscale distributions of the subject's and enrolled fingerprint images were the same one which was regarded as the 1D finite rectangular wave shown in the right side of Fig. 4. In this subsection, in order to investigate the SCF in the case that the subject's and enrolled fingerprint images are different from each other, the modified modeled fingerprint images, i.e., the modified finite rectangular waves, were used. The modified ones were produced by changing the positions of the ridges randomly from the regular positions of the ridges in the original finite rectangular wave. Concretely, the positions of ridges were changed obeying a Gaussian random statistics with zero mean. Moreover, the standard deviation of the variation of the positions of the ridges was normalized by a period of ridges of the original finite rectangular wave, 0.5mm. We call it the normalized standard deviation of the positions of ridges, expressed in terms of $\sigma_{pn}$. $\sigma_{pn}$ indicates the difference between the original and modified finite rectangular waves quantitatively.

Fig. 6 shows several examples of the normalized grayscale distributions of the modified modeled fingerprint images. Figs. 6(a), 6(b) and 6(c) correspond to the cases when the normalized standard deviations $\sigma_{pn}$ are 0.05, 0.1 and 0.2, respectively. Fig. 7 shows the normalized intensity distributions of the SCFs between the original finite rectangular wave shown in the right side of Fig. 4 and the modified ones shown in Figs. 6(a), 6(b) and 6(c). Concretely, Figs. 7(a), 7(b) and 7(c) are the results obtained using Figs. 4 and 6(a), Figs. 4 and 6(b) and Figs. 4 and 6(c), respectively. The obtained intensity distributions of the SCFs were normalized by the square root of the product of the peak value of the spatial-frequency autocorrelation function of the original finite rectangular wave and the one of the spatial-frequency autocorrelation function of the modified one. The peak values in Figs. 7(a), 7(b) and 7(c) are 0.832, 0.648 and 0.489, respectively. This result indicates the fact that the spatial-frequency correlation between the two fingerprint images gradually becomes low as the difference between the two becomes large.

Next, in order to investigate the behavior of the peak value of the normalized intensity distribution of the SCF, 1000 kinds of the modified modeled fingerprint images were used for each value of $\sigma_{pn}$. Fig. 8 indicates the dependence of the peak value of the normalized intensity distribution of the SCF on the normalized standard deviation of the positions of ridges of the modified finite rectangular wave, $\sigma_{pn}$. The symbol of circle denotes the averaged peak value of the normalized intensity distribution of the SCF and the error bar does the standard deviation of the peak values. As shown in the figure, the averaged peak values when $\sigma_{pn}$=0.05, 0.1, 0.2 and 0.3 are 0.789, 0.656, 0.428 and 0.290, respectively. In addition, the standard deviations of the peak values when $\sigma_{pn}$=0.05, 0.1, 0.2 and 0.3 are 0.0261, 0.0431, 0.0604 and 0.0555, respectively. That is, the peak value of the normalized intensity distribution of the SCF decreases with an increase in the normalized standard deviation of the positions of the ridges, $\sigma_{pn}$. As a result, it was shown quantitatively that the spatial-frequency correlation between the two fingerprint images becomes low as the difference between the two becomes large.

In the next subsection, the effect of random noise added to the subject's fingerprint image on the peak value of the normalized intensity distribution of the SCF is investigated quantitatively, in order to evaluate the effects of sweat, sebum and dust, etc., attached at the fingertip on the fingerprint recognition.

### 3.2.3 SCF between the modeled fingerprint images with and without random noise

In this subsection, the effect of the random noise corresponding to sweat, sebum and dust, etc., at the fingertip on the behavior of the peak value of the normalized intensity distribution of the SCF is analyzed.

Fig. 9 shows several examples of the normalized grayscale distributions of the modeled fingerprint images with random noise. Figs. 9(a), 9(b) and 9(c) correspond to the cases when the standard deviations of the normalized grayscale, $\sigma_{gn}$, are 0.02, 0.05 and 0.1, respectively. To obtain these figures, first, we added the Gaussian random noise with the averaged value of 0 and the standard deviation of $\sigma_{gn}$ to the original finite rectangular wave shown in the right side of Fig. 4. Next, we renormalized the obtained wave so as to have a range from 0 to 1. The reason why the renormalization was performed is that the renormalization of the grayscale of the fingerprint image would be conducted in the detecting process of a fingerprint by use of an optical scanner.
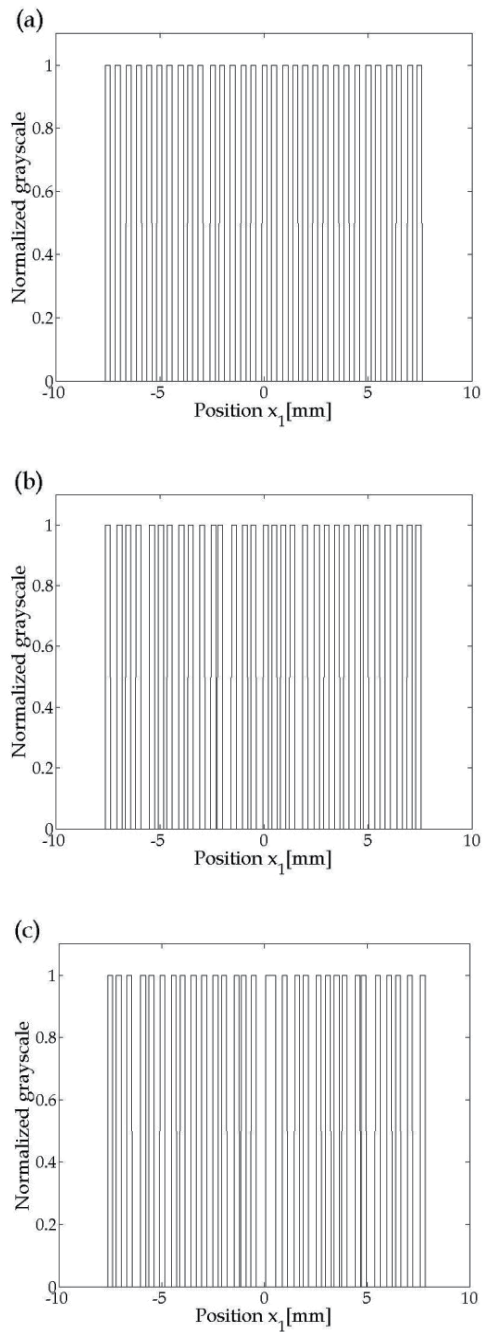
Fig. 6. Normalized grayscale distributions of the modified modeled fingerprint images when the normalized standard deviations of the positions of ridges, $\sigma_{pn}$, are (a)0.05, (b)0.1 and (c)0.2, respectively. $\sigma_{pn}$ is the standard deviation of the variation of the positions of the ridges of the modified rectangular wave, normalized by a period of ridges of the original rectangular wave, i.e., 0.5mm.
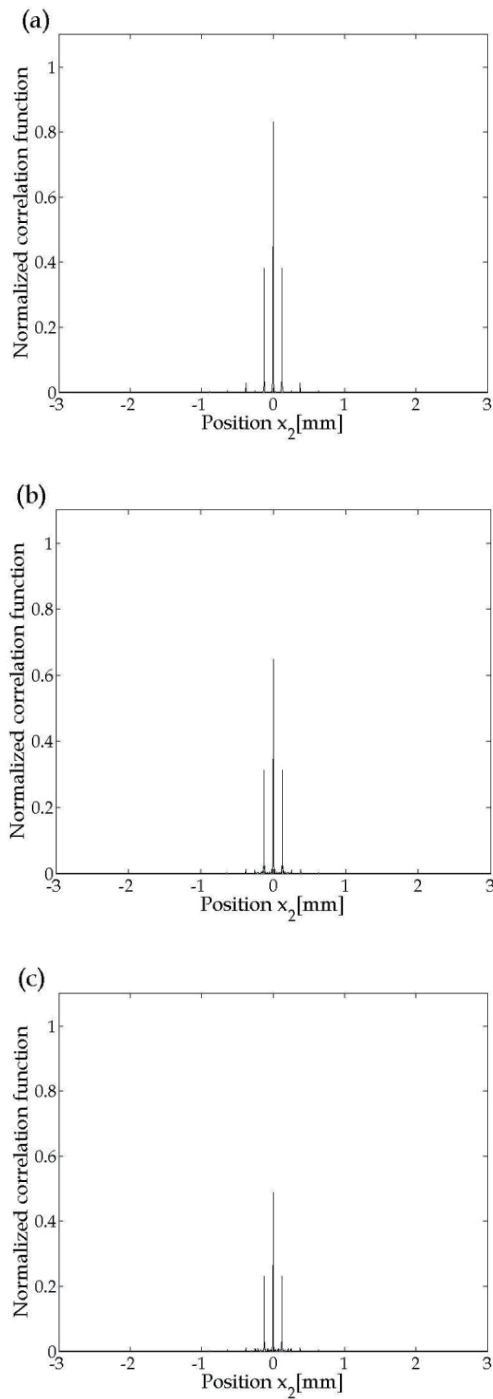
Fig. 7. Normalized intensity distributions of the SCFs between the original finite rectangular wave shown in the right side of Fig. 4 and the modified ones shown in Figs. 6(a), 6(b) and 6(c). The peak values in Figs. 7(a), 7(b) and 7(c) are 0.832, 0.648 and 0.489, respectively.
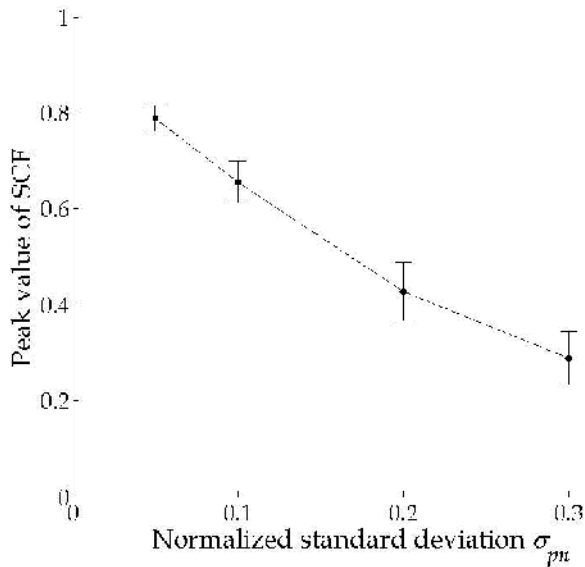
Fig. 8. Dependence of the peak value of the normalized intensity distribution of the SCF on the normalized standard deviation of the positions of ridges of the modified finite rectangular wave, $\sigma_{pn}$. The averaged peak values for $\sigma_{pn}$ of 0.05, 0.1, 0.2 and 0.3 are 0.789, 0.656, 0.428 and 0.290, respectively.

Fig. 10 shows the normalized intensity distributions of the SCFs between the original finite rectangular wave shown in the right side of Fig. 4 and the ones with the Gaussian random noise shown in Figs. 9(a), 9(b) and 9(c). Concretely, Figs. 10(a), 10(b) and 10(c) are the results obtained using Figs. 4 and 9(a), Figs. 4 and 9(b) and Figs. 4 and 9(c), respectively. The peak values in Figs. 10(a), 10(b) and 10(c) are 0.885, 0.759 and 0.652, respectively. This result indicates that the spatial-frequency correlation between the two fingerprint images gradually becomes low as the added random noise becomes large.

Next, in order to investigate the behavior of the peak value of the normalized intensity distribution of the SCF, 1000 kinds of the modeled fingerprint images with the Gaussian random noise were used for each value of $\sigma_{gn}$. Fig. 11 indicates the dependence of the peak value of the normalized intensity distribution of the SCF on the normalized standard deviation of the added random noise, $\sigma_{gn}$. The symbol of circle denotes the averaged peak value and the error bar does the standard deviation of the peak values. As shown in the figure, the averaged peak values when $\sigma_{gn}$ =0.02, 0.05 and 0.1 are 0.891, 0.775 and 0.653, respectively. In addition, the standard deviations of the peak values when $\sigma_{gn}$ =0.02, 0.05 and 0.1 are 0.0114, 0.0216 and 0.0294, respectively. That is, the peak value of the normalized intensity distribution of the SCF decreases with an increase in the normalized standard deviation of the added random noise, $\sigma_{gn}$. As a result, it was shown quantitatively that the spatial-frequency correlation between the two fingerprint images becomes low as the added random noise becomes large.

In the next subsection, we analyze the recognition accuracy of the OSC system by use of the modeled fingerprint images on the basis of the FAR, FRR and MER.

Fig. 9. Normalized grayscale distributions of the modeled fingerprint images with the Gaussian random noise when the standard deviations of the normalized grayscale, $\sigma_{gn}$, are (a)0.02, (b)0.05 and (c)0.1, respectively.

Fig. 10. Normalized intensity distributions of the SCFs between the original finite rectangular wave shown in the right side of Fig. 4 and the ones with the Gaussian random noise shown in Figs. 9(a), 9(b) and 9(c). The peak values in Figs. 10(a), 10(b) and 10(c) are 0.885, 0.759 and 0.652, respectively.

Fig. 11. Dependence of the peak value of the normalized intensity distribution of the SCF on the normalized standard deviation of the added random noise, $\sigma_{gn}$. The averaged peak values for $\sigma_{gn}$ of 0.02, 0.05 and 0.1 are 0.891, 0.775 and 0.653, respectively.

### 3.2.4 Recognition accuracy for the modeled fingerprint images

First, in order to derive the impostor distribution, for example, we paid attention to the result for $\sigma_{pn}$=0.3 in Fig. 8. Fig. 12 indicates the histogram of the peak value of normalized intensity distribution of the SCF between the original finite rectangular wave and the modified one with $\sigma_{pn}$=0.3. The averaged peak value was 0.290 and the standard deviation of the peak values was 0.0555 as already described in subsection 3.2.2.

Next, in order to derive the genuine distribution, for example, we paid attention to the result for $\sigma_{gn}$=0.1 in Fig. 11. Fig. 13 indicates the histogram of the peak value of normalized intensity distribution of the SCF between the original finite rectangular waves with and without the Gaussian random noise having the averaged value of 0 and $\sigma_{gn}$=0.1. The averaged peak value was 0.653 and the standard deviation of the peak values was 0.0294 as already described in subsection 3.2.3.

From the frequency distributions shown in Figs. 12 and 13, the impostor and genuine distributions shown in Fig. 2 can be obtained by fitting the normalized Gaussian distributions to these frequency distributions. Fig. 14 is the result. The left-side red and right-side blue curves correspond to the impostor and genuine distributions, respectively. In this figure, the MER where the FAR and FRR take the same value is $9.34 \times 10^{-4}$% when the authentication threshold is 0.527. As a result, it was found that the recognition accuracy of the OSC system is extremely high.

In the next subsection, we analyze the recognition accuracy of the OSC system by use of real fingerprint images on the basis of the FAR, FRR and MER.
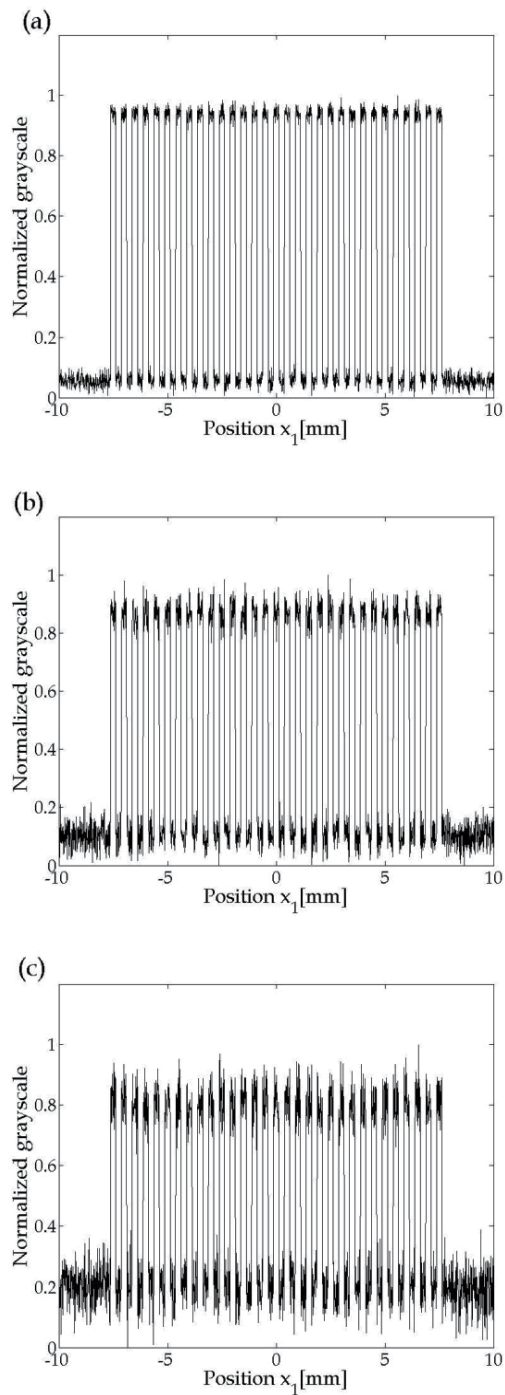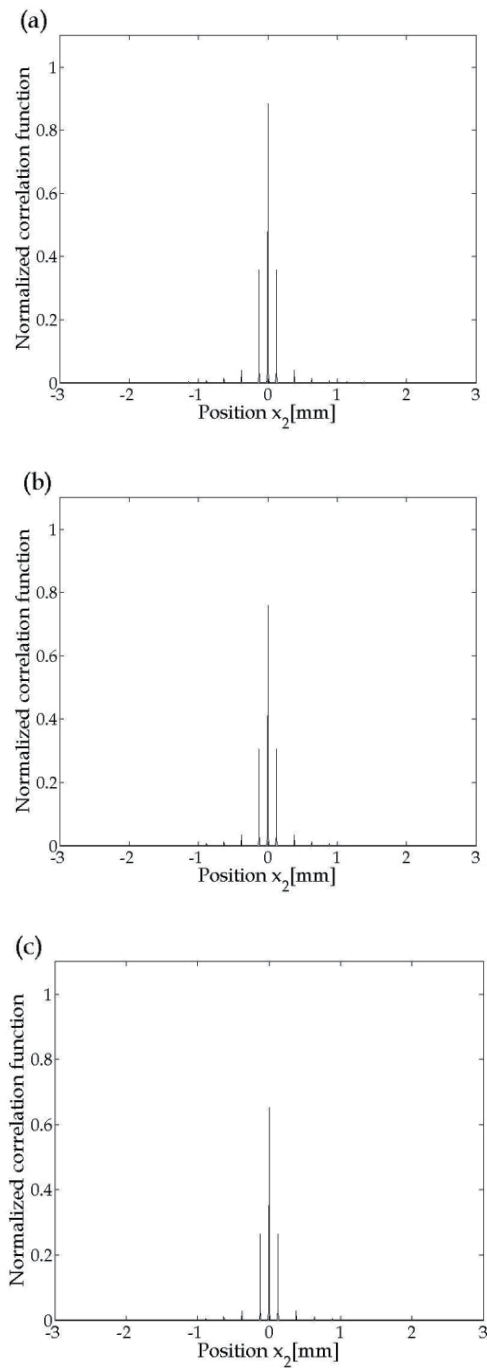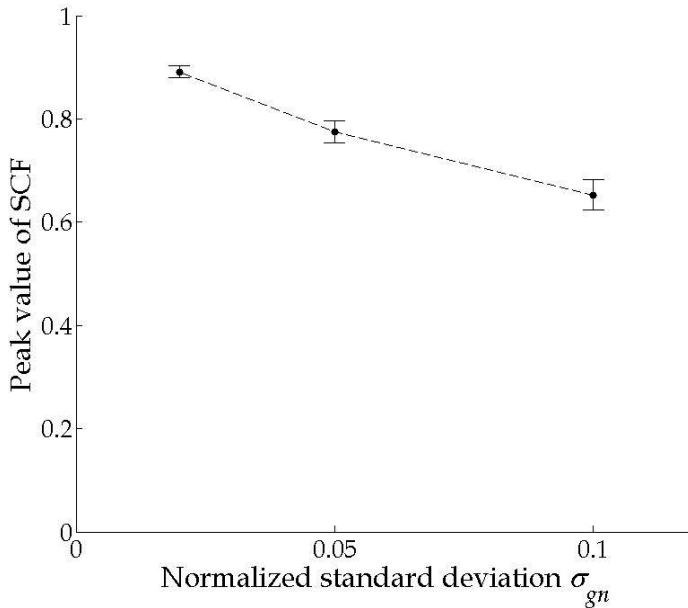
Fig. 12. Histogram of the peak value of the normalized intensity distribution of the SCF between the original finite rectangular wave and the modified one with $\sigma_{pn}$=0.3. The averaged peak value is 0.290 and the standard deviation of the peak values is 0.0555.
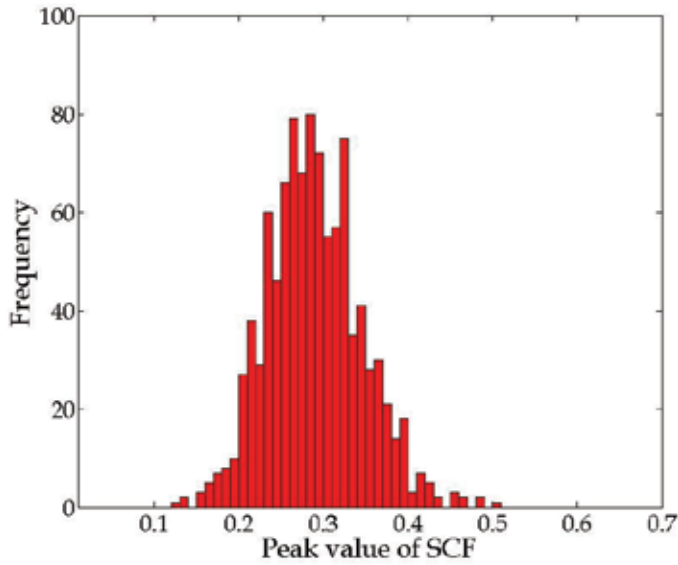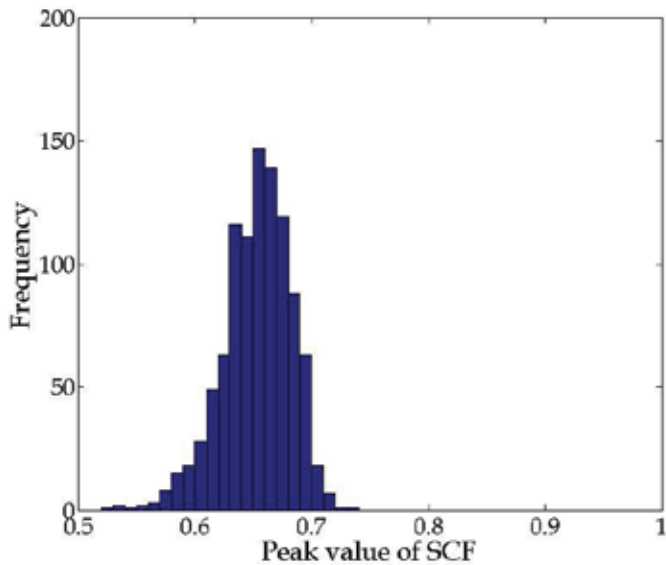


Fig. 13. Histogram of the peak value of the normalized intensity distribution of the SCF between the original finite rectangular waves with and without the Gaussian random noise having the averaged value of 0 and $\sigma_{gn}$=0.1. The averaged peak value is 0.653 and the standard deviation of the peak values is 0.0294.
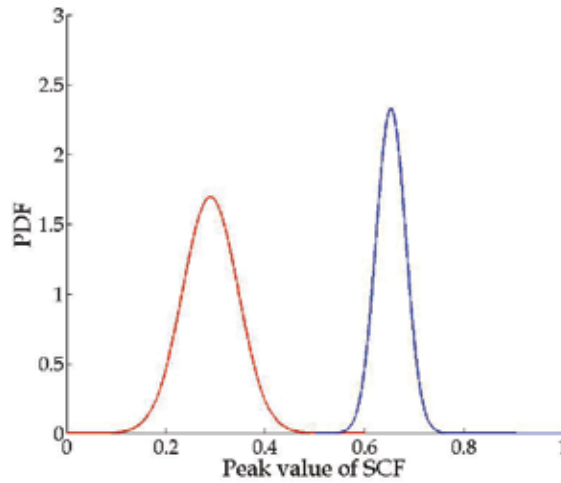
Fig. 14. Impostor and genuine distributions obtained from Figs. 12 and 13, respectively.

### 3.3 Recognition accuracy of the OSC system for real fingerprint images
In this subsection, the recongition accuracy of our proposed system is investigated by use of the real fingerprint images used in the FVC 2002. First, in subsection 3.3.1, the behavior of the peak value of the nomalized intensity distribution of the SCF between two different fingerprint images is shown. Next, in subsection 3.3.2, the behavior of the peak value of the normalized distribution of the SFC between the fingerprint images with and without random noise is also shown. Finally, in subsection 3.3.3, the recognition accuracy of the OSC system is indicated and compared with that of the marketed products of fingerprint recognition system.

### 3.3.1 Behavior of the peak value of the SCF between two different fingerprint images
First, in order to obtain the impostor distribution, we analyzed the frequency distribution of the peak value of the normalized intensity distribution of the SCF between two different fingerprint images. There are 880 fingerprint images for 110 kinds of fingertips in the database used in the FVC 2002. We used 110 fingerprint images which were selected one by one from 110 kinds of fingertips. Therefore, the total number of frequencies was $_{110}C_{109}$=5,995.

Fig. 15 indicates the histogram of the peak value of the normalized intensity distribution of the SCF between two different fingerprint images used in FVC2002. In the figure, the averaged peak value is 0.309 and the standard deviation of the peak values is 0.103. The obtained averaged peak value, 0.309, corresponds well to the result (0.290) when the normalized standard deviation of the positions of ridges, $\sigma_{pn}$, is 0.3, as shown in Fig. 12. However, the obtained standard deviation of the peak values, 0.103, does not corespond well to the result (0.0555) shown in Fig. 12.

Therefore, we may say from the viewpoint of the averaged property that the spatial-frequency correlation between two different real fingerprint images is equivalent to that between the modeled fingerprint image introduced in subsection 3.2.1 and the modified one with $\sigma_{pn}$=0.3 introduced in subsection 3.2.2. However, we found that the standard

deviations of the peak values, which correspond to the extent of the impostor distributions, are different from each other.
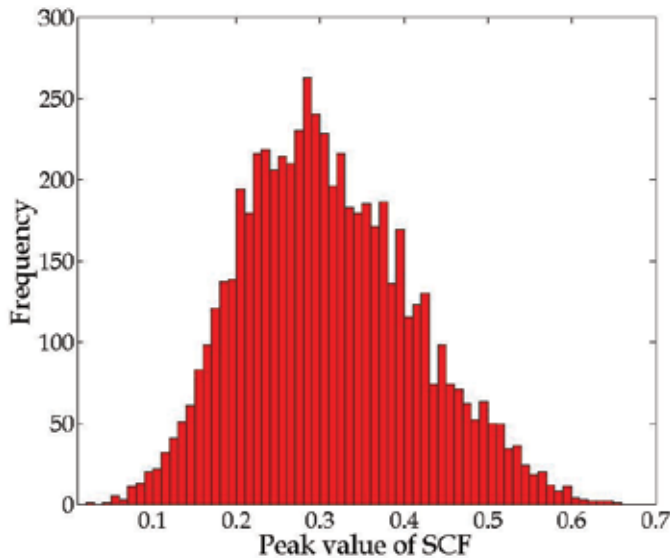


Fig. 15. Histogram of the peak value of the normalized intensity distribution of the SCF between two different fingerprint images used in FVC2002. The averaged peak value is 0.309 and the standard deviation of the peak values is 0.103.

### 3.3.2 Behavior of the peak value of the SCF between the fingerprint images with and without random noise

Next, in order to obtain the genuine distribution, we analyzed the frequency distribution of the peak value of the normalized intensity distribution of the SCF between the fingerprint images with and without random noise. Concretely, the Gaussian random noise with the standard deviation of the normalized grayscale, $\sigma_{gn}$, of 0.1 and the averaged value of 0 was added to the 110 fingerprint images selected in the previous subsection. For each selected fingerprint image, 50 fingerprint images with the Gaussian random noise having the same statistical properties mentioned above were produced. Therefore, the total number of frequencies was 5,500.

Fig. 16 indicates the histogram of the peak value of the normalized intensity distribution of the SCF between the fingerprint images with and without the Gaussian random noise. The averaged peak value is 0.889 and the standard deviation of the peak values is 0.0613. These obtained values of 0.889 and 0.0613 do not correspond well to the results (0.653 and 0.0294, respectively) shown in Fig. 13. In addition, the effect of random noise can be regarded as smaller in case of real fingerprint images because the averaged peak value has a higher value. The reason is considered that the 1D normalized grayscale distribution in a line of the real fingerprint image is not regular like the 1D finite rectangular wave. As a result, it was found that the genuine distribution obtained using the real fingerprint images is different from that obtained using the modeled fingerprint images.
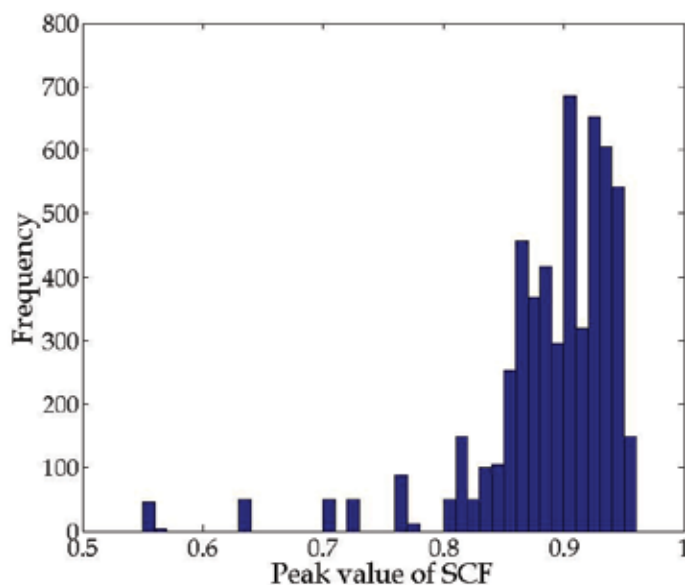
Fig. 16. Histogram of the peak value of the normalized intensity distribution of the SCF between the fingerprint images with and without the Gaussian random noise having the averaged value of 0 and $\sigma_{gn}$=0.1. The averaged peak value is 0.889 and the standard deviation of the peak values is 0.0613.

### 3.3.3 Recognition accuracy for real fingerprint images

From the frequency distributions shown in Figs. 15 and 16, the impostor and genuine distributions shown in Fig. 2 can be obtained by fitting the normalized Gaussian distributions to these frequency distributions. Fig. 17 is the result. The left-side red and right-side blue curves correspond to the impostor and genuine distributions, respectively. In this figure, the MER where the FAR and FRR take the same value is 0.021% when the authentication threshold is 0.672.

In Table 1, the relationship among the authentication threshold, FAR and FRR is summarized. The FAR and FRR are 0.01% and 0.042%, respectively, when the authentication threshold is 0.692. Moreover, the FAR and FRR are 0.001% and 0.26%, respectively, when the authentication threshold is 0.748. As already described in subsection 3.2.4, the MER was $9.34 \times 10^{-4}$%. Therefore, the recognition accuracy becomes low in case of using the real fingerprint images.

In Table 2, the FAR and FRR are shown for several marketed products of fingerprint recognition system. Our OSC system can be classified into a combination of the correlation-based and the frequency-based methods. From the comparison between Tables 1 and 2, it is found that the recognition accuracy of our OSC system is fully high in comparison with that of the existing marketed product named PUPPY FIU-600-N03 (SONY) based on the correlation method. In addition, we can see that the recognition accuracy of our OSC system is comparable to that of the other methods like the minutiae-based and the frequency analysis methods.
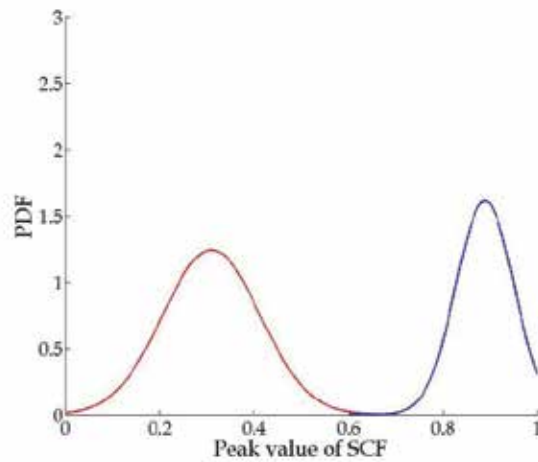
Fig. 17. Impostor and genuine distributions obtained from Figs. 15 and 16, respectively.

| Threshold | FAR(%) | FRR(%) |
|-----------|--------|--------|
| 0.672 | 0.021 | 0.021 |
| 0.692 | 0.01 | 0.042 |
| 0.748 | 0.001 | 0.26 |

Table 1. Relationship among the authentication threshold, FAR and FRR in the OSC system.

| Method | Product | Company | FAR(%) | FRR(%) | Reference |
|--------|---------|---------|--------|--------|-----------|
| Minutiae based | SX-Biometrics Suite | Silex Technology | 0.001 | 0.1 | [1] |
| Correlation based | PUPPY FIU-600-N03 | Sony | ≤0.01 | ≤1.0 | [2] |
| Frequency analysis | UB-safe | DDS | ≤0.001 | ≤0.1 | [3] |

Table 2. Several marketed products of the fingerprint recognition system and their recognition accuracy.

## 4. Conclusions

In this chapter, we have described the OSC system for the fingerprint recognition. Our system has the merit that high-speed authentication would be possible because it could be composed of all optical system. In addition, our system is very simple so that it could be composed in small size.

First, we analyzed the basic properties of the OSC system by use of the modeled fingerprint image of which the grayscale in a transverse line is the 1D finite rectangular wave with a period of 0.5mm and the whole width of the fingertip of 15mm. Concretely, the effect of transformation of the subject's fingerprint, such as variation of positions of ridges, on the fingerprint recognition in the OSC system was analyzed. Moreover, the effect of random noise, such as sweat, sebum and dust, etc., superimposed on the subject's fingerprint on the fingerprint recognition in the OSC system was analyzed. Next, we investigated the recognition accuracy of the OSC system by use of the real fingerprint images used in the FVC 2002 on the basis of the FAR, FRR and MER. As a result, we could make clear that our OSC system has high recognition accuracy of FAR=0.001% and FRR=0.26% in comparison with that in the marketed product based on the correlation-based method. Moreover, our OSC system has comparable recognition accuracy to that in the other marketed products based on the minutiae-based and the frequency analysis methods.

This study has been performed only on the basis of the numerical analysis. Therefore, as a further study, we would produce the OSC system by use of a laser, a lens, etc., and make clear the validity for our OSC system by evaluating our system experimentally from the viewpoint of the recognition accuracy such as the FAR, FRR and MER.

## 5. References

Cappelli, R. ; Ferrara, M. & Maltoni, D. (2010). Minutia Cylinder-Code : A New Representation and Matching Technique for Fingerprint Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 12, pp. 2128-2141, ISSN : 0162-8828

Goodman, J. W. (1996). *Introduction to Fourier Optics*, McGraw-Hill, ISBN : 0-07-024254-2, Singapore

Hashad, F. G. ; Halim, T. M. ; Diab, S. M. ; Sallam, B. M. & Abd El-Samie, F. E. (2010). Fingerprint Recognition Using Mel-Frequency Cepstral Coefficients, *Pattern Recognition and Image Analysis*, Vol. 20, No. 3, pp. 360-369, ISSN : 1054-6618

Jain, A. K. ; Feng, J. & Nandakumar, K. (2010). Fingerprint Matching, *Computer*, Vol. 43, No. 2, pp. 36-44, ISSN : 0018-9162

Kobayashi, Y. & Toyoda, H. (1999). Development of an Optical Joint Transform Correlation System for Fingerprint Recognition, *Optical Engineering,* Vol. 38, No. 7, pp. 1205-1210, ISSN : 0091-3286

Lindoso, A. ; Entrena, L. ; Liu-Jimenez, J. & Millan, E. S. (2007). Correlation-Based Fingerprint Matching with Orientation Field Alignment, In : *Lecture Notes in Computer Science, Vol. 4642, Advances in Biometrics,* Lee, S. –W. & Li, S. Z., (Eds.), pp. 713-721, Springer, ISBN: 978-3-540-74548-8, Berlin

Maltoni, D. & Maio, D. (2002). Download Page of FVC2002, Biometric System Laboratory, University of Bologna, Italy <http://bias.csr.unibo.it/fvc2002/download.asp>

Maltoni, D.; Maio, D.; Jain, A.K. & Prabhakar, S. (2003a). *Handbook of Fingerprint Recognition*, Springer, ISBN : 0-387-95431-7, New York

Maltoni, D.; Maio, D.; Jain, A.K. & Prabhakar, S. (2003b). DVD in the *Handbook of Fingerprint Recognition*, Springer, ISBN : 0-387-95431-7, New York

Nanni, L. & Lumini, A. (2009). Descriptions for Image-Based Fingerprint Matchers, *Expert Systems with Applications*, Vol. 36, No. 10, pp. 12414-12422, ISSN : 0957-4174

Sheng, W. ; Howells, G. ; Fairhurst, M. & Deravi, F. (2007). A Memetic Fingerprint Matching Algorithm, *IEEE Transactions on Information Forensics and Security*, Vol. 2, No. 3, pp. 402-412, ISSN : 1556-6013

Takeuchi, H. ; Umezaki, T. ; Matsumoto, N. & Hirabayashi, K. (2007). Evaluation of Low-Quality Images and Imaging Enhancement Methods for Fingerprint Verification, *Electronics and Communications in Japan, Part 3*, Vol. 90, No. 10, pp. 40-53, Online ISSN : 1520-6424

Xu, H. ; Veldhuis, R. N. J. ; Bazen, A. M. ; Kevenaar, T. A. M. ; Akkermans, T. A. H. M. & Gokberk, B. (2009a). Fingerprint Verification Using Spectral Minutiae Representations, *IEEE Transactions on Information Forensics and Security*, Vol. 4, No. 3, pp. 397-409, ISSN : 1556-6013

Xu, H. ; Veldhuis, R. N. J. ; Kevenaar, T. A. M. & Akkermans, T. A. H. M. (2009b). A Fast Minutiae-Based Fingerprint Recognition System, *IEEE Systems Journal*, Vol. 3, No. 4, pp. 418-427, ISSN : 1932-8184

Yang, J. C. & Park, D. S. (2008a). A Fingerprint Verification Algorithm Using Tessellated Invariant Moment Features, *Neurocomputing*, Vol. 71, Nos. 10-12, pp. 1939-1946, ISSN : 0925-2312

Yang, J. C. & Park, D. S. (2008b). Fingerprint Verification Based on Invariant Moment Features and Nonlinear BPNN, *International Journal of Control, Automation, and Systems,* Vol. 6, No. 6, pp. 800-808, ISSN : 1598-6446

Yoshimura, H. & Takeishi, K. (2009). Optical Spatial-Frequency Correlation System for Biometric Authentication, *Proceedings of SPIE*, Vol. 7442, *Optics and Photonics for Information Processing III*, Iftekharuddin, K. M. & Awwal, A. A. S., (Eds.), 74420N, ISBN : 9780819477323

[1] Specifications written in http://www.silexamerica.com/products/biometrics/sx-biometrics_suite.html

[2] http://www.sony.co.jp/Products/Media/puppy/pdf/FIU600_01.PDF

[3] http://www.dds.co.jp/en/technology/ub_safe.html

# 5

# On the Introduction of Secondary Fingerprint Classification

Ishmael S. Msiza[1], Jaisheel Mistry[1], Brain Leke-Betechuoh[1],
Fulufhelo V. Nelwamondo[1,2] and Tshilidzi Marwala[2]
*[1]Biometrics Research Group, CSIR Modelling & Digital Sciences*
*[2]Faculty of Engineering & the Built Environment, University of Johannesburg*
*Republic of South Africa*

## 1. Introduction

The concept of fingerprint classification is an important one because of the need to, before executing a database search procedure, virtually break the fingerprint template database into smaller, manageable partitions. This is done in order to avoid having to search the entire template database and, for this reason, minimize the database search time and improve the overall performance of an automated fingerprint recognition system. The commonly used primary fingerprint classes add up to a total of five (Msiza et al., 2009):

- Central Twins (CT),

- Left Loop (LL),

- Right Loop (RL),

- Tented Arch (TA), and

- Plain Arch (PA).

Many fingerprint classification practitioners, however, often reduce these five fingerprint classes to four. This is, at a high level, due to the difficulty in differentiating between the TA and the PA class. These two similar classes are often combined into what is referred to as the Arch (A) class. Recent examples of practitioners that have reduced the five-class problem to a four-class problem include Senior (2001), Jain & Minut (2002), and Yao et al. (2003). The not so recent examples include Wilson et al. (1992), Karu & Jain (1996), and Hong & Jain (1998). These four primary classes are sufficient in the performance improvement of small-scale applications such as access control systems and attendance registers of small to medium-sized institutions. They, however, may not be sufficient in the performance improvement of large-scale applications such as national Automatic Fingerprint Identification Systems (AFIS). In order to enforce visible performance improvement on such large-scale applications, this chapter introduces a two-stage classification system, by taking advantage of the extensibility of the classification rules that utilize the arrangement of the fingerprint global landmarks, known as the singular points (Huang et al., 2007) (Mathekga & Msiza, 2009).

The first classification stage produces the primary fingerprint classes and then the second classification stage breaks each primary class into a number of secondary classes. It is

important to note that the concept of secondary fingerprint classification is one that has not been exploited by fingerprint classification practitioners, and is being formally introduced in this chapter for the first time. The next section presents a detailed discussion of both the primary and the secondary fingerprint classes.

## 2. Primary and secondary fingerprint classes

This section presents the proposed primary and secondary fingerprint classes, together with the rules used to determine them. It is important to note that the rules used to determine these primary and the secondary classes are based on the arrangement of the fingerprint singular points, namely, the fingerprint core and the fingerprint delta. Forensically, a fingerprint core is defined as the innermost turning point where the fingerprint ridges form a loop, while the fingerprint delta is defined as the point where these ridges form a triangulating shape (Leonard, 1988). Figure 1 depicts a fingerprint with the core and delta denoted by the circle and the triangle, respectively.
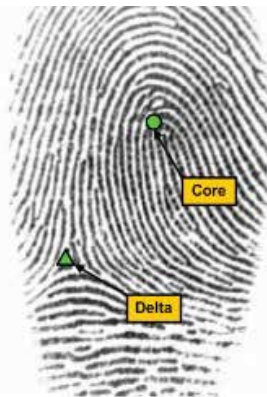


Fig. 1. A fingerprint showing clear markings of the core (circle) and the delta (triangle)

### 2.1 Central Twins (CT) primary class and its secondary classes

Fingerprints that belong to the CT class are, at a primary level, those that have ridges that either form (i) a circular pattern, or (ii) two loops, in the central area of the print. Some practitioners usually refer to the circular pattern as a whorl (Park & Park, 2005), while the two-loop pattern is referred to as a twin loop (Karu & Jain, 1996). The similarity, however, between the two patterns is that they both have cores located next to each other in the central area of the fingerprint, which is the main reason why Msiza et al. (2009) grouped these two patterns into the same class, called the Central Twins class. Figure 2(a) shows the whorl pattern, while the twin loop pattern is depicted on figure 2(b).

In addition to the two cores located in the central area, fingerprints belonging to CT class also have two deltas. These two deltas, however, are not located in the central area of the print, which immediately implies that there is a chance that one, or even both, may not be captured. All of this is dependent on how the user or subject impresses their finger, for capturing, on the surface of the fingerprint acquisition device. This is what brings into point the possibility of deriving secondary classes of this CT primary class.

The CT secondary classes derived in this chapter are depicted in figure 3, and they add up to a total of three. Figure 3(a) shows a CT class fingerprint that has all the singular points captured,
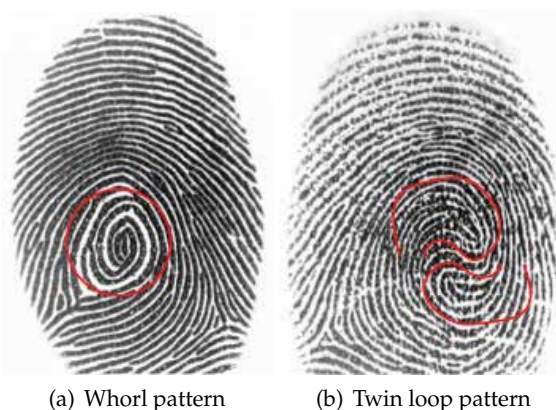
(a) Whorl pattern          (b) Twin loop pattern

Fig. 2. Fingerprint patterns that collectively belong to the CT primary class. The whorl pattern has a circular structure that forms two cores, and the twin loop pattern has two loops that form two cores

two cores and two deltas, which is an ideal case. Such a complete capture of information normally occurs in applications where fingerprints are rolled, instead of being slapped. This is because of the fact that deltas, in fingerprints that belong to the CT primary class, are normally located adjacent to the edges of the fingerprint ridge area. A CT class fingerprint that has two cores and two deltas captured, is assigned to what is introduced as the CT-1 secondary class. A CT class fingerprint that has two cores and one delta, as shown in figure 3(b), is assigned to what is introduced as the CT-2 secondary class while the one that has two cores and no delta, as depicted in figure 3(c), is assigned to what is introduced as the CT-3 secondary class.



(a) CT-1 secondary class     (b) CT-2 secondary class     (c) CT-3 secondary class

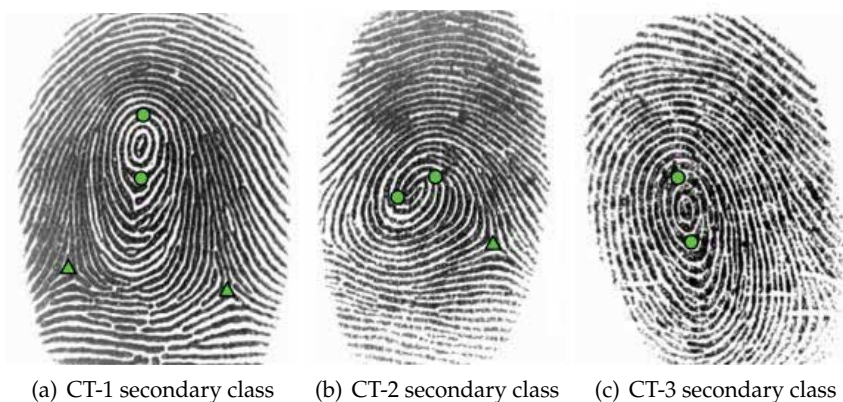Fig. 3. Fingerprint patterns that determine the CT secondary classes. CT-1 class: 2 cores & 2 deltas; CT-2 class: 2 cores & 1 delta; and CT-3 class: 2 cores & no delta

## 2.2 Arch (A) primary class and its secondary classes

Fingerprints that belong to the A class are, at a primary level, those that have ridges that appear to be entering the fingerprint on one side, rise in the middle area of the fingerprint,

and leave the fingerprint on the opposite side, as depicted in figure 4. Figure 4(a) shows a fingerprint pattern that some practitioners normally classify as a plain arch, while figure 4(b) depicts a pattern that some practitioners classify as a tented arch. The technical report of Hong & Jain (1998) is one example of the practice of ordering these two patterns into separate classes. A year later, however, Jain et al. (1999) realized that there is often a mis-classification between the two patterns, hence it is better to combine them into one class. Many other practitioners, including Msiza et al. (2009), have observed that combining the plain arch and the tented arch patterns into one class, does improve the classification accuracy.



(a) Plain arch pattern          (b) Tented arch pattern
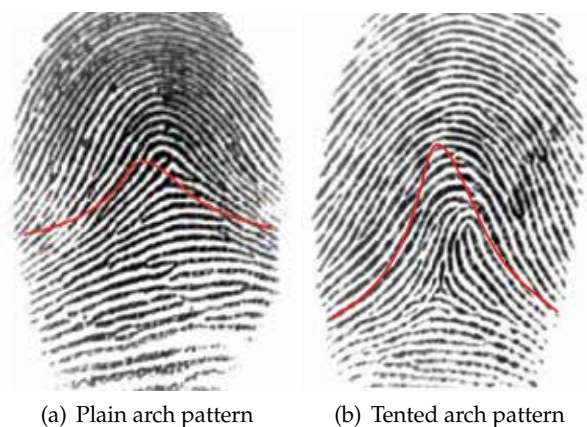
Fig. 4. Fingerprint patterns that collectively belong to the A primary class. The plain arch pattern has no singular points while the tented arch pattern has a core and a delta, with the delta located almost directly below the core

Because of this reality, it is proposed that these two patterns are better off at a secondary level of fingerprint classification. This immediately provides a platform for the proposition of a number of A class secondary rules. An A class fingerprint that is without both a core and a delta, is assigned to what is introduced as the A-1 secondary class. Msiza et al. (2009) suggest that, for an A class fingerprint that has a core and delta detected, the absolute difference between their $x$-coordinates, $\Delta x$, is less than or equal to 30 pixels. It is, for this reason, proposed that if an A class fingerprint has a core and delta detected, and:

$$pixels\ 15 \leqslant \Delta x \leqslant 30\ pixels, \tag{1}$$

then the fingerprint is assigned to what is introduced as the A-2 secondary class, else if:

$$pixels\ 0 \leqslant \Delta x < 15\ pixels, \tag{2}$$

then fingerprint is assigned to what is introduced as the A-3 secondary class. Equation 2 is used for the instances where the rise of the ridges in the middle part of the fingerprint is extremely acute, hence $\Delta x$ is extremely small. Figure 5 depicts all three A secondary fingerprint classes.

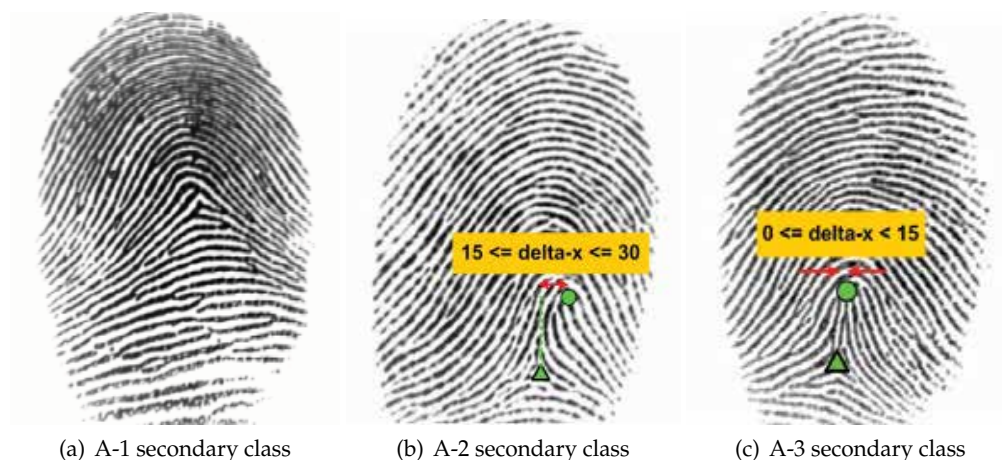(a) A-1 secondary class    (b) A-2 secondary class    (c) A-3 secondary class

Fig. 5. Fingerprint patterns that determine the A secondary classes. A-1 class: 0 cores & 0 deltas; A-2 class: equation 1; and A-3 class: equation 2

### 2.3 Left Loop (LL) primary class and its secondary classes

Fingerprints that belong to the LL class are, at a primary level, those that have ridges that appear to be entering the fingerprint on the left hand side, make a loop in the middle area of the fingerprint, and leave the fingerprint on the same side where they entered. The loop in the middle area is what forms the core of the print. An example of a fingerprint that belongs to this class is depicted on figure 6. In addition to the core that is formed by the loop in the middle area, an LL fingerprint has a delta located at the bottom of the loop, adjacent to the right hand side edge of the print. Depending on how the finger is impressed against the surface of the capturing device, there is always a chance that the delta may not be captured, more especially because it is adjacent to the edge of the fingerprint. This, therefore, presents an opportunity for the formulation of two LL secondary classes.



Fig. 6. A fingerprint pattern that belongs to the LL primary class. The ridges enter the print on the left hand side, make a loop in the middle, and leave on the same side

If a fingerprint that belongs to the LL class has (i) both a core and a delta detected, (ii) the conjugate slope (C-Slope) of the line joining the core and the delta is negative, and (iii) $\Delta x > 30$

pixels, then this fingerprint is assigned to what is introduced as the LL-1 secondary fingerprint class. The said C-Slope is just a complement of the conventional slope, because its reference point, or origin, is not the geometric center of the fingerprint image, but is the top left hand corner of the image. The LL-1 classification rules are summarized in figure 7(a).

If a fingerprint that belongs to the LL class has (i) only a core detected, and (ii) the auxiliary ($\theta$) is less than 90 degrees, then the fingerprint is assigned to what is introduced as the LL-2 secondary class. The auxiliary ($\theta$) is mathematically defined:

$$\theta = arctan(M) \tag{3}$$

where $M$ is the C-Slope of the line joining the core and the pedestrian point (Msiza et al., 2009). The pedestrian is a point located along the bottom of the fingerprint image, exactly below the True Fingerprint Center Point (TFCP), as shown in figure 7(b). Its $x$-coordinate is exactly the same as the one of the TFCP, and its $y$-coordinate has the same value as the height of the fingerprint image. The TFCP is defined as the geometric center of the fingerprint ridge area, that is, the fingerprint foreground (Msiza et al., 2011). Figure 7(b) shows the TFCP marked by the point of intersection of the two Cartesian axes.



(a) LL-1 secondary class          (b) LL-2 secondary class

Fig. 7. Fingerprint patterns that determine the LL secondary classes. LL-1 secondary class: 1 core & 1 delta, with C-Slope < 0; and LL-2 secondary class: 1 core & 0 delta, with $\theta < 90$ degrees

### 2.4 Right Loop (RL) primary class and its secondary classes

Fingerprints that belong to the RL class are, at a primary level, those that have ridges that appear to be entering the fingerprint on the right hand side, make a loop (which forms a core) in the middle area of the fingerprint, and leave the fingerprint on the same side where they entered. An example of a fingerprint that belongs to this RL class is depicted on figure 8. In addition to the core that is formed by the loop in the middle, an RL fingerprint has a delta located at the bottom of the loop, adjacent to the left hand side edge of the print. Similarly, depending on how the finger is impressed against the capturing device, there is always a chance that the delta may not be captured, more especially because it is adjacent to the edge

of the fingerprint. This, therefore, presents an opportunity for the formulation of two RL secondary classes.



Fig. 8. A fingerprint pattern that belongs to the RL primary class. The ridges enter the print on the right hand side, make a loop in the middle, and leave on the same side

If a fingerprint that belongs to the RL class has (i) both a core and a delta detected, (ii) the C-Slope of the line joining the core and the delta is positive, and (iii) $\Delta x \leqslant 30$ pixels, then this fingerprint is assigned to what is introduced as the RL-1 secondary fingerprint class. If a fingerprint that belongs to the RL class has (i) only a core detected, and (ii) the auxiliary ($\theta$) is greater than or equal to 90 degrees, then the fingerprint is assigned to what is introduced as the RL-2 secondary fingerprint class. These two secondary classification rules are summarized in figure 9.
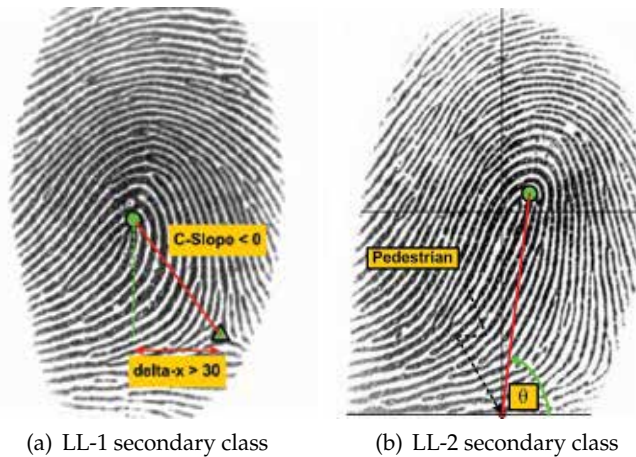


(a) RL-1 secondary class          (b) RL-2 secondary class
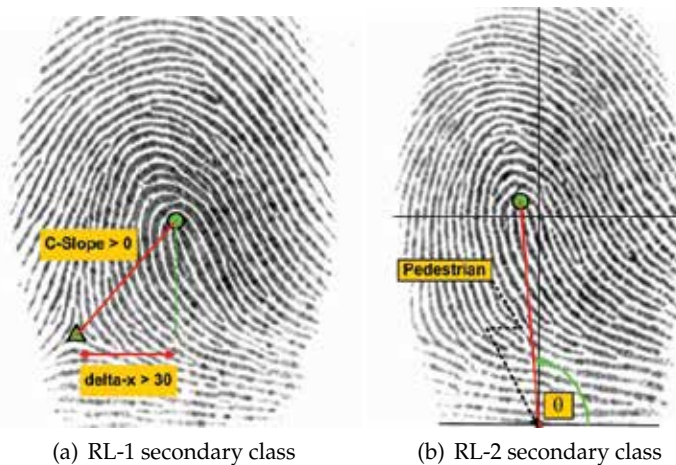
Fig. 9. Fingerprint patterns that determine the RL secondary classes. RL-1 secondary class: 1 core & 1 delta, with C-Slope > 0; and LL-2 secondary class: 1 core & 0 delta, with $\theta \geqslant 90$ degrees

**2.5 Classes overview**

Following the proposed primary and secondary classes, figure 10 presents a combined picture that shows the relationship between all of them. The primary classification layer consists of 4 instances, while the secondary classification layer consists of a total of 10 instances.



Fig. 10. An overview of the proposed primary and secondary fingerprint classes

## 3. Implementation of the proposed classification scheme

The implementability of the proposed classification scheme is demonstrated through the pseudo-code presented in algorithm 1. It is important to note that, before classification can be done, the captured fingerprint has to go through some pre-processing. These pre-processes include:

- contrast enhancement (Hong et al., 1998),

- ridge segmentation (Maltoni et al., 2009),

- orientation image computation and smoothing (Ratha et al., 1995), and

- singular point detection (Mathekga & Msiza, 2009).

The credibility of this proposed classification scheme is evaluated, in two different ways, in the next section.

---

**Algorithm 1:** The main procedure that, when presented with singular points, determines both a fingerprint's primary and secondary class

---

**Input** : Fingerprint singular points
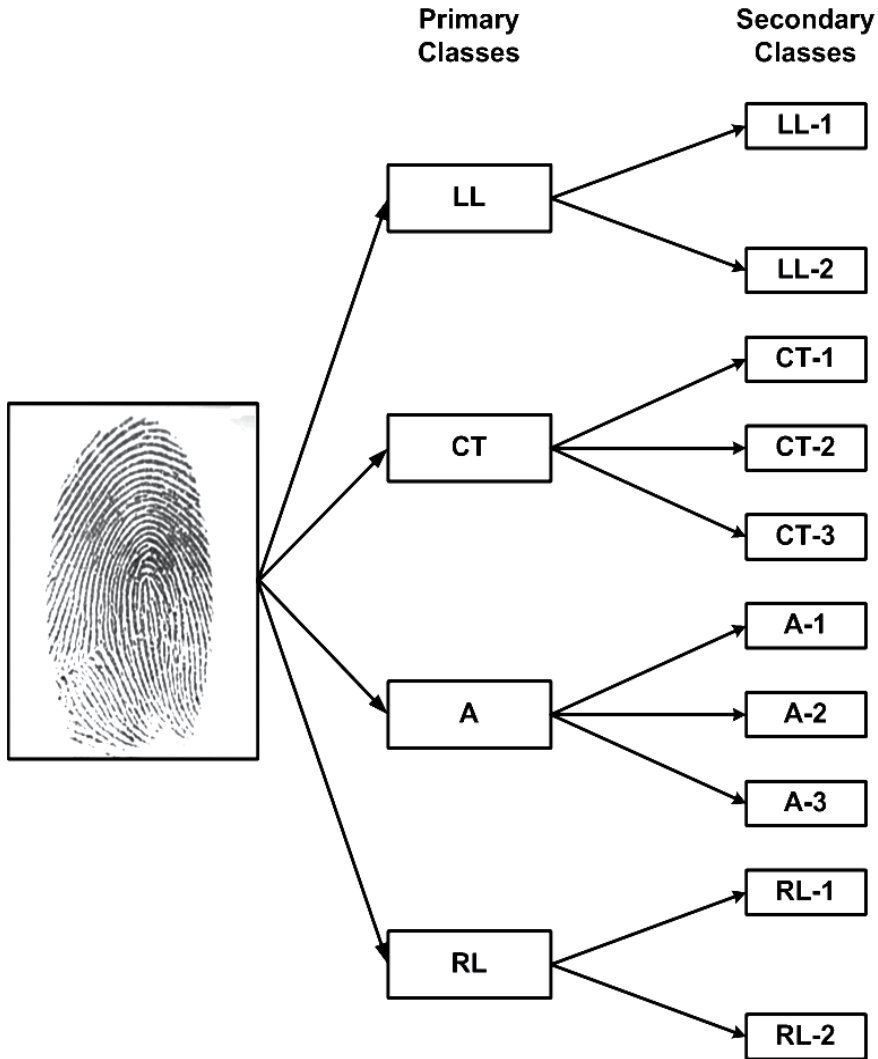**Output**: Fingerprint primary and secondary class

**begin**
    initialize: primary class = unknown, and secondary class = unknown;
    calculate: the number of cores, $N_C$, and the number of deltas, $N_D$, detected;

    **if** $N_C = 0$ *and* $N_D = 0$ **then**
        | use algorithm 2 for classification;
    **end**
    **else if** $N_C = 1$ *and* $N_D = 0$ **then**
        | use algorithm 3 for classification;
    **end**
    **else if** $N_C = 1$ *and* $N_D = 1$ **then**
        | use algorithm 4 for classification;
    **end**
    **else if** $N_C = 2$ *and* $N_D$ *is between* 0 *and* 2 **then**
        | use algorithm 5 for classification;
    **end**
**end**

---

**Algorithm 2:** A procedure that, when presented with neither core nor delta, determines both a fingerprint's primary and secondary class

---

**Input** : Zero core and zero delta
**Output**: Fingerprint primary and secondary class

**begin**
    | primary class = A;
    | secondary class = A-1;
**end**

---

## 4. Classifier performance evaluation

In order to evaluate the credibility of the idea of secondary fingerprint classification, it is important to measure the accuracy of both the primary and the secondary classification module. If this idea is indeed credible, the difference between the accuracy value of the primary module and the one of the secondary module should be small. It should be small to an

---

**Algorithm 3:** A procedure that, when presented with one core and no delta, determines both a fingerprint's primary and secondary class

---

**Input**   : One core and zero delta
**Output**: Fingerprint primary and secondary class

**begin**
    compute: the coordinates of the pedestrian;
    compute: the C-Slope, *M*, of the line joining the core and the pedestrian;
    compute: the auxiliary, $\theta$, using equation 3;

    **if** $\theta < 90$ *degrees* **then**
        primary class = LL;
        secondary class = LL-2;
    **end**
    **else if** $\theta \geqslant 90$ *degrees* **then**
        primary class = RL;
        secondary class = RL-2;
    **end**
**end**

---

extent that it should tempt any fingerprint classification practitioner to, in future applications, consider using the proposed secondary fingerprint classes as primary classes.

In addition to the accuracy values, the proposed classification scheme's credibility should be evaluated through observing the time it takes a fingerprint recognition system to search through a template database (i) without any classification, (ii) with only the primary classification module, and (ii) with the secondary classification module. For this classification scheme to be regarded as credible, the average database search time for cases (ii) and (iii) must be less than that for case (i), and the one for case (iii) should be less than the one for for case (ii), while the matching rates remain significantly unchanged. For the purposes of this evaluation, the CSIR-Wits Fingerprint Database (CWFD) which was jointly collected, for academic research purposes, by the Council for Scientific & Industrial Research (CSIR) and the University of the Witwatersrand (Wits), both in the Republic of South Africa.

### 4.1 Classification rates

This section presents the classification accuracy values, in the form of confusion matrices, of both the primary and the secondary classification modules. A confusion matrix is a table that shows a summary of the classes assigned by the automated fingerprint classifier, measured against those assigned by a human fingerprint classification expert. The classification accuracy value is mathematically expressed as:

$$Accuracy = \frac{M}{T} \times 100\%, \qquad (4)$$

where *M* is the sum of the main diagonal of the matrix, and *T* is the sum of all the instances of data in the chosen database. Evaluated on a database that contains 946 instances, table 1 shows the confusion matrix for the primary classification module, while table 2 shows the confusion matrix of the secondary classification module.

---

**Algorithm 4:** A procedure that, when presented with one core and one delta, determines both a fingerprint's primary and secondary class

---

**Input** : One core and one delta
**Output**: Fingerprint primary and secondary class

**begin**
    compute: the absolute difference, $\Delta x$, between the $x$-coordinates;

    **if** $\Delta x \leqslant 30$ *pixels* **then**
        primary class = A;

        **if** $\Delta x \leqslant 15$ *pixels* **then**
          | secondary class = A-2;
        **end**
        **else if** $\Delta x < 15$ *pixels* **then**
          | secondary class = A-3;
        **end**
    **end**
    **if** $\Delta x > 30$ *pixels* **then**
        compute the C-Slope of the line joining the core and the delta;

        **if** *C-Slope is Positive* **then**
          primary class = RL;
          secondary class = RL-1;
        **end**
        **else if** *C-Slope is Negative* **then**
          primary class = LL;
          secondary class = LL-1;
        **end**
    **end**
**end**

---

Table 1 displays a classification accuracy of 80.4%, which is an acceptable figure for a four-class problem. As an example, Senior (1997) obtained a classification accuracy of 81.6% for his four-class problem. Some of the A class fingerprints are mis-classified as LL and RL because it is not all of them that have a $\Delta x$ that is less than 30 pixels. Possible future improvements, therefore, involve a bit more experimentation on a range of $\Delta x$ values. Some of the CT class fingerprints are mis-classified as A possibly because the singular point detection module was unable to detect the cores of the fingerprints. A possible future improvement, therefore, involves working on the functionality of the singular point detection module. Some of the LL class fingerprints are mis-classified as A because it is not all the LL fingerprints that have a $\Delta x$ that is greater than 30 pixels, and the same reasoning can be attributed to the mis-classification of some of the RL class fingerprints. Possible future improvements, again, involve a bit more experimentation on a range of $\Delta x$ values.

The secondary classification accuracy in table 2 has a value of 76.8%, which is an encouraging figure for a newly introduced concept. This implies that there is a difference of only 3.6% between the primary and the secondary classification modules. This, therefore, provides future opportunities for a classification practitioner to fine-tune the secondary classification

---

**Algorithm 5:** A procedure that, when presented with two cores and zero or a few deltas, determines both a fingerprint's primary and secondary class

---

**Input** : Two cores and zero or a few deltas
**Output**: Fingerprint primary and secondary class

**begin**
    primary class = CT;
    calculate: the exact number of deltas, $N_D$, detected;
    **if** $N_D = 0$ **then**
        |  secondary class = CT-3;
    **end**
    **else if** $N_D = 1$ **then**
        |  secondary class = CT-2;
    **end**
    **else if** $N_D = 2$ **then**
        |  secondary class = CT-1;
    **end**
**end**

---

| Actual | | As | | | |
|---|---|---|---|---|---|
| | A | CT | LL | RL | Total |
| A | **200** | 03 | 25 | 36 | 264 |
| CT | 18 | **187** | 05 | 10 | 220 |
| LL | 10 | 08 | **152** | 06 | 176 |
| RL | 29 | 11 | 24 | **222** | 286 |
| | | | 80.4% | | 946 |

Table 1. The primary class experimental results tested on the CWFD, which contains 946 instances of data

rules in order to further close down the gap between the two classification modules. As soon as this gap approaches zero, these newly introduced secondary classes can be used as primary classes and, with a total of 10 primary classes, there will be countless opportunities to further reduce the database search time. This is achievable through the introduction of another set of secondary classes by using unsupervised techniques such as artificial neural networks (Marwala, 2007).

### 4.2 Average search times and matching rates
To further demonstrate the credibility of the proposed classification scheme, this section presents its performance when measured through the average database search time, together with the matching rates, also done on the CWFD. These matching rates are listed as follows:

- True Match Rate (TMR)

- False Match Rate (FMR)

- True Non-Match Rate (TNMR)

| Actual | As | | | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | A-1 | A-2 | A-3 | CT-1 | CT-2 | CT-3 | LL-1 | LL-2 | RL-1 | RL-2 | |
| A-1 | **118** | 01 | 01 | 00 | 00 | 03 | 03 | 05 | 00 | 23 | 154 |
| A-2 | 05 | **17** | 02 | 00 | 00 | 00 | 02 | 01 | 00 | 05 | 33 |
| A-3 | 06 | 02 | **48** | 00 | 00 | 00 | 05 | 09 | 00 | 08 | 77 |
| CT-1 | 00 | 00 | 00 | **00** | 00 | 00 | 00 | 00 | 00 | 00 | 00 |
| CT-2 | 03 | 00 | 00 | 00 | **16** | 03 | 00 | 00 | 00 | 00 | 22 |
| CT-3 | 14 | 00 | 01 | 00 | 07 | **161** | 00 | 05 | 01 | 09 | 198 |
| LL-1 | 00 | 00 | 00 | 00 | 00 | 00 | **09** | 00 | 00 | 02 | 11 |
| LL-2 | 07 | 03 | 00 | 00 | 03 | 05 | 03 | **140** | 01 | 03 | 165 |
| RL-1 | 02 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | **09** | 00 | 11 |
| RL-2 | 24 | 03 | 00 | 00 | 00 | 11 | 02 | 22 | 04 | **209** | 275 |
| | | | | | 76.8% | | | | | | 946 |

Table 2. The secondary class experimental results tested on the CWFD, which contains 946 data instances

- False Non-Match Rate (FNMR)

A true match occurs when a fingerprint recognition system correctly regards a genuine comparison, $C_G$, as genuine. Given a matching threshold $T$, the TMR value of $T$ is the number of genuine comparisons with match scores greater than $T$, divided by the total number of genuine samples, $S_G$, presented for comparison. Mathematically, this is modeled as:

$$TMR = \frac{Count\{C_G \geqslant T\}}{S_G} \times 100\%. \tag{5}$$

A false match occurs when a fingerprint recognition system regards an impostor comparison, $C_I$, as genuine. The FMR value of $T$ is the number of impostor comparisons with match scores greater than $T$, divided by the total number of impostor samples, $S_I$, presented for comparison. Mathematically, the FMR can be modeled as:

$$FMR = \frac{Count\{C_I \geqslant T\}}{S_I} \times 100\%. \tag{6}$$

A true non-match occurs when a fingerprint recognition system correctly regards an impostor comparison as an impostor. The TNMR value of $T$ is the number of impostor comparisons with match scores less than $T$, divided by the total number of impostor samples presented for comparison. Mathematically, this can be modeled as:

$$TNMR = \frac{Count\{C_I < T\}}{S_I} \times 100\%. \tag{7}$$

A false non-match occurs when the fingerprint recognition system regards a genuine comparison as an impostor. The FNMR value of $T$ is the number of genuine comparisons with match scores less than $T$, divided by the total number of genuine samples presented for

comparison. Mathematically, this can be modeled as:

$$FNMR = \frac{Count\{C_G < T\}}{S_G} \times 100\%. \tag{8}$$

Table 3 shows the results obtained from the evaluation, where 3 instances of the same fingerprint were enrolled into the template database, in order to make the system more accurate. The template database, for this reason, ended up with a total of $3 \times 86 = 258$ instances. The credibility of the proposed classification scheme is verified by the fact that the average database search time (AST) is improved from 2 426 ms to 645 ms and 492 ms by the primary and the secondary classification module, respectively, while the matching rates remain significantly unchanged.

|  | No Classification | Primary Classification | Secondary Classification |
| --- | --- | --- | --- |
| True Match Rate (TMR) | 78.3% | 70.4% | 66.2% |
| False Match Rate (FMR) | 0.7% | 0.2% | 0.1% |
| True Non-Match Rate (TNMR) | 99.3% | 99.1% | 99.2% |
| False Non-Match Rate (FNMR) | 21.6% | 32.2% | 30.2% |
| Average Search Time (AST) | 2 426 ms | 645 ms | 492 ms |

Table 3. A summary of the match and non-match rates together with the average database search times, tested on the CWFD

Because the TMR and the FNMR are complements of each other, their values should add up to a 100%. For the same reason, the values of the FMR and the TNMR should add up to a 100%. The reason why this is not case in the third and the fourth columns of table 3 is that the database search was done continuously per group of fingerprint instances of a common subject, which leads to a loss of data. This loss of data is, in essence, attributable to a combination of possible mis-classifications and failure to meet the matching threshold.

## 5. Discussions and conclusions

This chapter presented the concept of automatic fingerprint classification, in general, and introduced the concept of secondary fingerprint classification, in particular. Secondary fingerprint classification was introduced in order to further reduce the time it takes for an automated fingerprint recognition system to search through a database of templates. The key fingerprint features employed in the proposed classification scheme are the core and the delta, with a total of 4 primary fingerprint classes; namely: CT, A, LL, and RL; and 10 secondary fingerprint classes, namely: CT-1, CT-2, CT-3, A-1, A-2, A-3, LL-1, LL-2, RL-1, and RL-2. Using a confusion matrix as a performance measure, the primary fingerprint classification module registered an accuracy of 80.4%, while the secondary classification module registered an accuracy of 76.8%. This 3.6% gap is indicative of the fact that, in future applications, there is a chance to fine-tune the secondary classification rules and, after improving the accuracy, there is even a good chance to use these secondary classes at a primary level. With a total of 10 fingerprint classes at a primary level, there is a good chance of decreasing the database search time even further, while the change in matching rates remains acceptably small.

## 6. References

Hong, L. & Jain, A.K. (1998). Classification of Fingerprint Images. *Michigan State University (MSU) Technical Report*, Jan. 1998, MSUCPS: TR98-18

Hong, L.; Wan, Y. & Jain, A.K. (1998). Fingerprint Image Enhancement: Algorithm and Performance Evaluation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 20, No. 08, Aug. 1998, pp. 777–789, ISSN: 0162-8828

Huang, C-Y.; Liu, L-M. & Douglas Hung, D.C. (2007). Fingerprint Analysis and Singular Point Detection. *Pattern Recognition Letters*, Vol. 28, No. 04, Apr. 2007, pp. 1937–1945, ISSN: 0167-8655

Jain, A.K.; Prabhakar, S. & Hong, L. (1999). A Multichannel Approach to Fingerprint Classification. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 21, No. 04, Apr. 1999, pp. 348–359, ISSN: 0162-8828

Jain, A.K. & Minut, S. (2002). Hierarchical Kernel Fitting for Fingerprint Classification and Alignment, *Proceedings of the 16$^{th}$ International Conference on Pattern Recognition - Volume* 2, pp. 469–473, ISBN: 0-7695-1695-X, Quebec, Canada, Aug. 2002

Karu, K. & Jain, A.K. (1996). Fingerprint Classification. *Pattern Recognition*, Vol. 29, No. 03, Mar. 1996, pp. 389–404, ISSN: 0031-3203

Leonard, B. (1988). *Science of Fingerprints: Classification and Uses*, Diane Publishing Co., ISBN: 0-16-050541-0, Darby, Pennsylvania

Maltoni, D.; Maio, D.; Jain, A.K. & Prabhakar, S. (2009). *Handbook of Fingerprint Recognition*, Springer, ISBN: 978-84882-253-5, London, UK

Marwala, T. (2007). Bayesian Training of Neural Networks Using Genetic Programming. *Pattern Recognition Letters*, Vol. 28, No. 12, Dec. 2007, pp. 1452–1458, ISSN: 0167-8655

Mathekga, M.E. & Msiza, I.S. (2009). A Singular Point Detection Algorithm Based on the Transition Line of the Fingerprint Orientation Image, *Proceedings of the 20$^{th}$ Annual Symposium of the Pattern Recognition Association of South Africa*, pp. 01–06, ISBN: 978-0-7992-2356-9, Stellenbosch, South Africa, Nov. 2009

Msiza, I.S.; Leke-Betechuoh, B.; Nelwamondo, F.V. & Msimang, N. (2009). A Fingerprint Pattern Classification Approach Based on the Coordinate Geometry of Singularities, *Proceedings of the IEEE International Conference on Systems, Man & Cybernetics*, pp. 510–517, ISBN: 978-1-4244-2793-2, San Antonio, Texas, Oct. 2009

Msiza, I.S.; Leke-Betechuoh, B. & Malumedzha, T. (2011). Fingerprint Re-alignment: A Solution Based on the True Fingerprint Center Point, *Proceedings of the IEEE International Conference on Machine Learning & Computing – Vol.* 2, pp. 338–343, ISBN: 978-1-4244-9253-4, Little India, Singapore, Feb. 2011

Park, C.H. & Park, H. (2005). Fingerprint Classification Using Fast Fourier Transform and Nonlinear Discriminant Analysis. *Pattern Recognition*, Vol. 38, No. 04, Apr. 2005, pp. 495–503, ISSN: 0031-3203

Ratha, A.K.; Chen, S. & Jain, A.K. (1995). Adaptive Flow Orientation-Based Feature Extraction in Fingerprint Images. *Pattern Recognition*, Vol. 28, No. 11, Nov. 1995, pp. 1657–1672, ISSN: 0031-3203

Senior, A. (1997). A Hidden Markov Model Fingerprint Classifier, *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems, & Computers*, pp. 306–310, ISBN: 0-8186-8316-3, Pacific Grove, California, Nov. 1997

Senior, A. (2001). A Combination Fingerprint Classifier. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 23, No. 10, Oct. 2001, pp. 1165–1174, ISSN: 0162-8828

Wilson, C.L.; Candela, G.T.; Grother, P.J.; Watson, C.I. & Wilkinson R.A. (1992). Massively Parallel Neural Network Fingerprint Classification System. *National Institute of Standards and Technology (NIST) Technical Report*, Jan. 1992, NISTIR 4880

Yao, Y.; Marcialis, G.; Pontil, M.; Frasconi, P. & Roli, F. (2003). Combining Flat and Structured Representations for Fingerprint Classification with Recursive Neural Networks and Support Vector Machines. *Pattern Recognition*, Vol. 36, No. 02, Feb. 2003, pp. 397–406, ISSN: 0031-3203

# Part 2

## Face Recognition

# Biologically Inspired Processing for Lighting Robust Face Recognition

Ngoc-Son Vu and Alice Caplier
*Grenoble Institute of Technology*
*France*

## 1. Introduction

Due to its wide variety of real-life applications, ranging from user-authentication (access control, ATM) to video surveillance and law enforcements, face recognition has been one of the most active research topics in computer vision and pattern recognition. Also, it has obvious advantages over other biometric techniques, since it is natural, socially well accepted, and notably *non-intrusive*. In reality, several reliable biometrics authentication techniques are available and widely used nowadays (such as iris or fingerprint), but they mostly rely on an active participation of the user. On the contrary, facial biometric demands very little cooperation from the user; thanks to this user-friendly capability, face recognition is said to be non-intrusive.

Over the last decades, significant progress has been achieved in face recognition area. Since the seminal work of Turk and Pentland (Turk & Pentland, 1991), where the Principal Component Analysis (PCA) is proposed to apply to face images (Eigenfaces), more sophisticated techniques for face recognition appear, such as Fisherfaces (Belhumeur et al., 1997), based on linear discriminant analysis (LDA), Elastic Bunch Graph Matching (EBGM) (Wiskott et al., 1997), as well as approaches based upon Support Vector Machines (SVM) (Phillips, 1999), or Hidden Markov Models (HMM) (Nefian & III, 1998; Vu & Caplier, 2010b), etc.

Nevertheless, face recognition, notably under *uncontrolled scenarios*, remains active and unsolved. Among many factors affecting the performance of face recognition systems, illumination is known to be one of the most significant. Indeed, it was proven, both theoretically (Moses et al., 1994) and experimentally (Adini et al., 1997) that image variation due to lighting changes is more significant than that due to different personal identities. In other words, the difference between two face images of the same individual taken under varying lighting conditions is larger than the difference between any two face images taken under the same lighting conditions, as illustrated in Fig. 1.

Inspired by the great ability of human retina that enables the eyes to see objects in different lighting conditions, we present in this chapter a novel method of illumination normalization by simulating the performance of its two layers: the photoreceptors and the outer plexiform layer. Thus, we say the algorithm biologically inspired.

The rest of the chapter is structured as follows: Section 2 briefly discusses the related work; Section 3 presents the model of retinal processing and its advantage. In Section 4, the

Fig. 1. Face appearance varies significantly due to different lighting conditions: (left) face images of two people taken under the same lighting conditions; (right) two face images of the same individual taken under varying lighting conditions.

proposed method is described in detail. Experimental results are presented in Section 5, and conclusion is finally given in Section 6.

## 2. Related work

It is possible to deal with problems of illumination at three different stages in the pipeline of face recognition: during the preprocessing, the feature extraction and the classification. Therefore, existing methods are usually divided into the three following categories:

### 2.1 Illumination invariant feature extraction

The methods of this category try to extract image features which are invariant to illumination changes. It was shown theoretically in (Moses et al., 1994) that in the general case there are no functions of images that are illumination invariant. In (Adini et al., 1997), the authors empirically showed that classical image representations such as edge maps, derivatives of the gray level as well as the image filtered with 2D Gabor-like functions are not sufficient for recognition task under a wide variety of lighting conditions. This observation was later formally proved in (Chen et al., 2000), where the authors showed that for any two images, there is always a family of surfaces, albedos and light sources that could have produced them. Although more recent work, such as Local Binary Patterns (LBP) (Ahonen et al., 2004), Patterns of Oriented Edge Magnitudes (POEM) (Vu & Caplier, 2010a), reveals that certain features are less sensitive to lighting conditions, face recognition based on feature extraction only performs not reliably enough under extreme lighting variations.

### 2.2 Illumination modeling

These approaches require a training set containing several images of the same individual under varying illumination conditions. A training phase is then performed so as to derive a *model* for every identy, which will be used for recognition task. Examples are Illumination Cone (Belhumeur & Kriegman, 1998), Spherical Harmonics (Basri & Jacobs, 2003). Although providing the high quality results in general, these algorithms are costly and in particular they require several images obtained under different lighting conditions for each individual to be recognized. They are therefore impractical for many applications, such as surveillance where there is strict constraint upon the computational time or face recognition in one sample circumstances.

### 2.3 Suppression of illumination variation

The most suitable choice is to deal with lighting variation during the preprocessing step, prior to other stages. Such algorithms transform the image to a canonical form where the illumination variation is erased. Classical algorithms such as histogram equalization, gamma

correction are simple examples whereas the more elaborated techniques are mostly based on properties of the human visual system, evidenced in (Land & McCann, 1971), known as the Retinex theory.

The Retinex theory aims to describe how the human visual system perceives the color/lightness of a natural scene. Our vision ensures that the perceived color/lightness of objects remains relatively constant under varying illumination conditions. This feature helps us identifying objects. Physics says applying red light on a green apple is not the same as applying white light on the same green apple, but our vision attempts to see the same color, regardless of the applied light. The goal of Land's Retinex theory was thus to understand and to find a computational model of how our vision system process the physical stimuli in such a way that color/lightness consistancy is preserved (in the remainder, only the gray images are considered since face recognition techniques perform well on gray images). Assuming that the intensity signal $I(x, y)$ is the product of the illumination $L(x, y)$ and the surface reflectance $R(x, y)$, i.e. $I(x, y) = L(x, y)R(x, y)$, the authors supposed that the reflectance value of a pixel $R(x, y)$ can be computed by taking the ratio of the pixel intensity with the illumination. The problem of obtaining $R$ from an input image $I$ can be solved therefore by estimating $L$. Using this observation, several methods have been presented, such as Single Scale Retinex (SSR), Multi Scale Retinex (MSR) (Jobson et al., 1997) as well as Self-Quotient Image (SQI) (Wang et al., 2004).

Actually, these algorithms are widely used for illumination normalization and also reach the state-of-the-art results. However, they still can not exactly estimate $L$, so large illumination variations are not completely removed. Another disadvantage of those algorithms is that the computational time is still relatively high: both MSR and SQI are "multi-scale" methods which require to estimate the illumination at various "scales". It is also worth noting that the term Retinex coming from the words "Retina" and "Cortex", meaning that both the eyes and the brain are involved in the process. However, to the best of our knowledge, the rule of brain is rather to build a visual representation with vivid details, whereas the natural properties of retina allow our eyes to see and to identify objects in different lighting conditions. That is the motivation for our retina based illumination normalization method.

Before going into details of our retina filter, we need to distinguish the difference between the method proposed in (Tan & Triggs, 2007), referred as PS in the follows, and ours. Although both consist of three steps (see (Tan & Triggs, 2007) for details of the PS method and Section 4 for ours), algorithms used in each step (except the second stage) are different. In their work, the authors do not point out that their algorithm is basically based on the performance of retina. Moreover, we will show that our algorithm is both more efficient and of lower complexity. It is also worth noting that our Gipsa-lab is one of pioneer laboratories on modeling the behavior of the retina, such as (Beaudot, 1994).

## 3. Retina: properties and modeling

The retina lies at the back of the eye (Fig. 2). Basically, it is made of three layers: the photoreceptor layer with cones and rods; the outer plexiform layer (OPL) with horizontal, bipolar and amacrine cells; and the inner plexiform layer (IPL) with ganglion cells. The goal here is not to precisely model the dynamics of retinal processing, such as is done, for example, in (Benoit, 2007). We aim at identifying which processing acts on the retinal signal for illumination normalization. This section demonstrates that bipolar cells not only remove illumination variations and noise but also enhance the image edges. It is worth noting that the
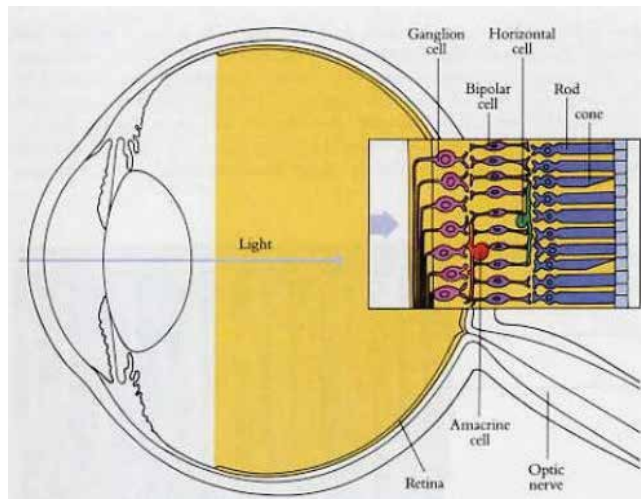
Fig. 2. The retina lies at the back of the eye. Light passes through the bipolar and amacrine cells and reaches the photoreceptors layers where it returns [http://hubel.med.harvard.edu/bio.htm].

retina is capable to process both spatial and temporal signals but working on static images, we consider only the spatial processing in the retina.

### 3.1 Photoreceptors: light adaptation filter

Rods and cones have quite different properties: rods have the ability to see at night, under conditions of very low illumination (night vision) whereas cones have the ability to deal with bright signals (day vision). In other words, the photoreceptors are is able to adjust the dynamics of light intensity they receive: it plays a crucial rule as light adaptation filter. This property is also called the *adaptive* or *logarithmic compression*.

To exploit and mimic this property, an adaptive nonlinear function is usually applied on the input signal (Benoit, 2007):

$$y = \frac{x}{x + x_0},\tag{1}$$

where $x$ represents the input light intensity, $x_0$ is the adaptation factor, and $y$ is the adapted signal.

Fig. 3 illustrates the adaptive nonlinear function for different values of $x_0$. If $x_0$ is small, the output has increased sensitively, otherwise when $x_0$ is large, there is not much change in sensitivity.

For an automatic operator, several methods are proposed to determine the adaptation factor $x_0$. One solution is to take $x_0$ equal to the average image intensity. This works if the image intensity is roughly balanced, meaning that the histogram is relatively flat around the mean value. However, when image regions are not lighted similarly, such a function will equalize those image regions identically (see Fig. 4(b)).

Therefore, $x_0$ should vary for each pixel. It can be obtained by applying a low pass filtering on the input image (lighting adapted image using these factors are shown in Fig. 4(c)). Another solution is to combine these two approaches: a low-pass filtering is applied on the input image and for each pixel, the adaptation factor is the sum of the image average intensity and the
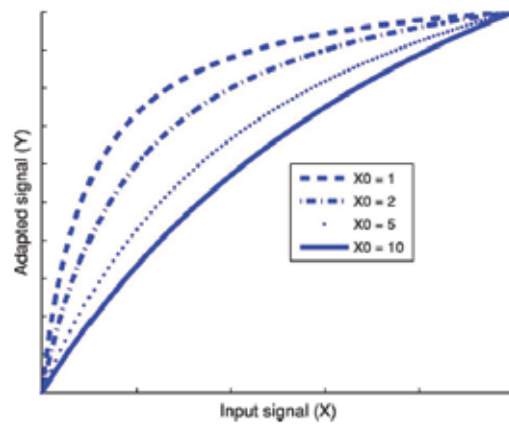
Fig. 3. Performance of nonlinear operations with different adaptation factors $x_0$.
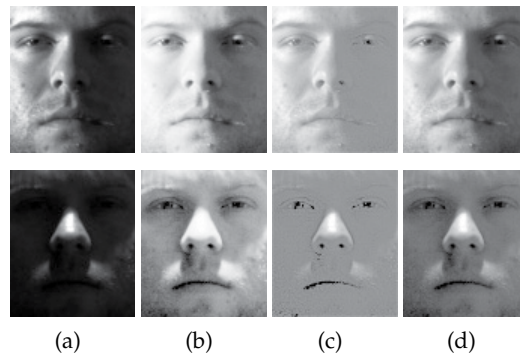


|  (a)  |  (b)  |  (c)  |  (d)  |

Fig. 4. (a): original images; images obtained with adaptation factor equal to: (b) the average of image intensity; (c): intensity of low-pass filtered image; (d) the sum of both (b) and (c).

intensity of the low-pass filtered image. Resulting images in Fig. 4 show that after applying the adaptive operators, the local dynamic range in dark regions are enhanced whilst bright regions remain almost unchanged. Among the images (b),(c) and (d), the images (d) are the best lighting adapted. Consequently, this combinational approach will be used in our model.

### 3.2 Outer Plexiform Layer (OPL)

Photoreceptors perform not only as a light adaptation filter but also as a low pass filter. This leads to an image in which the high frequency noise is strongly attenuated and the low frequency visual information is preserved. The signal is then transmitted and processed by horizontal cells which acts as a second low pass filter. Bipolar cells calculate the difference between photoreceptor and horizontal cell responses, meaning that bipolar cells act as a band pass filter: the high frequency noise and low frequency illumination are removed.

To model the behavior of bipolar cells, two low pass filters with different cutoff frequencies corresponding to performance of photoreceptors and horizontal cells (the cutoff frequency of horizontal cells is lower than that of photoreceptors) are often used, and then the difference of these responses is calculated. In our algorithm, two Gaussian low pass filters with different standard deviations corresponding to the effects of photoreceptors and horizontal cells are

used, and bipolar cells act like a Difference of Gaussians filter (*DoG*). As can be seen from Fig. 5, the very high and very low frequencies are eliminated whilst the middle ones are preserved. Note that, another advantage of the DoG filter is the enhancement of the image edges, which is believed useful for recognition task.
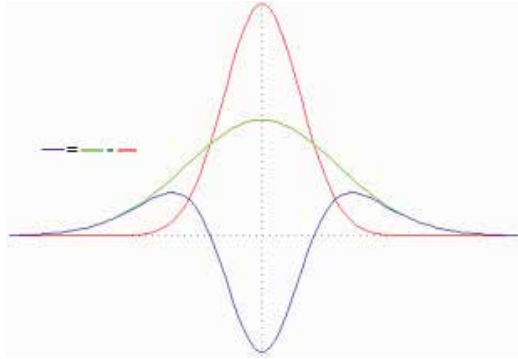


Fig. 5. Difference of Gaussians filter.

### 3.3 Inter Plexiform Layer (IPL)

In this last processing stage of the retina before the optic nerve, the information obtained through the processing of OPL is processed by the ganglion and amacrine cells. However, this layer rather deals with temporal information or movement and therefore is not related to this work.

## 4. Proposed method in detail

As pointed out above, a model with a nonlinear operator and a band pass filter can be used for illumination variation removal. In our model, *multiple* consecutive nonlinear operations are used for a more efficient light adaptation filter. Also, a truncation is used after the band pass filter to enhance the global image contrast.

### 4.1 Multiple logarithmic compressions

In (Meylan et al., 2007), being interested in the property of light adaptation of the retina, the authors modeled the behavior of the *entire* retina by two adaptive compressions, which correspond to the effects of OPL and IPL, respectively. By experiments, they showed that these duplex operations lead to a very good light adaptation filter with a good visual discrimination. Inspired by this observation, we propose to apply several adaptive operations in the first step of our model. In reality, (Vu & Caplier, 2009) already pointed out that using two consecutive adaptive functions leads to the optimal performance on the Yale B database (when images with the most neutral light sources are used ad reference). For an algorithm with generality, this work will automatically determines the optimal number of compressions.

The adaptation factor ($x_0$ in Equation 1) of the first nonlinear function is computed as the sum of the average intensity of the input image and the intensity of the low pass filtered image:

$$F_1(p) = I_{in}(p) * G_1 + \frac{\overline{I_{in}}}{2} \tag{2}$$

where $p = \{x, y\}$ is the current pixel; $F_1(p)$ is the adaptation factor at pixel $p$; $I_{in}$ is the intensity of the input image; $*$ denotes the convolution operation; $\overline{I_{in}}$ is the mean value of the input; and $G_1$ is a 2D Gaussian low pass filter with standard deviation $\sigma_1$:

$$G_1(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2 + y^2}{2\sigma_1^2}} \tag{3}$$

The input image is then processed according to Equation 1 using the adaptation factor $F_1$, leading to $I_{la_1}$ image:

$$I_{la_1}(p) = (I_{in}(max) + F_1(p)) \frac{I_{in}(p)}{I_{in}(p) + F_1(p)} \tag{4}$$

The term $I_{in}(max) + F_1(p)$ is a normalization factor where $I_{in}(max)$ is the maximal value of the image intensity.

The second nonlinear function works similarly, the light adaptation image $I_{la_2}$ is obtained by:

$$I_{la_2}(p) = (I_{la_1}(max) + F_2(p)) \frac{I_{la_1}(p)}{I_{la_1}(p) + F_2(p)} \tag{5}$$

with

$$F_2(p) = I_{la_1}(p) * G_2 + \frac{\overline{I_{la_1}}}{2} \tag{6}$$

and

$$G_2(x, y) = \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2 + y^2}{2\sigma_2^2}} \tag{7}$$

When using more compressions, the next operations work in the same way and we finally get the image $I_{la_n}$ where $n$ is the number of nonlinear operators used.



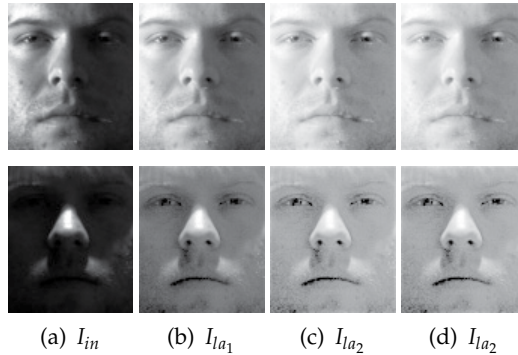(a) $I_{in}$    (b) $I_{la_1}$    (c) $I_{la_2}$    (d) $I_{la_2}$

Fig. 6. Performance of photoreceptors with different parameters. (a): original images; (b): images after one adaptive operator; (c) & (d): images after two operators with different parameters.

Fig. 6 shows the effect of the adaptive nonlinear operator on two images with different parameters. Visually, we observe that the images after two operations (Fig. 6(c),(d)) are

better adapted to lighting than those after a single operation (Fig. 6(b)). Another advantage of two consecutive logarithmic compressions, as argued in (Meylan et al., 2007), is that the resulting image does not depend on low pass filter parameters: we see any difference between the images in Fig. 6(c) ($\sigma_1 = 1, \sigma_2 = 1$) and those in Fig. 6(d) ($\sigma_1 = 1, \sigma_2 = 3$).

### 4.2 Difference of Gaussians filter and truncation

The image $I_{la_n}$ is then transmitted to bipolar cells and processed by using a Difference of Gaussians ($DoG$) filter:

$$I_{bip} = DoG * I_{la_n} \tag{8}$$

where $DoG$ is given by:

$$DoG = \frac{1}{2\pi\sigma_{Ph}^2} e^{-\frac{x^2+y^2}{2\sigma_{Ph}^2}} - \frac{1}{2\pi\sigma_H^2} e^{-\frac{x^2+y^2}{2\sigma_H^2}} \tag{9}$$

The terms $\sigma_{Ph}$ and $\sigma_H$ correspond to the standard deviations of the low pass filters modeling the effects of photoreceptors and horizontal cells.

In fact, the output image at bipolar cells $I_{bip}$ is the difference between the output image at the photoreceptors $I_{Ph}$ and that at horizontal cells $I_H$: $I_{bip} = I_{Ph} - I_H$, where $I_{Ph}$ and $I_H$ are obtained by applying low pass Gaussian filters on the image $I_{la_2}$: $I_{Ph} = G_{Ph} * I_{la_n}, I_H = G_H * I_{la_n}$.



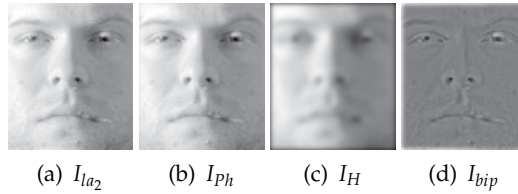(a) $I_{la_2}$          (b) $I_{Ph}$          (c) $I_H$          (d) $I_{bip}$

Fig. 7. Effect of the Difference of Gaussians filter: $I_{bip} = I_{Ph} - I_H$. (a) image after two non-linear operators; (b) image after by the $1^{st}$ low pass filter at photoreceptors; (c) image after by the $2^{nd}$ low pass filter at horizontal cells; (d) output image at bipolar cells.

Fig. 7 shows the effect of Difference of Gaussians filter on the output image of two adaptive operators (Fig. 7(a)). We observe that the illumination variations and noise are well suppressed (Fig. 7(d)).

A drawback of the DoG filter is to reduce the inherent global contrast of the image. To improve image contrast, we further propose to remove several extreme values by truncation. To facilitate the truncation, we first use a zero-mean normalization to rescale the dynamic range of the image. The substraction of the mean $\mu_{I_{bip}}$ is not necessary because it is near to 0.

$$I_{nor}(p) = \frac{I_{bip}(p) - \mu_{I_{bip}}}{\sigma_{I_{bip}}} = \frac{I_{bip}(p)}{\sqrt{E(I_{bip}^2)}} \tag{10}$$

After normalization, the image values are well spread and are mainly lied around 0, some extreme values are removed by a truncation with a threshold $Th$ according to:

$$I_{pp}(p) = \begin{cases} max(Th, |I_{nor}(p)|) & \text{if } I_{nor}(p) \geq 0 \\ -max(Th, |I_{nor}(p)|) & \text{otherwise} \end{cases} \tag{11}$$

The threshold *Th* is selected in such a way that the truncation can remove approximately 2-4% extreme values of image. Fig. 8 shows the main steps of the algorithm. We observe that after the truncation, the overall contrast of the image is improved (Fig. 8(d)). As properties, the proposed algorithm not only removes the illumination variations and noise, but also reinforces the image contours.



(a) $I_{in}$          (b) $I_{la_2}$          (c) $I_{bip}$          (d) $I_{pp}$
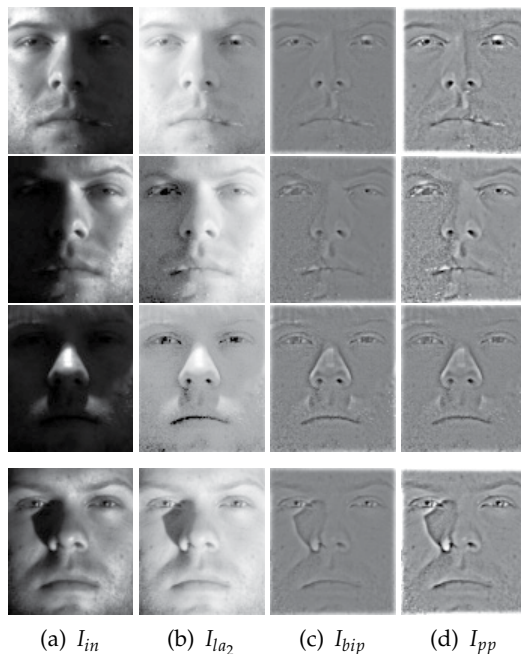
Fig. 8. Effects of different stages of the proposed algorithm. (a) input images; (b) images after the light adaptation filter; (c) output images at the bipolar cells; (d) final processed images.

## 5. Experimental results

### 5.1 Experiment setting

The performance of the proposed preprocessing algorithm is evaluated regarding face recognition application. Three recognition methods are considered, including Eigenface (Turk & Pentland, 1991), the Local Binary Patterns (LBP) (Ahonen et al., 2004) based and the Gabor filter based (Liu & Wechsler, 2002) methods:

- Although the Eigenface method is very simple (many other methods lead to better recognition rate), the recognition results obtained by blending our preprocessing method with this recognition technique is very interesting because they show the effectiveness of our algorithm.

- Both LBP and Gabor features are considered to be robust to lighting changes, we report the recognition results when combining our preprocessing algorithm with these illumination robust feature based techniques in order to show that such features are not sufficient for wide variety of lighting changes and that our illumination normalization method improves significantly the performance of these methods.

The simple nearest neighbor classification is used to calculate the identification rates in all tests. Experiments are conducted on three databases: Extended Yale B, FERET (frontal faces) and AR, which are described in the rest of this section.

### 5.1.1 Extended Yale B database

The Yale B face dataset (Georghiades & Belhumeur, 2001) containing 10 people under 64 different illumination conditions has been a *de facto* standard for studying face recognition under variable lighting over the past decade. It was recently updated to the Extended Yale B database (Lee et al., 2005), containing 38 subjects under 64 different lighting conditions. In both cases the images are divided into five subsets according to the angle between the light source direction and the central camera axis ($0°$–$12°$; $13°$–$25°$; $26°$–$50°$; $51°$–$77°$; $++78°$) (c.f. Fig. 9). Although containing few subjects and little variability of expression, aging, the extreme lighting conditions of the Extended Yale B database still make it challenging for most face recognition methods.

Normally, the images with the most neutral light sources (named "A+000E+00" and an example is shown in the first image of Fig. 9) are used as reference, and all other images are used as probes. But in this work, we conduct more difficult experiments where reference database also contains images of *non-neutral* light sources. We use face images already aligned by the authors and then resize them which are originally of $192 \times 168$ pixels to $96 \times 84$ pixels.



(a) Subset 1 ($0°$–$12°$)          (b) Subset 2 ($13°$–$25°$)          (c) Subset 3 ($26°$–$50°$)



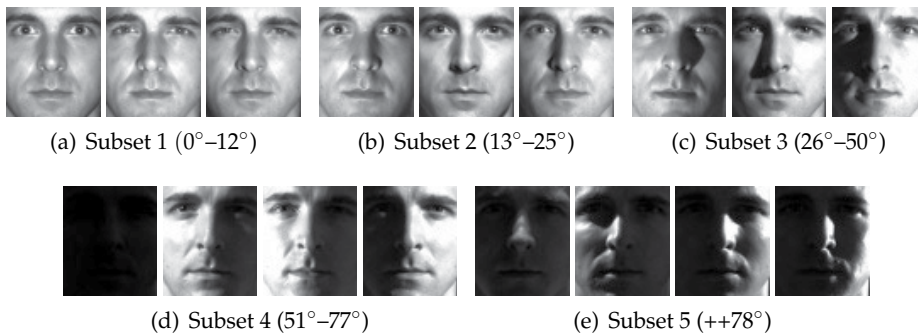(d) Subset 4 ($51°$–$77°$)          (e) Subset 5 ($++78°$)

Fig. 9. Example images from the Extended Yale B database. Images from this challenging database is divided into 5 subsets according to the angle between the light source direction and the central camera axis. (a) Subset 1 with angle between $0°$ and $12°$; (b) Subset 2 with angle between $13°$ and $25°$; (c) Subset 3 with angle between $26°$ and $50°$; (d) Subset 4 with angle between $51°$ and $77°$; (e) Subset 5 with angle larger than $78°$. Example of filtered images can be seen in Fig. 8.

### 5.1.2 FERET database

In the FERET database (Phillips et al., 2000), the most widely adopted benchmark for the evaluation of face recognition algorithms, all frontal face pictures are divided into five categories: Fa, Fb, Fc, Dup1, and Dup2 (see example images in Fig. 10). Fb pictures were taken at the same day as Fa pictures and with the same camera and illumination condition. Fc pictures were taken at the same day as Fa pictures but with different cameras and illumination. Dup1 pictures were taken on different days than Fa pictures but within a year. Dup2 pictures were taken at least one year later than Fa pictures. We follow the standard FERET tests, meaning that 1196 Fa pictures are gallery samples whilst 1195 Fb, 194 Fc, 722 Dup1, and

234 Dup2 pictures are named as Fb, Fc, Dup1, and Dup2 probes, respectively. As alignment, thanks to the available coordinates of eyes, facial images are geometrically aligned in such a way that centers of the two eyes are at fixed positions and images are resized to $96 \times 96$ pixels.



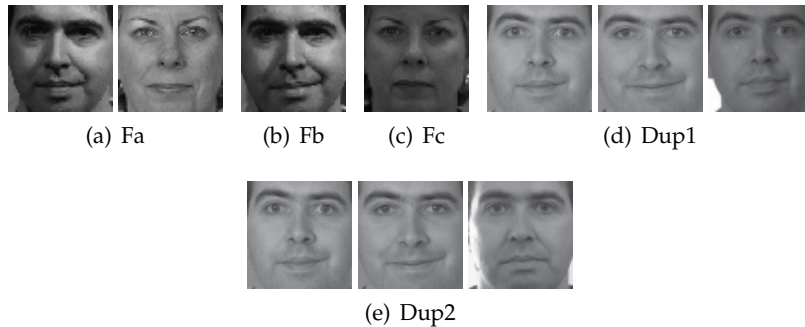(a) Fa          (b) Fb          (c) Fc          (d) Dup1



(e) Dup2

Fig. 10. Example images of the FERET database. (a): Fa pictures with neutral expressions; (b): Fb pictures taken at the same day as Fa pictures and with the same camera and illumination condition; (c): Fc pictures taken at the same day as Fa pictures but with different cameras and illumination; (d): Dup1 pictures were taken on different days than Fa pictures but within a year; (e): Dup2 pictures were taken at least one year later than Fa pictures.

### 5.1.3 AR database

The AR database  (Martinez & Benavente, 1998) contains over 4000 mug shots of 126 individuals (70 men and 56 women) with different facial expressions, illumination conditions and occlusions. Each subject has up to 26 pictures in two sessions. The first session, containing 13 pictures, named from "AR–01" to "AR–13", includes neutral expression (01), smile (02), anger (03), screaming (04), different lighting (05 - 07), and different occlusions under different lighting (08 - 13).  The second session exactly duplicates the first session two weeks later. We used 126 "01" images, one from each subject, as reference and the other images in the first section as probes.  In total, we have 12 probe sets, named from "AR–02" to "AR–13". The AR images are cropped and aligned in a similar way as images in the FERET database. Fig.  11 shows example images from this dataset whereas the images shown in Fig.  12 are preprocessed images, from the probe set "AR–07", containing images of two side lights on.

### 5.2 Parameter selection

This section considers how the parameters effect to the filter performance. Parameters varied include the number of compressions and the standard deviations associated ($\sigma_1, \sigma_2$ in the case of two compressions), the standard deviations $\sigma_P, \sigma_H$ and the threshold $Th$. As recognition algorithm, in this section, we use the simple Eigenface method associated and the cosine distance. The Yale B database containing images of 10 different people is used.

### 5.2.1 Number of compressions et parameters

The compression number $n$ used in the first step is turned variable whilst other parameters are fixed ($\sigma_P = 0.5, \sigma_H = 3.5$, et $Th = 4$). Nearly eight hundred tests were carried out:

1.  $n = 1, 2, 3, 4$. In the follows, the corresponding filters are denoted $\mathbf{F}^1, \mathbf{F}^2, \mathbf{F}^3, \mathbf{F}^4$.

(a) 1       (b) 2-4       (c) 5-7
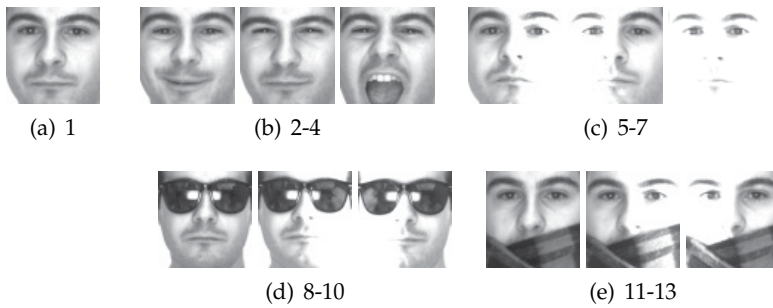
(d) 8-10       (e) 11-13

Fig. 11. Example images from the AR database. (a) image of neutral expression; (b) images of different expressions: smile (02), anger (03), screaming (04); (c) images of different lighting conditions (05 - 07); (d) & (e) images of different occlusions under different lighting (08 - 13).
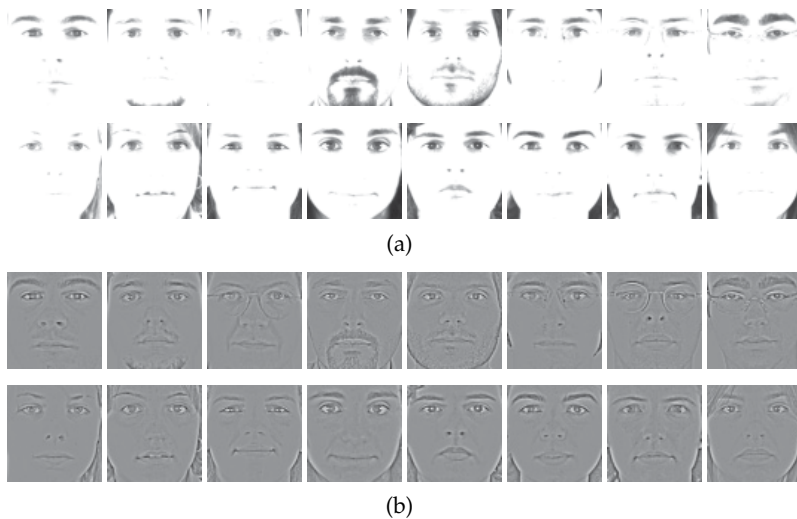


(a)

(b)

Fig. 12. Examples of the AR-07 subset: (a) original images; (b) processed images.

2. For each $n$, we vary the standard deviations $\sigma = 1, 2, 3$ (those are used to calculate the adaptation factors). The corresponding filters are referred as $\mathbf{F}^n_{(\sigma_1,...,\sigma_n)}$. In total, 12 filters (3 for each n) are considered.

3. For each filter, 64 different experiments are carried out: for a given illumination angle, 10 images of 10 people in this angle are used as reference and the rest of the database is used as test.

4. We then calculate the average of results obtained across the subset to which reference images belong.

Fig. 13 shows the average recognition rates obtained when the reference images belong to different subsets with different filters (for clarity, we depict only the results of 7 filters). On the horizontal axis of this figure, $(i_1, i_2, ..., i_n)$ means $\mathbf{F}^n_{(\sigma_1,...,\sigma_n)}$. For example, (1) corresponds to $\mathbf{F}^1$ with the standard deviation $\sigma_1 = 1$. It is clear that:

1. Multiple adaptive operations always lead to better compression rates than only one.

| $\sigma_{Ph}, \sigma_H$ | Sous-ensembles | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 0.5 & 3 | 100 | 98.3 | 99.8 | 98.6 | 100 |
| 0.5 & 3.5 | 100 | 98.3 | 99.4 | 97.9 | 100 |
| 1 & 3.5 | 100 | 98.3 | 99.0 | 96.8 | 99.7 |
| 1 & 4 | 99.8 | 97.9 | 96.4 | 95.7 | 99.3 |

Table 1. Recognition rates on the Yale B database for different $\sigma_{Ph}$ & $\sigma_H$.

2. Regarding the performance of multiple compressions, all $\mathbf{F}^2$, $\mathbf{F}^3$, and $\mathbf{F}^4$ perform very well.

3. The performance of $\mathbf{F}^n$ ($n = 2, 3, 4$) is similar: *the values $\sigma_i$ are therefore not important.* In reality, the final results of $\mathbf{F}^3$ are slightly better than those of $\mathbf{F}^2$ and $\mathbf{F}^4$ with the differences of 0.2 and 0.3% respectively. But for complexity constraints, we use $\mathbf{F}^2$ with $\sigma_1 = \sigma_2 = 1$.

The proposed method produces very good results in all cases. When reference images belong to the first four subsets, the subset 1 is the easiest query; when the reference images belong to subset 5, the subset 5 is the easiest test. This is easy to understand. However, surprisingly, the subset 5 is often easier to process than subsets 2, 3, 4 (see Fig. 13 (b), (c) and (d)).

### 5.2.2 Parameters of DoG and truncation

*DoG* filter parameters are the two standard deviations $\sigma_{Ph}$ and $\sigma_H$ which define the low and high cutoff frequencies of the band pass filter, respectively (refer to Fig. 5). A critical constraint is $\sigma_{Ph} < \sigma_H$. In this work, we choose the values $\sigma_{Ph} \in \{0.5; 1\}$, $\sigma_H \in \{3; 4\}$ [1]. By varying these values in corresponding intervals, we find that $\sigma_{Ph} = 0.5$ and $\sigma_H = 3$ give better results than others. Table 1 shows the average rates obtained when the reference images belong to subset 3.

Regarding the truncation threshold *Th*, we first analyze the distribution of image values. Remind that these values are mainly lied around 0 and their average is 0. By choosing randomly 20 images $I_{nor}$, we find that with a threshold $Th \in \{3, 4\}$, we can remove in average 3-4% extreme values per image. We then evaluate the effect of the retina filter by varying *Th* in this range and we observe that the obtained results are almost similar. However, if the truncation is not applied, the recognition rate is degraded about 1-2%, depending on the subsets. This shows the effectiveness of the truncation.

**The optimal parameters are:** $\sigma_1 = \sigma_2 = 1, \sigma_{Ph} = 0.5, \sigma_H = 3$, and $Th = 3.5$.

### 5.3 Results on the Extended Yale B database

In the follows, we compare the performance of our method with the state of the art methods. Thanks to the "INface tool" software [2], we have codes in Matlab of several illumination normalization methods. Among the available methods, we consider the most representative methods, such as MSR, SQI, and PS [3]. Parameters are used as recommended by the authors.

Table 2 presents results obtained on the Extended Yale B dataset when the reference set contains images acquired under ideal lighting condition (frontal lighting with angle 0°) and the test contains the rest of database. The reported results are divided into two groups: ones

---

[1] The cutoff frequencies should depend on the quality of images. With a blurred image whose information lies mainly in low frequency, applying a filter with $\sigma_{Ph}$ "too high" will cause a lost of information. For a fully automatic parameter choice, a quality metric images should be used: $\sigma_{Ph}$ and $\sigma_H$ should be chosen in such a way that not two much information of image is removed.

[2] http://uni-lj.academia.edu/VitomirStruc

[3] The code for this method is available from http://parnec.nuaa.edu.cn/xtan/

(a) Images in set 1 as reference

(b) Images in set 2 as reference

(c) Images in set 3 as reference

(d) Images in set 4 as reference

(e) Images in set 5 as reference

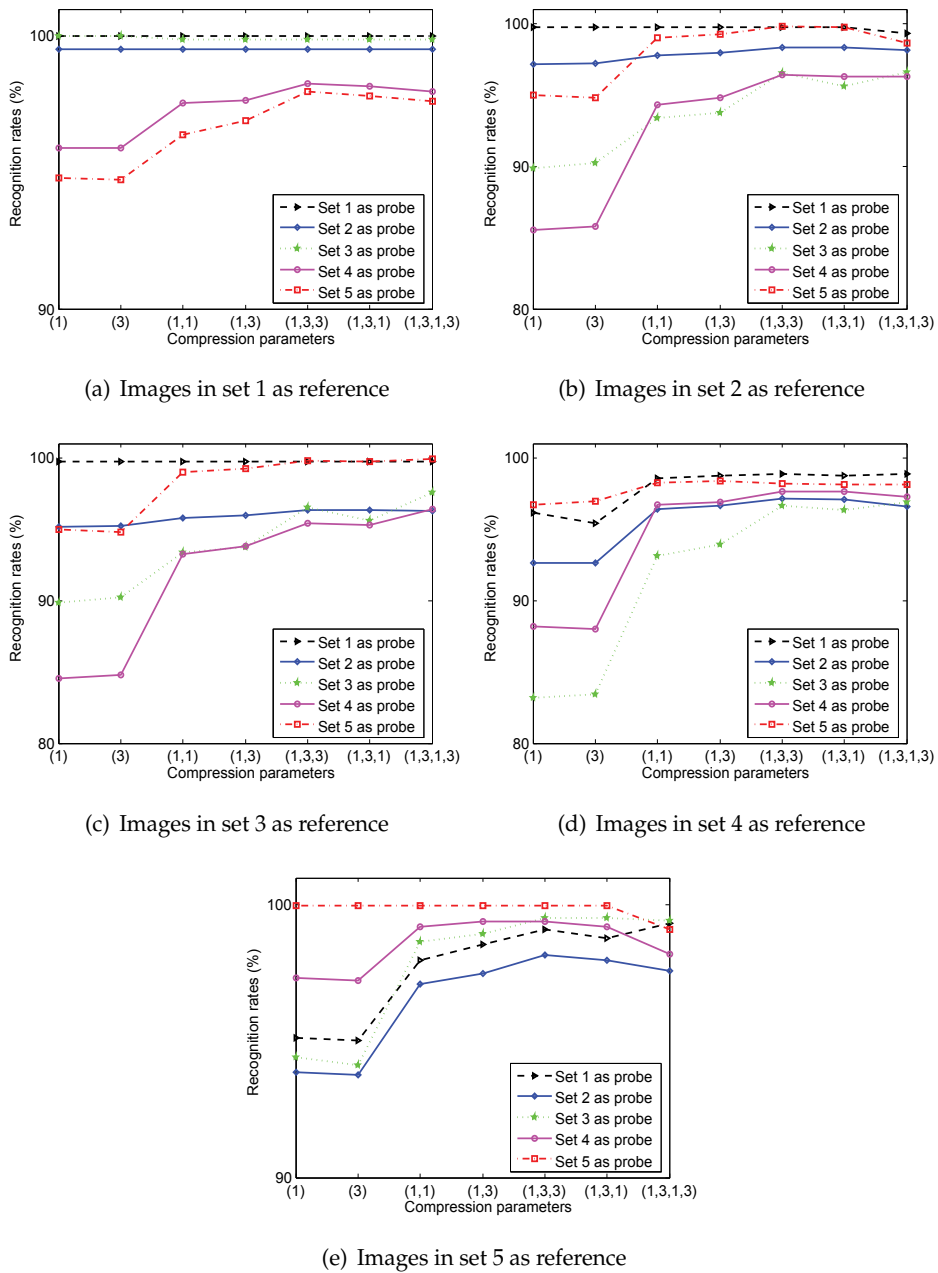Fig. 13. Recognition rates on the Yale B database for different adaptive operations with different parameters.

| | Methods | Subsets | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 10 individuals | Without preprocessing | 100 | 98.3 | 64.2 | 32.9 | 13.7 |
| | Histogram equalization (HE) | 100 | 98.3 | 65.8 | 35 | 32.6 |
| | MSR | 100 | 100 | 96.7 | 85 | 72.1 |
| | SQI (Wang et al., 2004) | 100 | 100 | 98.3 | 88.5 | 79.5 |
| | PS (Tan & Triggs, 2007) | 100 | 100 | 98.4 | 97.9 | 96.7 |
| | LTV (Chen et al., 2006)[+] | 100 | 100 | 100 | 100 | 100 |
| | **Retina filter** | **100** | **100** | **100** | **100** | **100** |
| | *Cone-cast (Georghiades & Belhumeur, 2001)** | *100* | *100* | *100* | *100* | *-* |
| | *Harmonic image (Basri & Jacobs, 2003)** | *100* | *100* | *99.7* | *96.9* | *-* |
| 38 individuals | HE | 98.9 | 97.6 | 56.5 | 23.6 | 21.4 |
| | MSR | 100 | 100 | 96.7 | 79.5 | 65.7 |
| | SQI | 100 | 99.8 | 94.0 | 85.5 | 77.0 |
| | PS | 100 | 99.8 | 99.3 | 99.0 | 96.6 |
| | **Retina filter** | **100** | **100** | **99.7** | **99.3** | **98.8** |

[+] *This method is about 500 times lower than ours.*
\* *These methods belong to the second category which aims at modeling the illumination. These methods require a training set of several images of the same individual under different lighting conditions (e.g. 9 images per person for the Cone cast method). The authors do not report on the subset 5.*

Table 2. Recognition rate obtained on the Extended Yale B when using the Eigenface recognition technique associated with different preprocessing methods and using frontal lighting images as reference.

obtained on the Yale B database containing only 10 individuals and the others obtained on the Extended Yale B database containing 38 individuals (in fact, researchers mostly conduct the experiments on the Yale B database). It can be seen from Table 2 that:

1. The recognition performance drops dramatically very significantly on subsets 4 and 5 when any preprocessing is used (the first row of the table).

2. The proposed method reaches the very strong results and outperforms all competing algorithms on both datasets. On the Yale B database, we obtain the perfect rates, even on the most challenging subset. The LTV method also performs very well but it is about 500 times slower than our algorithm.

We now consider more difficult tests when there are illumination variations on both reference and test images. 30 different tests are carried out on the Extended Yale B set. For each experiment, the reference database contains 38 images of 38 persons for a given angle of illumination and the test database contains all the rest (in fact, there are in total 64 different tests but we randomly choose 30 different lighting conditions: 5 conditions for each of the first four subsets and 10 for the subset 5). Presented in Table 3 are the average recognition rates which clearly show that the proposed method works very well even in challenging test where there are illumination variations on both the reference and probe images.

### 5.4 Results on the FERET database
The aim of this section is to prove the following advantages of our method:

1. It enhances the methods based on features which are considered to be robust to illumination variations.

| Method | MSR | SQI | PS | Proposed |
|--------|-----|-----|------|----------|
| Rate | 75.5 | 81.6 | 97.8 | 99.1 |

Table 3. Average recognition rates obtained on the Extended Yale B database when combining the simple Eigenface recognition technique with different preprocessing methods.

2. It improves the face recognition performance in all cases whether there are or not illumination variations on images.

To this end, two features being considered illumination invariant are used for representing face, including LBP (Ahonen et al., 2004) and Gabor wavelets (Liu & Wechsler, 2002). We follow the standard FERET evaluation protocol for reporting the performance: the subset Fa is reference whilst Fb, Fc, Dup1 & Dup2 subsets are probes.



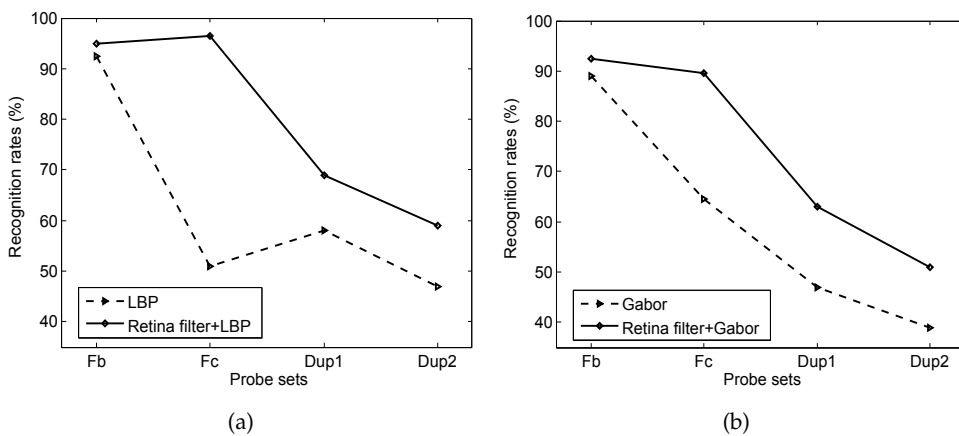(a)                                                                                                      (b)

Fig. 14. Performance of the retina filter on the FERET database when combining with different recognition methods: (a) LBP; (b): Gabor

Fig. 14 clearly shows that the retina preprocessing improves significantly the performance of the two considered methods on all four probes. The considerable improvements obtained on the Fc set confirm that when a good illumination normalization method is used in prior, the robustness of facial features is increased even if these features are considered to be robust to lighting changes. Regarding the results obtained on Fb, Dup1 & Dup2 sets, we can see that the retina filtering is useful for face recognition in all cases whether there are or not lighting variations on images. The reason is that our filter not only removes illumination variations but also enhances the image contours which are important cues for distinguishing individuals.

### 5.5 Results on the AR database

We repeat the similar experiments as in the previous section on the AR database. All 126 images "AR-01" (one for each individual) are used as reference. The recognition results are assessed on 12 probe sets, from "AR02" to "AR13", and are shown in Fig. 15. As can be seen from this figure, our method always leads to very good results. This again proves high efficiency of retinal filter.
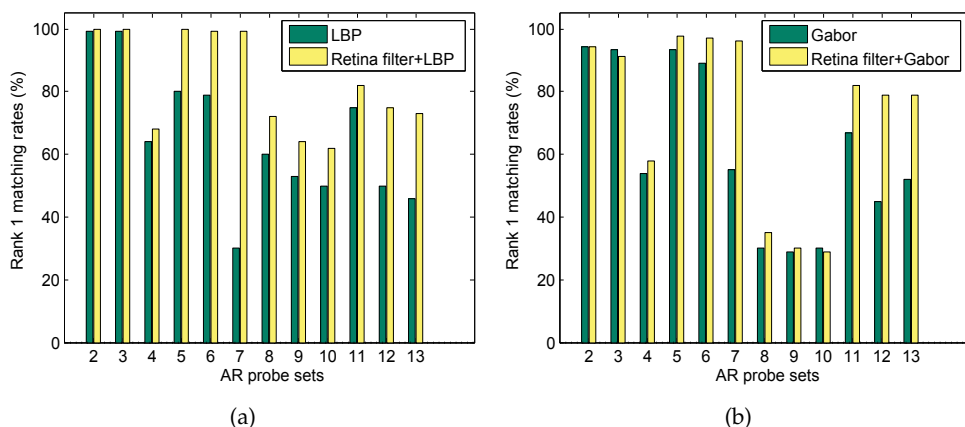
Fig. 15. Performance of the retina filter on the AR database when combining with different recognition methods:: (a) LBP; (b): Gabor

### 5.6 Computational time

This section compares the complexity of different illumination normalization methods to show the advantage of our algorithm. We consider the time required for processing 2000 images of $192 \times 168$ pixels and show the average time for one image in Table 4.

| Method | LTV$^+$ | SQI | MSR | PS | Proposed |
|--------|---------|-----|-----|-----|----------|
| Time (s) | 7.3$^+$ | 1.703 | 0.126 | 0.0245 | 0.0156 |

$^+$ In (Tan & Triggs, 2007), the authors showed that the LTV method is about 300 times slower than the PS method. We estimate therefore that this algorithm is about 500 times slower than ours.

Table 4. Time required to process an image of 192x168 pixels.

As can be seen from Table 4 or more visually from Figure 16, our method is of very low complexity. Using the code implemented in Matlab (on a desktop of Dual core 2.4 GHz, 2Gb Ram), we can process about 65 images of 192x168 pixels per second; our algorithm is a real-time one. Our method is about 1.57 times faster than the PS method and significantly faster than the others. It maybe worth noting that the most consuming stages in our method are convolutions. Let $mn$ be the image size, $w^2$ the mask size. To reduce the complexity, instead of directly using a 2D mask, which leads to the complexity of $O(mn \times w^2)$, we can use two successive 1D convolutions, leading to a *linear complexity* $O(mn \times 2w)$. In the model, $w = 3\sigma$ where $\sigma$ is the standard deviation of the Gaussian filter. As we use small deviations ($\sigma_1 = \sigma_2 = 1$), the computational time of convolutions is not important. On the contrary, the standard deviations in MSR and SQI are much bigger (e.g. $\sigma_1 = 7, \sigma_2 = 13, \sigma_3 = 20$). That is the reason why our method is very fast.

### 5.7 Illumination normalization for face detection

This section shows another application of our retina filter, i.e. lighting normalization for face detection. Regarding the effects of each step of the proposed model, we observe that using the output image at photoreceptor layer (light adaptation filter) may improve well the face detection performance. For validation, we use the face detector proposed in (Garcia & Delakis, 2004) and calculate the face detection rates of 640 original images in the Yale
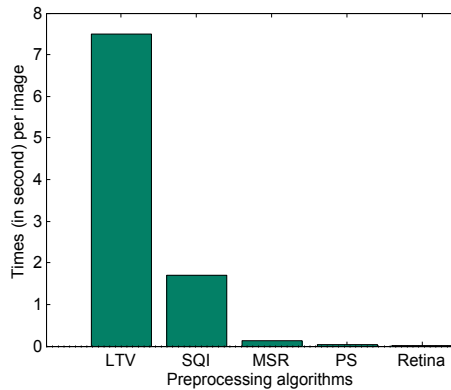
Fig. 16. Average computational time of different algorithms on a $192 \times 168$ image.

B database being preprocessed by different methods. Fig. 17(b) shows an example of correct detection and Table 5 compares the performance of different normalization methods. It is clear that our method improves significantly the face detection performance and also outperforms other preprocessing algorithms.



(a) Without preprocessing, face detector does not work

(b) With preprocessing, face detector works well

Fig. 17. Illustration of performance of the proposed algorithm for face detection.

| Method | Without preprocessing | HE | MSR | Proposed |
|---|---|---|---|---|
| Rate (%) | 12 | 98 | 99.0 | 99.5 |

Table 5. Face detection rate on the Yale B database with different preprocessing methods.

## 6. Conclusion

Face recognition has obvious advantages over other biometric techniques, since it is natural, socially well accepted, and non-intrusive. It has attracted substantial attention from various disciplines and contributed to a skyrocketing growth in the literature. Although these attempts, unconstrained face recognition remains active and unsolved. One of the remaining challenges is face recognition across illumination, which is addressed in this chapter. Inspired

by the natural ability of human retina that enables the eyes to see objects in varying illumination conditions, we propose a novel illumination normalization method simulating the performance of retina by combining two adaptive nonlinear functions, a Difference of Gaussian filter and a truncation. The proposed algorithm not only removes the illumination variations and noise, but also reinforces the image contours. Experiments are conducted on three databases (Extended Yale B, FERET and AR) using different face recognition techniques (PCA, LBP, Gabor filters). The very high recognition rates obtained in all tests prove the strength of our algorithm. Considering the computational complexity, ours is a real time algorithm and is faster than many competing methods. The proposed algorithm is also useful for face detection.

## 7. References

Adini, Y., Moses, Y. & Ullman, S. (1997). Face recognition: The problem of compensating for changes in illumination directions, *IEEE Trans. PAMI* 19: 721–732.

Ahonen, T., Hadid, A. & Pietikainen, M. (2004). Face recognition with local binary patterns, *European Conference on Computer Vision*, pp. 469–481.

Basri, R. & Jacobs, D. W. (2003). Lambertian reflectance and linear subspaces, *IEEE Trans. PAMI* 25(2): 218–233.
    URL: *http://dx.doi.org/10.1109/TPAMI.2003.1177153*

Beaudot, W. (1994). *The neural information processing in the vertebrate retina: A melting pot of ideas for artificial vision*, PhD thesis, Grenoble Institute of Technology, Grenoble, France.

Belhumeur, P., Hespanha, J. & Kriegman, D. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, *IEEE Trans. PAMI* .
    URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.3247*

Belhumeur, P. & Kriegman, D. (1998). What is the set of images of an object under all possible illumination conditions?, *Int. J. Comput. Vision* 28(3): 245–260.
    URL: *http://dx.doi.org/10.1023/A:1008005721484*

Benoit, A. (2007). *The human visual system as a complete solution for image processing*, PhD thesis, Grenoble Institute of Technology, Grenoble, France.

Chen, H., Belhumeur, P. & Jacobs, D. (2000). In search of illumination invariants, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*.

Chen, T., Yin, W., Zhou, X., Comaniciu, D. & Huang, T. (2006). Total variation models for variable lighting face recognition, *IEEE Trans. PAMI* 28(9): 1519–1524.
    URL: *http://dx.doi.org/10.1109/TPAMI.2006.195*

Garcia, C. & Delakis, M. (2004). Convolutional face finder: A neural architecture for fast and robust face detection, *IEEE Trans. PAMI* 26(11): 1408–1423.
    URL: *http://dx.doi.org/10.1109/TPAMI.2004.97*

Georghiades, A. & Belhumeur, P. (2001). From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. PAMI* 23: 643–660.

Jobson, D., Rahman, Z. & Woodell, G. (1997). A multiscale retinex for ridging the gap between color images and the human observation of scenes, *IEEE Trans. On Image Processing* 6: 965–976.

Land, E. & McCann, J. (1971). Lightness and retinex theory, *J. Opt. Soc. Am.* 61(1): 1–11.
    URL: *http://dx.doi.org/10.1364/JOSA.61.000001*

Lee, K., Ho, J. & Kriegman, D. J. (2005). Acquiring linear subspaces for face recognition under variable lighting, *IEEE Trans. PAMI* 27(5): 684–698.
    URL: *http://dx.doi.org/10.1109/TPAMI.2005.92*

Liu, C. & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Trans. Image Processing* 11: 467–476.

Martinez, A. M. & Benavente, R. (1998). The ar face database, *Technical report*.

Meylan, L., Alleysson, D. & Susstrunk, S. (2007). Model of retinal local adaptation for the tone mapping of color filter array images, *Journal of the Optical Society of America A* 24: 2807–2816.
   URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.109.2728*

Moses, Y., Adini, Y. & Ullman, S. (1994). Face recognition: The problem of compensating for changes in illumination direction, *European Conference on Computer Vision*.

Nefian, A. & III, M. H. (1998). Hidden markov models for face recognition, *ICASSP*, pp. 2721–2724.

Phillips, J. (1999). Support vector machines applied to face recognition, *NIPS*, Vol. 11, pp. 803–809.
   URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.2690*

Phillips, J., H.Moon & et al., S. R. (2000). The feret evaluation methodology for face-recognition algorithms, *IEEE Trans. PAMI* 22: 1090–1104.

Tan, X. & Triggs, B. (2007). Enhanced local texture feature sets for face recognition under difficult lighting conditions, *AMFG*, pp. 168–182.
   URL: *http://portal.acm.org/citation.cfm?id=1775256.1775272*

Turk, M. & Pentland, A. (1991). Eigenfaces for recognition, *J. Cognitive Neuroscience 3* pp. 71–86.

Vu, N.-S. & Caplier, A. (2009). Illumination-robust face recognition using the retina modelling, *International Conference on Image Processing*, IEEE.

Vu, N.-S. & Caplier, A. (2010a). Face recognition with patterns of oriented edge magnitudes, *European Conference on Computer Vision*.

Vu, N.-S. & Caplier, A. (2010b). Patch-based similarity hmms for face recognition with a single reference image, *International Conference on Pattern Recognition*.

Wang, H., Li, S. & Wang, Y. (2004). Generalized quotient image, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (2)*, pp. 498–505.

Wiskott, L., Fellous, J. M., Kuiger, N. & von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching, *IEEE Trans. PAMI* 19(7): 775–779.
   URL: *http://dx.doi.org/10.1109/34.598235*

# Temporal Synchronization and Normalization of Speech Videos for Face Recognition

Usman Saeed[1] and Jean-Luc Dugelay[2]
*[1]Department of Computer Science,*
*COMSATS Institute of Information Technology,*
*[2]Multimedia Communication Department,*
*Eurecom-Sophia Antipolis,*
*[1]Pakistan*
*[2]France*

## 1. Introduction

Automatic Face Recognition (AFR) is a domain that provides various advantages over other biometrics, such as acceptability and ease of use, but due to the current trends, the identification rates are still low as compared to more traditional biometrics, such as fingerprints. Image based face recognition, was the mainstay of AFR for several decades but quickly gave way to video based AFR with the arrival of inexpensive video cameras and enhanced processing power.

Video based face recognition has several advantages over image based techniques, the two main being, more data for pixel-based techniques, and availability of temporal information. But with these advantages there are some inconveniences also, the foremost being the augmentation of variation. In the classical image based face recognition degraded performance has mostly been attributed to three main sources of variation in the human face, these being pose, illumination and expression. Among these, pose has been quite problematic both in its effects on the recognition results and the difficulty to compensate for it. Techniques that have been studied for handling pose in face recognition can be classified in 3 categories, first are the ones that estimates an explicit 3D model of the face (Blanz & Vetter, 2003) and then use the parameters of the model for pose compensation, second are subspace based such as eigenspace (Matta & Dugelay, 2008) and the third type are those which build separate subspaces for each pose of the face such as view-based eigenspace (Lee & Kriegman, 2005).

Managing illumination variation in videos has been relatively less studied as compared to pose, mostly image based techniques are extended to video. The two classical image based techniques that have been extended for video with relative success are illumination cones (Georghiades et al., 1998) and 3D morphable models (Blanz & Vetter, 2003). Lastly expression invariant face recognition techniques can be divided in two categories, first are based on subspace methods that model the facial deformations (Tsai et al., 2007). Next are techniques that use morphing techniques (Ramachandran et al., 2005), who morph a smiling into a neutral face.

In this chapter we have focus on another mode of variation that has been conveniently neglected by the research community that is caused by speech. The deformation caused by lip motion during speech can be considered a major cause of low recognition results, especially in videos that have been recorded in studio conditions where illumination and pose variations are minimal. In this chapter we present a novel method of handling this variation by using temporal synchronization and normalization based on lip motion.

The chapter is divided into two main parts; in the first part we propose a temporal synchronization method that, given a group of videos for a person repeating the same phrase in all videos, studies the lip motion in one of the videos and selects synchronization frames based on a criterion of significance (optical flow). The next module then compares the motion of these synchronization frames with the rest of the videos and selects frames with similar motion as synchronization frames. For evaluation of our proposed method we use the classical eigenface algorithm to compare synchronization frames extracted from the videos and random frames to observe the improvement in face recognition results. The second part of this chapter consists of a temporal normalization algorithm that takes the synchronization frames from the previous module and normalizes the length of the video by lip morphing. Firstly the videos are divided into segments defined by the location of the synchronization frames. Next the normalization is carried out independently for each segment of the video by first selecting an optimal number of frames for each segment and then adding and removing frames to normalize the length of the video. For evaluation of our normalization algorithm we have devised a spatio-temporal person recognition algorithm using video information. By applying discrete video tomography, our algorithm summarizes the facial dynamics of a sequence into a single image, which is then analyzed by a modified version of the eigenface for improvement in a face recognition scenario.

The rest of the chapter is divided as follows. In Section 2 we elaborate the lip detection method. In Section 3 we give the synchronization method, after that we present the normalization method in Section 4 and in section 5 we give the concluding remarks and future works.

## 2. Lip detection

In this section we present a lip detection method to extract the outer lip contour that combines edge based and segmentation based algorithms. The results from the two methods are then combined by OR fusion. The novelty lies in the fusion of two methods, which have different characteristics and thus exhibit different type of strengths and weaknesses. The other significance of this study lies in the extensive testing and evaluation of the detection algorithm on a realistic database. Most previous studies either never carried out empirical comparisons to the ground truth or sufficed by using a limited dataset. Some studies (Liew et al., 2003; Guan, 2008) do exist that have presented results on considerably large datasets but these mostly consists of high resolution images with constant lighting conditions. Figure 1 gives an overview of the lip detection algorithm. Given a database image containing a human face the first step is to select the mouth Region of Interest (ROI) using the tracking points provided with the database. The next step involves the detection, where the same ROI is provided to the edge and segmentation based methods. Finally the results from the two methods are fused to obtain the final outer lip contour.
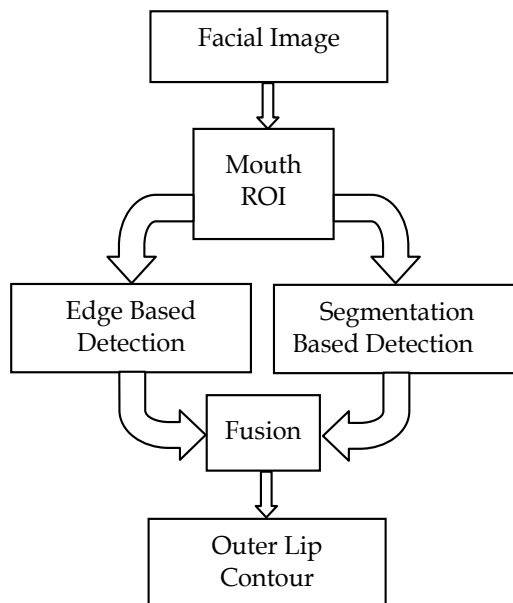
Fig. 1. Overview of lip detection.

## 2.1 Edge based detection

The first algorithm is based on a well accepted edge detection method, it consists of two steps, the first one is a lip enhancing color transform and the second one is edge detection based on active contours. Several color transforms have already been proposed for either enhancing the lip region independently or with respect to the skin. Here, after evaluating several transforms we have selected the color transform (equation 1) proposed by (Canzler & Dziurzyk, 2002). It is based on the principle that blue component has reduced role in lip / skin color discrimination.

$$I = \frac{2G - R - 0.5B}{4} \tag{1}$$

Where R,G,B are the Red, Green and Blue components of the mouth ROI. The next step is the extraction of the outer lip contour, for this we have used active contours (Michael et al., 1987). Active contours (cf. Figure 2) are an edge detection method based on the minimization of an energy associated to the contour. This energy is the sum of internal and external energies; the aim of the internal energy is to maintain the shape as regular and smooth as possible. The most straightforward approach grants high energy to elongated contours (elastic force) and to high curvature contours (rigid force). The external energy models the edge of the object and is supposed to be minimal when the active contours (snake) is at the object boundary. The simplest approach consists of using regularized gradient as the external energy. In our study the contour was initialized as an oval half the size of the ROI with node separation of four pixels.

Since we have applied active contours which have the possibility of detecting multiple objects, on a ROI which may include other features such as the nose tip, jaw line etc. an additional cleanup step needs to be carried out. This consists of selecting the largest detected

object approximately in the middle of the image as the lip and discarding the rest of the detected objects.
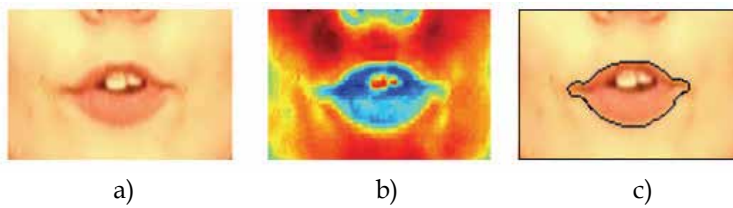


Fig. 2. a) Mouth ROI, b) Color Transform, c) Edge Detection.

### 2.2 Segmentation based detection

In contrast to the edge based technique the second approach is segmentation based after a color transform in the YIQ domain (cf. Figure 3) . As in the first approach we experimented with several color transform presented in the literature to find the one that is most appropriate for lip segmentation. (Thejaswi & Sengupta, 2008) have presented that skin/lip discrimination can be achieved successfully in the YIQ domain, which firstly de-couples the luminance and chrominance information. They have also suggested that the I channel is most discriminant for skin detection and the Q channel for lip enhancement. Thus we transformed the mouth ROI form RGB to YIQ color space using the equation 2 and retained the Q channel for further processing.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -0.522591 & 0.31135 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (2)$$

In classical active contours the external energy is modelled as an edge detector using the gradient of the image, to stop the evolution of the curve on the boundary of the desired object while maintaining smoothness in the curve. This is a major limitation of the active contours as they can only detect objects with reasonably defined edges. Thus for the second method we selected a technique called "active contours without edges" (Chan & Vese, 2001), which models the intensities in different region of the image and uses it as the stopping term in active contours. More precisely this model (Chan & Vese, 2001) is based on Mumford–Shah functional and level sets. In the level set formulation, the problem becomes a mean-curvature flow evolving the active contour, which will stop on the desired boundary. However, the stopping term does not depend on the gradient of the image, as in the classical active contour models, but is instead based on Mumford–Shah functional for segmentation.
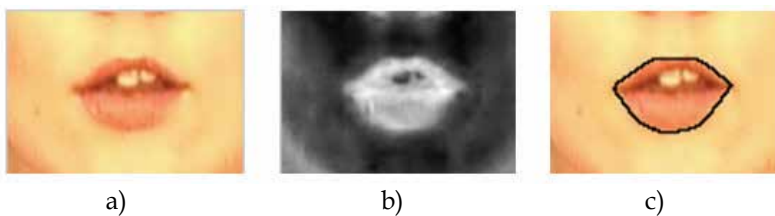


Fig. 3. a) Mouth ROI, b) Color Transform, c) Region Detection

### 2.3 Error detection and fusion

Lip detection being an intricate problem is prone to errors, especially the lower lip as reported by (Bourel et al., 2000). We faced two types of errors and propose appropriate error detection and correction techniques. The first type of error, which was commonly observed, was caused when the lip was missed altogether and some other feature was selected. This error can easily be detected by applying feature value and locality constraints such as the lip cannot be connected to the ROI's boundary and cannot have an area value less than one-third of the average area value in the entire video sequence. If this error was observed, the detection results were discarded.

The second type occurs when the lip is not detected in its entirety, e.g. missing the lower lip, such errors are difficult to detect thus we proposed to use fusion as a corrective measure, under the assumption that both the detection techniques will not fail simultaneously.

The detection results from the above described methods were then fused using OR logical operator. The outer lip contours are used to create binary masks which describe the interior and the exterior of the outer lip contour. These were then fused using OR Logical Operator defined as

| A | B | V |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

Table 1 presents the commonly observed errors and the effect of OR fusion on the results.

| | Type 1 Error | Type 2 Error | No Error |
|---|---|---|---|
| Segmentation Based |  |  |  |
| Edge Based |  |  |  |
| OR Fusion |  |  |  |

Table 1. Errors and OR Fusion

## 2.4 Experiments and results

In this section we elaborate the experimental setup and discuss the results obtained. Tests were carried out on Valid Database (Fox et al., 2005) which consists of five recording sessions of 106 subjects using the third utterance. One image was extracted from each of the five videos to create a database of 530 facial images. The reason for selecting one image per video was that the database did not contain any ground truth for lip detection, so ground truth had to be created manually, which is a time consuming task. The images contained both illumination and shape variation; illumination from the fact that they were extracted from all five videos, and shape as they were extracted from random frames of speaker videos.

As already described above the database did not contain any ground truth with respect to the outer lip contour. Thus the ground truth was established manually by a single operator using Adobe Photoshop. The outer lip contour was marked using the magnetic lasso tool which separated the interior and exterior of the outer lip contour by setting the exterior to zero and the interior to one.

To evaluate the lip detection algorithm we used the following two measures proposed by (Guan, 2008), the first measure, equation 3, determines the percentage of overlap (OL) between the segmented lip region A and the ground truth AG. It is defined in Equation 3.

$$OL = \frac{2(A \cap A_G)}{A + A_G} * 100 \tag{3}$$

Using this measure, total agreement will have an overlap of 100%. The second measure, equation 4, is the segmentation error (SE) defined as

$$SE = \frac{OLE + ILE}{2 * TL} * 100 \tag{4}$$

LE (outer lip error) is the number of non-lip pixels being classified as lip pixels and ILE (inner lip error) is the number of lip-pixels classified as non-lip ones. TL denotes the number of lip-pixels in the ground truth. Total agreement will have an SE of 0%.

Initially we calculated the overlap and segmentation errors for edge and segmentation based methods individually, and it was visually observed that edges based method was more accurate but not robust and on several occasions missed almost half of the lip. This can also be observed in the histogram (cf. Figure 4) of segmentation errors; although the majority of lips are detected with 10% or less error but a large number of lip images exhibit approximately 50% of segmentation error. On the other hand segmentation based method was less accurate as majority of lips detected are with 20% error but was quite robust and always succeeded in detecting the lip.

The minimum segmentation, Table 2, error obtained was around 15%, which might seem quite large, but on visual inspection of Figure 4, it is evident that missing the lip corners or including a bit of the skin region can lead to this level of error. Another aspect of the experiment that must be kept in mind is the ground truth. Although every effort was made to establish an ideal ground truth but due to limited time and resources some compromises had to be made. "OR Fusion on 1st Video" are the results that were obtained when OR fusion was applied to only the images from the first video, which are recorded in studio conditions.
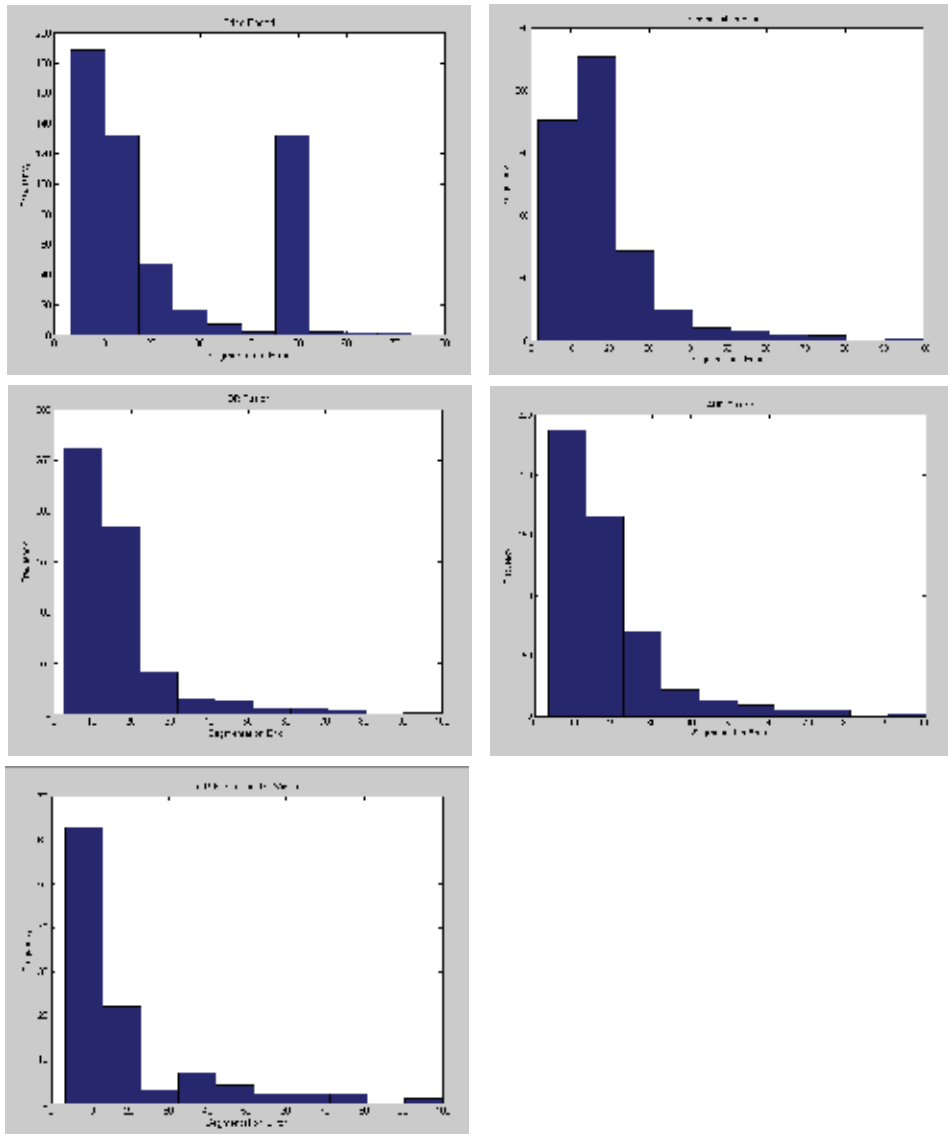
Fig. 4. Histograms for Segmentation Errors

| Lip Detection Method | Mean Segmentation Error (SE) % | Mean Overlap (OL) % |
|---|---|---|
| Segmentation Based | 17.8225 | 83.6419 |
| Edge Based | 22.3665 | 65.6430 |
| OR Fusion | 15.6524 | 83.9321 |
| AND Fusion | 18.4067 | 84.2452 |
| OR Fusion on 1st Video | 13.9964 | 87.1492 |

Table 2. Lip detection Results

Fig. 5. Example of Images with 15 % Segmentation Error

## 3. Synchronization

In this section we propose a temporal synchronization method that, given a group of videos for a person repeating the same phrase in all videos, studies the lip motion in one of the videos and selects synchronization frames based on a criterion of significance (optical flow). The next module then compares the motion of these synchronization frames with the rest of the videos and selects frames with similar motion as synchronization frames. For evaluation of our proposed method we use the classical eigenface algorithm to compare synchronization frames extracted from the videos and random frames to observe the improvement in a face recognition results.

The proposed synchronization method can be divided into two main parts; first is a selection method which selects frames in one of the video that are considered significant, second is a search algorithm in which the synchronization frames selected in the first video are synchronized with the remaining videos.

### 3.1 Synchronization frame selection

The aim of this module is to select synchronization frames from the first video of the group of videos for a specific person. Given a group of videos $V_i$ for the person $p$, where $i$ is the video index in the group, this module takes the first video $V_1$ for each person as input and selects synchronization frames $SF_1$ , that are considered useful for synchronization with the rest of the videos.  The criterion for significance is based on amount of lip motion, hence frames that exhibit more lip motion as compared to the frames around them are considered significant. First for the video $V_1$ the mouth region of interest (ROI) $MI_t$ for each frame $t$ is isolated based on tracking points provided with the database. Then frame by frame optical flow is calculated using the Lucas Kanake method (cf. Figure 6) for the entire video resulting in a matrix of horizontal and vertical motion vectors. As we are interested in a general description of the amount of lip motion in the frame we then calculate the mean of the motion vectors $Of_t$  (cf. Figure 8) for each mouth ROI $MI_t$.

$$for\ t \leftarrow 1\ to\ N-1$$
$$[u_{m,n,t}\ v_{m,n,t}] = LK(MI_t, MI_{t+1})$$
$$Of_t = \sum_{m=1}^{M} \sum_{n=1}^{N} (abs(u_{m,n,t}) + abs(v_{m,n,t}))$$
$$end$$

Fig. 6. Mean optical flow algorithim

Where $T$ is the number of frames in the video $V_i$ , $LK()$ calculates the Lucas Kanade optical flow. $u_{m,n,t}$  and $v_{m,n,t}$ are the horizontal and vertical components of the motion vectors at  row $m$ and column $n$ of the frame $t$.
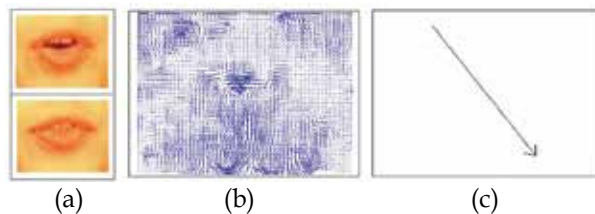
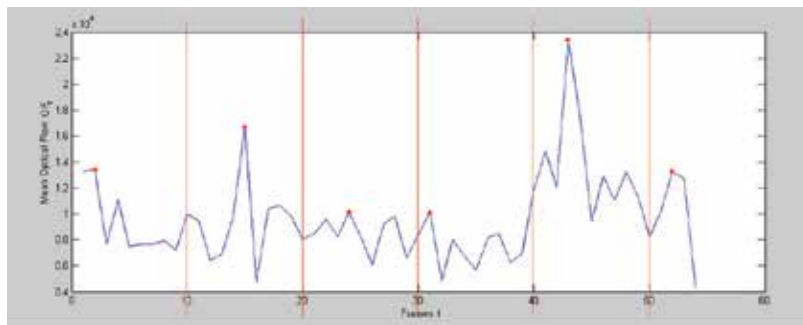Fig. 7. (a) Mouth ROI. (b) *LK* optical flow. (c) Mean vector.



Fig. 8. Mean optical flow $Of_t$ for video

The next step is to select synchronization frames $SF_1$ based on the mean optical flow $Of_t$, if we select frames that exhibit maximum lip motion there is a possibility that these frames might lie in close vicinity to each other. Thus we decided to divide the video into predefined segments (cf. Figure 8) and then select the frame with local maxima as synchronization frames.

$$for\ t \leftarrow 1\ to\ (N - D)\ with\ incriments\ of\ D$$
$$SF_1 = Frame\ with\ value\ (\max(Of_t\ to\ Of_{t+D}))$$
$$end$$
$$where\ D = \frac{N}{K}$$

Fig. 9. Syncronization frame selecion algorithim

Where $T$ is the total number of frames in the video. $K$ is the number of synchronization frames, its value is predefined and is based on the average temporal length of the videos in the database and will be given in the experiments and results section.

### 3.2 Synchronization frame matching
In the previous module we have selected synchronization frames from the first video of a person and in this module we try to match these frames with the remaining videos in the group. This module can be broken down into several sub-modules, the first one is a feature extractor where we extracted two features related to lip motion. The second is an alignment algorithm that aligns the extracted lip features before matching, and the last sub-module is a search algorithm that matches the lip features using an adapted mean-square error algorithm. This results in the synchronization frame matrix $SF_i$ for each person.

### 3.2.1 Feature extraction

In this section we have studied the utility of two mouth features, the first one is quite simply the mouth ROI ($MI_t$) as used in the previous module, the second is based on lip shape and appearance ($LSA_t$) and its is based on the outer lip contour extracted in Section 1. Once the outer lip contour is detected the background is then removed and the final feature is obtained as depicted in Figure 9. It contains the shape information in the form of lip contour and the appearance as pixel values inside the outer lip contour. Thus the feature image $J$ may consist of either $MI_t$ or $LSA_t$.
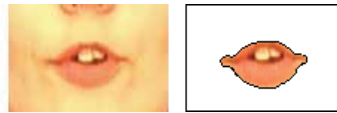


Fig. 10. Lip Feature Image

### 3.2.2 Alignment

Before the actual matching step, it is imperative that the feature images $J$ ($MI_t$, $LSA_t$) are properly aligned, the reason being that some feature images maybe naturally aligned and thus have unfair advantage in matching. The alignment process is based on minimization of mean square error between feature images.

### 3.2.3 Synchronization frame matching

The last module consists of a search algorithm, which tries to find frames having similar lip motion as synchronization frames selected from the first video in the rest of the videos. The algorithm is based on minimizing the mean square error, adapted for sequences of images.

Let $J_{f(k),i,w}$ be the feature image, where $k$ is the synchronization frame index, $f(k)$ is the location of the synchronization frame in the video, $i$ describes the video number and $w$ the search window, which is fixed to +/-5 frames. Thus the search algorithm tries to find synchronization frames $SF_i$ by matching the current feature image $J_{f(k),1,0}$ previous feature image $J_{f(k)-1,1,0}$ and the future feature image $J_{f(k)+1,1,0}$ from the first video with the rest of the videos within a search window $w$. The search window $w$ is created in the rest of the video centred at the location of the synchronization frame from the first video given by $f(k)$.

$$
\begin{aligned}
&for\ \ k \leftarrow 1\ to\ No\ of\ Synchronization\ Frames \\
&\quad for\ i \leftarrow 2\ to\ No\ of\ Videos\ Per\ Person \\
&\quad\quad for\ w \leftarrow f(k) - 5\ to\ f(k) + 5 \\
&SF_i = \underset{J}{\arg\min} \frac{\sum\sum((J_{f(k)-1,1,0})^2 - (J_{f(k)-1,i,w})^2) + \sum\sum((J_{f(k),1,0})^2 - (J_{f(k),i,w})^2) + \sum\sum((J_{f(k)+1,1,0})^2 - (J_{f(k)+1,i,w})^2)}{(M * N)}
\end{aligned}
$$

Fig. 11. Syncronization frame matching algorithim

Where $SFi$ is the final matrix that contains the synchronization frames for all the videos $Vi$ for one person.

### 3.3 Person recognition

Classification was carried out using the eigenface technique (Turk & Pentland, 1991). The pre-processing step consists of histogram equalisation and image vectorisation (image pixels are arranged in long vectors).

We apply a linear transformation from the high dimensional image space, to a lower dimensional space (called the face space). More precisely, each vectorised image $\mathbf{s}_n$ is approximated with its projection in the face space $\mathbf{v}_n \in \Re^D$ by the following linear transformation, equation 5.

$$v_n = W^T(s_n - \mu) \tag{5}$$

where $\mathbf{W}$ is a projection matrix with orthonormal columns, and $\mu \in \Re^D$ is the mean image vector of the whole training set, equation 6.

$$\mathbf{\mu} = \frac{1}{JN}\sum_{j=1}^{J}\sum_{n=1}^{N}\mathbf{s}_{j,n} \tag{6}$$

in which $J$ is the total number of sequences in the training set, and $\mathbf{s}_{j,n}$ is the $n$-th vectorised image belonging to video $\Phi_j$. The optimal projection matrix W is computed using the principal component analysis (PCA).

After the image data set is projected into the face space, the classification is carried out using a nearest neighbour classifier which compares unknown feature vectors with client models in feature space. The similarity measure adopted $S$, equation 7, is inversely proportional to the cosine distance.

$$S(y_i, y_j) = 1 - \frac{y_i^T y_j}{||y_i|| \; ||y_j||} \tag{7}$$

and has the property to be bounded into the interval [0, 1].

### 3.4 Experiments and results

Tests were carried out on Valid Database (Fox et al., 2005) which consists of five recording sessions of 106 subjects using the third utterance. The videos contain head and shoulder region of the subjects and the subjects are present in front of the camera from the beginning till the end.

The first video $V_1$ was selected for the synchronization frame selection module and the rest of the 4 videos were then matched with the first video using the synchronization frame matching module. To estimate the improvement due to our synchronization process we have compared the synchronization frames $SF_i$ and randomly selected frames using the person recognition module. The first video was excluded from training and testing due to its unrealistic recording conditions, 2nd and 3rd videos were used for training and 4th and 5th were used for testing both synchronization and random frames.

We apply PCA to the enrolment subset to compute a reduced face space of 243 dimensions. Then, the client models are registered into the system using their centroid vectors, which are calculated by taking the average of the feature vectors in the enrolment subset; in the end, recognition is achieved using a nearest neighbour classifier with cosine distances.

We have created 8 datasets from our database by varying the parameters such as selection method, the type of feature image and the number of synchronization frames. The results are summarized in Table 3, the first column gives dataset number, the second column the method for selecting frames, the first 4 datasets use the proposed synchronization frame selection method and the last 4 datasets were created by selecting random frames from the

videos. The third column signifies which lip features were used in the synchronization frame matching module. The fourth column is the number of synchronization frames $K$ that were used for each video, in this study we have limited $K$ to only 7 and 10 frames as most of the video in our database ranged from 60 to 110 frames. In case of last 4 datasets the number of synchronization frames simply signifies the number of random frames selected. The last column gives the identification rates.

| Dataset | Method | Lip Feature | Number of Synchronization Frames | Identification Rates |
|---------|--------|-------------|----------------------------------|----------------------|
| 1 | Synchronization | MI | 7 | 71.80 % |
| 2 | Synchronization | MI | 10 | 74.18 % |
| 3 | Synchronization | LSA | 7 | 72.28 % |
| 4 | Synchronization | LSA | 10 | 74.02 % |
| 5 | Random | - | 7 | 69.01 % |
| 6 | Random | - | 10 | 69.92 % |
| 7 | Random | - | 7 | 69.64 % |
| 8 | Random | - | 10 | 68.85 % |

Table 3. Person Recognition Results

The main result of this study is the overall improvement of identification results from synchronization frames as compared to random frames, which is evident from the Table 3. If we compare the identification results from the first 4 and last 4 datasets, it is obvious that there is an average improvement of around 4% between the 2 group of datasets. The second result that can be deduced is the improvement of recognition rates when more synchronization frames are used. The number of synchronization frames in the case of random frames simply signifies how many random frames were used and as it can be seen from the Table 3, using more random frames has no impact on the identification results. The third is insignificant change with regards to using *MI* or *LSA* as features. Here we would like to emphasize that the amount of testing for the second and third results is rather limited but this was not the main focus of this study.

## 4. Normalization

This section of the chapter consists of a temporal normalization algorithm that takes the synchronization frames from the previous module and normalizes the length of the video by lip morphing. Firstly the videos are divided into segments defined by the location of the synchronization frames. Next the normalization is carried out independently for each segment of the video by first selecting an optimal number of frames for each segment and then adding and removing frames to normalize the length of the video. The evaluation is carried out by comparing normalized videos with the original videos in a person recognition scenario.

### 4.1 Optimal number of frames

Given the video $V_i$ , it is first divided into segments $S_q$ , where $q$ is the number of segments and is equal to the number of synchronization frames plus one. Next the optimal number of frames $O_q$ for each corresponding segment $S_q$ is calculated by averaging the number of frames $F_{i,q}$ in the corresponding segment of the videos $V_i$.

$$for\ q \leftarrow 1\ to\ Q$$
$$for\ i \leftarrow 1\ to\ I$$
$$O_q = \frac{\sum_{i=1}^{I} F_{i,q}}{I}$$

Fig. 12. Optimal number algorithim

## 4.2 Transcoding

The next step is to add/remove frames (commonly known as transcoding) from each segment of the video so as to make them equal to the optimal number of frames. The simplest techniques for transcoding like up/down-sampling and interpolation results in jerky and blurred videos respectively. Advanced technique such as motion compensated frame rate conversion (Ugiyama et al., 2005), use block matching to estimate and compensate for motion but are imperfect as they lack information about the type of motion and thus frequently consider a uniform linear model of motion. As for this study we already have an estimation of lip motion from previous modules, we decided to use image morphing instead of block matching/compensation which results in visually superior results.

Morphing is the process of creating intermediate or missing frames from existing frames. Mesh morphing (Wolberg, 1996), one of the well studied techniques consists of creating a morphed frame $I_m$ from source frame $I_s$ and target frame $I_t$ by selecting corresponding feature points in $I_s$ and $I_t$, creating a mesh based on these feature points, warping $I_s$ and $I_t$ and finally interpolating warped frames to obtain the morphed frame $I_m$. In our study morphing was carried out only on the lip ROI as this region exhibits the most significant motion in the video. Lip ROI was first isolated and outer lip contour detected as in the previous section. These Lip ROI formed the $I_s$ and $I_t$ frames, feature points consisted of the 4 extremas of the outer lip contour (top, bottom, left, right). Mesh morphing was then carried out. Finally the morphed Lip ROI was superimposed on the original image to obtain the morphed frame (cf. Figure 13).
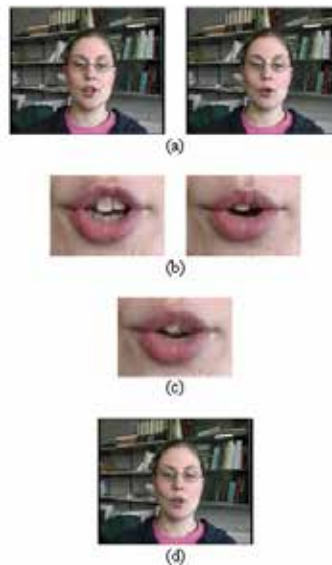


Fig. 13. (a) Existing Frames  (b) Lip ROI  (c) Morphed Lip ROI  (d) Morphed Frame

Decision regarding the number of frames to be added/removed is taken by comparing the number of frames in each segment $S_q$ to the optimal number of frames; the frames are then added/removed at regularly spaced intervals of the segment. Addition of a frame consists of creating a morphed frame $I_i$ from previously existing frames, $I_{i-1}$ and $I_{i+1}$. Similarly frame $I_i$ is removed by morphing frames $I_{i-1}$ and $I_i$ and replacing $I_{i-1}$ with the morphed frame, and replacing frame $I_{i+1}$ with the morphed frame from $I_i$ and $I_{i+1}$. Finally deleting the frame $I_i$. Thus

$$Frame\ Addition$$
$$I_i \leftarrow Morph(I_{i-1}, I_{i+1})$$

$$Frame\ Removal$$
$$I_{i-1} \leftarrow Morph(I_{i-1}, I_i)$$
$$I_{i+1} \leftarrow Morph(I_{i+1}, I_i)$$
$$Delete(I_i)$$

Fig. 14. Frame addition/deletion algorithim

### 4.3 Person recognition

For testing our normalization algorithm we used a spatio-temporal method proposed by (Matta & Dugelay, 2008). It consists of two modules: Feature Extraction, which transforms input videos into "X-ray images" and extracts low dimensional feature vectors, and Person Recognition, which generates user models for the client database (enrolment phase) and matches unknown feature vectors with stored models (recognition phase).

### 4.3.1 Feature extraction

Inspired by the application of discrete video tomography (Akutsu & Tonomura, 1994) for camera motion estimation, we compute the temporal X-ray transformation of a video sequence, to summarize the facial motion information of a person into a single X-ray image. It is important to notice that we restrict our framework to a fixed camera; hence, the video X-ray images represent the motion of the facial features and some appearance information, which is the information that we use to discriminate identities.

Given an input video of length $T_i$, $V_i \equiv \{I_i,1, \dots, I_i,T_i\}$, the Feature Extractor module first calculates the edge image sequence $E_i$, obtained by applying the Canny edge-finding method (Canny, 1986) frame by frame, equation 8.

$$E \equiv \{J_{i,1}, \dots, J_{i,T},i\} = f_{EF}(V_i) \tag{8}$$

Then, the resulting binary frames, $J_{i,t}$, are temporally added up to generate the X-ray image of the sequence, equation 9.

$$X_i = C \sum_{t=1}^{T_i} J_{i,t} \tag{9}$$

where $C$ is a scaling factor to adjust the upper range value of the X-ray image.

Fig. 15. Original Frames and Temporal X-ray Image.

After that, the Feature Extractor reduces the X-ray image space to a low dimensional feature space, by applying the principal component analysis (PCA) (also called the Karhunen-Loeve transform (KLT)): PCA computes a set of orthonormal vectors, which optimally represent the distribution of the training data in the root mean squares sense. In the end, the optimal projection matrix, $P$, is obtained by retaining the eigenvectors corresponding to the $M$ largest eigenvalues, and the X-ray image is approximated by its feature vector, $y_i \in \Re^M$ calculated using the linear projectionin equation 10.

$$y_i = P^T (x_i - \mu) \tag{10}$$

where $x_i$ is the X-ray image in a vectorial form and $\mu$ is the mean value.

### 4.3.2 Person recognition

During the enrolment phase, the Person Recognition module generates the client models and stores them into the system. These representative models of the users are the cluster centres in feature space that are obtained using the enrolment data set.

For the recognition phase, the system implements a nearest neighbour classifier which compares unknown feature vectors with client models in feature space. The similarity measure adopted $S$, equation 11, is inversely proportional to the cosine distance.

$$S(y_i, y_j) = 1 - \frac{y_i^T y_j}{||y_i|| \, ||y_j||} \tag{11}$$

and has the property to be bounded into the interval [0, 1].

### 4.4 Experiments and results

Tests were carried out on Valid Database (Fox et al., 2005) which consists of five recording sessions of 106 subjects using the third utterance. The first video was selected for the synchronization frame selection module and the rest of the 4 videos were then synchronized with the first video using the synchronization frame matching module. Finally all videos were temporally normalized.

To estimate the improvement due to our normalization process we have compared the normalized videos generated by our algorithm to original non-normalized videos using the person recognition module described above. First 3 videos were used for training and the rest 2 were used for testing. The number of synchronization frames in this study have been set to 7, as the average number of frames per video in our database was approximately 70. The recognition system has been tested using a feature space of size 190, constructed with the enrolment data set. The video frames are also pre-processed using histogram equalization, in order to reduce the illumination variations between different sequences.

| Method | CIR % (1st ) | CIR % (5th ) | CIR % (10th ) | EER % |
|--------|--------------|--------------|---------------|-------|
| Normalized Video | 69.02 % | 82.60 % | 89.13 % | 10.1 % |
| Original Video | 65.21 % | 81.52 % | 85.86 % | 11.9 % |

Table 4. Person Recognition Results

The identification and verification results are summarized in Table 4; its columns report the correct identification rates (CIR), computed using the best, 5-best and 10-best matches, and the equal error rates (EER) for the verification mode. We notice that the recognition system using normalized videos performs better than the analogous one working with non-normalized videos. Detailed Identification and EER Rates are given in figure 16.
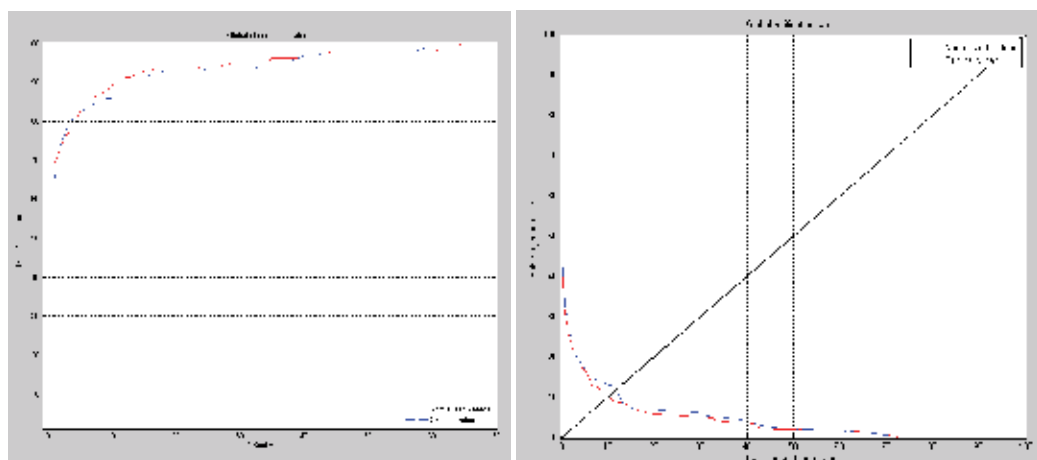


Fig. 16. Correct Identification Rates (CIR) and Verification Rates (EER)

## 5. Conclusions

In this chapter at first, we have presented a novel lip detection method based on the fusion of edge based and segmentation based methods, along with empirical results on a dataset of considerable size with illumination and speech variation. We observed that the edge based technique is comparatively more accurate, but is not so robust and fails if lighting conditions are not favourable, thus it ends up selecting some other facial feature. On the other hand the segmentation based method is robust to lighting but is not as accurate as the edge based method. Thus by fusing the results from the two techniques we achieve comparatively better results which can be achieved by using only one method. The proposed methods

were tested on a real world database of considerable size and illumination/speech variation with adequate results.

Then we have presented a temporal synchronization algorithm based on mouth motion for compensating variation caused by visual speech. From a group of videos we studied the lip motion in one of the videos and selected synchronization frames based on a criterion of significance. Next we compared the motion of these synchronization frames with the rest of the videos and selects frames with similar motion as synchronization frames. For evaluation of our proposed method we use the classical eigenface algorithm to compare synchronization frames and random frames extracted from the videos and observed an improvement of 4%.

Lastly we have presented a temporal normalization algorithm based on mouth motion for compensating variation caused by visual speech. Using the synchronization frames from the previous module we normalized the length of the video. Firstly the videos were divided into segments defined by the location of the synchronization frames. Next normalization was carried out independently for each segment of the video by first selecting an optimal number of frames and then adding/removing frames to normalize the length of the video. The evaluation was carried out by using a spatio-temporal person recognition algorithm to compare our normalized videos with non-normalized original videos, an improvement of around 4% was observed.

## 6. References

Blanz, V. and Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *PAMI*, Vol. 9, (2003), pp. 1063-1074

Matta, F. Dugelay, J-L. (2008). Tomofaces: eigenfaces extended to videos of speakers, *In Proc. of International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, USA, March 2008

Lee, K. and Kriegman, D. (2005). Online learning of probabilistic appearance manifolds for video-based recognition and tracking, *In Proc of CVPR*, San Diago, USA, June 2005

Georghiades, A. S. Kriegman, D. J. and Belhumeur, P. N. (1998). Illumination cones for recognition under variable lighting: Faces, *In Proc of CVPR*, Santa Barbara, USA, June 1998

Tsai, P. Jan, T. Hintz, T. (2007). Kernel-based Subspace Analysis for Face Recognition, *In Proc of International Joint Conference on Neural Networks*, Orlando, USA, August 2007

Ramachandran, M. Zhou, S.K. Jhalani, D. Chellappa, R. (2005). A method for converting a smiling face to a neutral face with applications to face recognition, *In Proc of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, USA, March 2005.

Liew, A.W.-C. Shu Hung, L. Wing Hong, L. (2003). Segmentation of color lip images by spatial fuzzy clustering, *IEEE Transactions on Fuzzy Systems*, Vol.11, No.4, (2003), pp. 542-549

Guan, Y.-P. (2008). Automatic extraction of lips based on multi-scale wavelet edge detection, *IET Computer Vision*, Vol.2, No.1, March 2008, pp.23-33

Canzler, U. and Dziurzyk, T. (2002). Extraction of Non Manual Features for Videobased Sign Language Recognition, *In Proceedings of the IAPR Workshop on Machine Vision Application*, Nara, Japan, June 2002

Michael, K. Andrew, W. and Demetri, T. (1987). Snakes: active Contour models, *International Journal of Computer Vision*, Demetri, Vol. 1, (1987), pp. 259-268

Thejaswi N. S and Sengupta, S. (2008). Lip Localization and Viseme Recognition from Video Sequences, *In Proc of Fourteenth National Conference on Communications*, Bombay, India, 2008

Chan, T.F. Vese, L.A. (2001). Active contours without edges, *IEEE Transactions on Image Processing* , Vol.10, No.2, (2001) pp.266-277

Bourel, F. Chibelushi, C. C. and Low, A. (2000). Robust Facial Feature Tracking, *In Proceedings of the 11th British Machine Vision Conference,* Bristol, UK, September 2000

Fox, N. O'Mullane, A. B. and Reilly, R.B. (2005). The realistic multi-modal VALID database and visual speaker identification comparison experiments, *In Proc of    5th International Conference on Audio- and Video-Based Biometric Person Authentication*, New York, USA,  July 2005

Turk, M.  and Pentland, A. (1991). Eigenfaces for recognition*, J. Cog. Neurosci.* Vol. 3, (1991) pp. 71–86

Ugiyama, K.  Aoki, T. Hangai, S. (2005). Motion compensated frame rate conversion using normalized motion estimation, *In Proc. IEEE Workshop on Signal Processing Systems Design and Implementation*, Athens, Greece, November 2005.

Wolberg, G. (1996). Recent Advances in Image Morphing, In Proceedings of the International Conference on Computer Graphics, USA, 1996

Huang,  C.L. and Huang, Y.M. (1997). Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification, *Journal of Visual Communication and Image Representation*, Vol. 8, (1997), pp. 278-290

Akutsu,  A. and Tonomura, Y. (1994). Video tomography: an efficient method for camerawork extraction and motion analysis, *In Proceedings of the Second ACM international Conference on Multimedia*, USA, 1994

Canny, J.  (1986). A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell*. Vol. 8, (1986), pp. 679-698

# Part 3

## Iris Recognition

# Personal Identity Recognition Approach Based on Iris Pattern

Qichuan Tian[1], Hua Qu[2], Lanfang Zhang[3] and Ruishan Zong[1]

*[1]College of electronic and information engineering,*
*Taiyuan University of Science and Technology,*
*[2]College of Science, Tianjin Polytechnic University,*
*[3]Taiyuan University of Science and Technology*
*China*

## 1. Introduction

Personal identification based on biometrics technology is a trend in the future. Traditional approaches, for example, keys, ID cards, username and password, are neither satisfactory nor reliable enough in many security fields, biometrics authorizations based on face, iris, fingerprint have become a hot research filed. In those methods, iris recognition is regarded as a high accuracy verification technology, so that many countries have the same idea of adopting iris recognition to improve the safety of their key departments.

The human iris can also be considered a valid biometrics for personal identification [Richard P W, 1996]. Biometrics recognition based on iris patterns is a hotspot as face recognition and fingerprint recognition recently years. The iris is the colored ring on the human eye between the pupil and the white sclera. Lots of physical biometric can be found in the colored ring of tissue that surrounds the pupil, such as corona, crypts, filaments, flecks, pits, radial furrows and striations. The iris features can be encoded by mathematical representation so that the patterns can be compared easily.

In real-time iris recognition application system, iris localization is a very important step for iris recognition. The iris regions segmentation accuracy and localization real-time performance will affect the whole recognition system's correct rate and effectiveness for large-scale database.

Because iris region is a small object and has low grey value, it is very difficult to capture high contrast iris image clearly. In order to improve iris image contrast, usually some illuminations such as near infrared light source are used to increase intensity; however these illuminations may result in some faculas in iris image and affect iris segmentation and iris features.

Here, we will discuss iris recognition system's algorithm, all steps of iris recognition system will be introduced in details. Finally, we will show the experimental results based on iris database.

## 2. Iris recognition system principle

Iris feature is convenience for a person to prove his/her identity based on him/her biometrics at any place and at any time. Iris recognition may become the most important

identify and verify approach in many departments such as navigation, finance, and so on.

Iris recognition system main includes iris capturing, image pre-processing, iris region segmentation, iris region normalization, iris feature extraction and pattern matching. Every part is very important for correct recognition person identity.

There are plenty of features in iris regions of human eye image. Because iris is a small and black object, iris image capturing is not an easy work. Iris must be captured at a short distance about 4cm-13cm and under a good illumination environment. Near infrared is a better light resource for many visible image recognition systems, such as face recognition. Near infrared can perform good illumination for enhancing image contrast and it is harmless to human eyes. In order to capture ideal iris image, it is necessary that a friend cooperation of user and captured camera is the base of iris recognition. A good cooperation can decrease the quantity of iris pre-processing and improve iris recognition real-time character. However, the demand for cooperation may affect user's feeling and result in users doesn't accept iris recognition system because of high rejection rate. So, many researchers begin to study imperfect iris recognition theory under in-cooperation conditions, iris recognition system will have more width application fields based on imperfect iris recognition theory under in-cooperation conditions.

Because of motion blur or defocus blur or the occluder like eyelids and eyelashes, objective evaluation algorithm of image quality can be used to select a high quality eye-image for iris recognition. Now, some literatures evaluate images by using features of frequency domain and spatial domain and by calculating the rate of effective iris regions' pixels to whole iris regions' pixels. Image quality evaluation is a step to select an eye-image for iris recognition, and this procedure can decrease processing work according to lower quality eye-images.
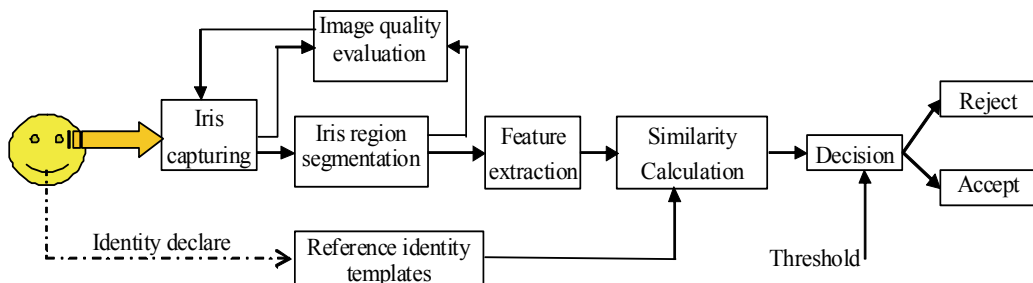


Fig. 1. Iris recognition system principle

## 3. Iris region segmentation

### 3.1 Iris segmentation background

In iris recognition system, iris region is the part between pupil and sclerotic, the aim of iris boundary localization is to locate the boundary of iris/pupil and the boundary of iris/sclerotic. Both inter boundary and outer boundary of iris are alike circles, so many iris localization methods are to locate iris boundaries using circle detector. John Daugman's Integral-differential method and Wildes's Hough transform method are effective methods for iris localization precision [John Daugman, 1993; Richard P W, 1997], but those methods cost lots of compute times because of large parameters search space. In order to improve iris localization real-time performance, many modified algorithms adopting some known

information of iris image are introduced in other literature. Pupil position can be estimated easily because of the lower grey level in pupil region and then iris boundary localization speed can be improved based on pupil position localization. Fast iris localization based on Hough transforms and inter-gradational method can be realized at the base of pupil position estimation [Tian Qi-chuan, 2006].

Usually, there is much interference in iris regions, and the interference can cause iris texture and grey value change, so high accuracy iris segmentation also need to remove the interference.

### 3.2 Iris boundary localization based on Hough transforms

In iris boundary localization methods, John Daugman's Integral-differential method and Wildes's Hough transform method with high iris localization precision are the most popular and effective methods, but the real-time character of those methods can't be satisfied. At the same time, these methods also have its disadvantages. Integral-differential method will be affected by local gradient maximum easily and then iris boundaries are located in these wrong positions, main reason is that light-spots will produce great gradient change. Hough transform for circle parameters voting can decrease local gradient effect, but the threshold for extracting edge points will affect the number of edge points and finally these edge points will cause iris boundary localization failure.

Fast boundaries localization based on prior pupil centre position estimation can improve iris boundary localization real-time, the main idea of this algorithm is: firstly, pupil centre coarse localization, secondly, edge detection based on canny operation; thirdly, iris inter boundary localization in a small image block selected; fourthly, edge extraction based on local grey gradient extreme value; finally, outer boundary localization in image block selected based Hough transform.

### 3.2.1 Pupil center coarse localization

In eye image, there are obvious lower grey levels in pupil regions than other parts as shown in Fig.2. Firstly, a binary threshold can be selected based on Histogram adopting p-tail method. Usually, iris image is captured in a distance, so pupil size is limited to a range in eye image, we can select threshold depend on the set rate of pupil pixels number to whole image pixels in histogram. Faculas can be seen in eye image as shown in Fig.2, these interferences must be removed, or they will affect iris boundary localization.
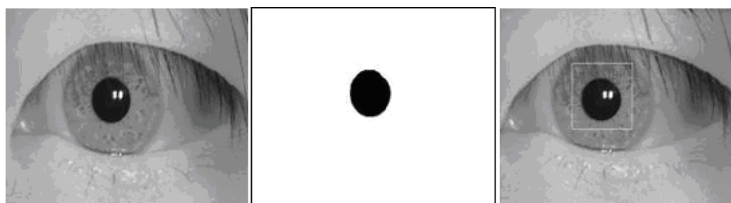


a) Original eye image     b) Binarized result     c) Pupil coarse estimation

Fig. 2. Pupil centre coarse localization

According to high grey pixels caused by faculas in pupil parts, morphological operation can be used for filling with these holes and remove noise such as eyelashes by using close-operation. So the original image of Fig.2 (a) can be transformed into the image of Fig.2 (b).

After removing light-spots, pupil centre position will be estimated more accuracy, it is helpful to precision iris localization.
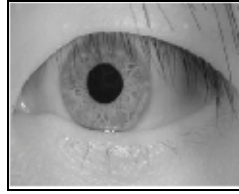


Fig. 3. Results of morphological operation

We can find a position using a moving window, when the window move to a position, the number of pixels in the window can be calculated, thus the position responding to the minimum sum can be regarded as pupil centre. The ordinate parameters (x, y) of pupil centre can be calculated as formula (1).

$$\begin{cases} x = \dfrac{1}{N}\sum_{i=1}^{N} x_i \\ y = \dfrac{1}{N}\sum_{i=1}^{N} y_i \end{cases} \tag{1}$$

Where, $(x_i, y_i)$ indicate 0-pixel ordinates, N is the number of 0-pixel points.
Pupil centre coarse estimation result is shown in Fig.2 (c).

### 3.2.2 Iris inter boundary delicate localization

At the base of pupil centre coarse localization, small region as n*n can be selected for pupil boundary localization, in this small region, we can achieve several local gradient extreme value points as binary edge points, then these points are divided into two sets (named left-set and right-set) according to their position direction to the pupil centre estimation, and we can take every point of left-set and every point of right-set as a pair points, so we can achieve n*n pairs data, the centre and radius of pupil boundary can be confirmed by every pair data voting for $h(x,y,r)$ as follows formula (2). Fig.4 is the principle of pupil boundary localization.
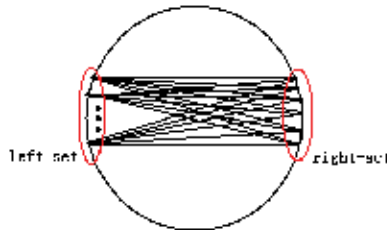


Fig. 4. Result of pupil boundary localization

$$h\left( \frac{x_i + x_j}{2}, \frac{y_i + y_j}{2}, \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right) + + \tag{2}$$

When finished all voting, we can achieve the pupil boundary parameters by using formula (3). Fig.5 is the pupil localization result.

$$H(x_p, y_p, r_p) = \max\left\{\bigcup h(x, y, r)\right\} \qquad (3)$$
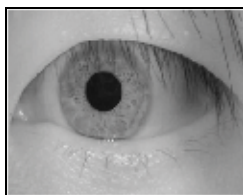


Fig. 5. Results of morphological operation

### 3.2.3 Iris outer boundary localization

In this paper, in order to realize iris boundary fast localization, a new method is proposed to extract iris edge points by using local gradient extreme value for each line of image, and then to achieve iris boundaries' parameters based on new voting approach. The upper eyelid and the lower eyelid often corrupt iris outer boundary, so we locate iris outer boundary by using part edge information.

Because iris outer boundary has lower gradient than iris inter boundary (pupil boundary), it is very difficult to extract edges by comparing with a set threshold. If the threshold for edge extraction in iris gradient image is lower, then the more edge points extracted will be not helpful to improve real-time character of iris recognition system and the more edge points also may cause iris localization failure; if the threshold for edge extraction in iris gradient image is higher, then many edge points may be lost and it also can cause iris localization failure. So, we want to extract edge information by using local grey gradient extreme value, then to locate outer boundary based on Hough transform.

Here, we don't select the threshold to extract edge points, instead of this, while we achieve binary edge points by comparing every line's grey gradient extreme value. The principle is shown in Fig.6. Same as pupil region selection, an image block for iris outer boundary localization can be selected so that we can locate iris boundary in a small region. Because upper eyelid usually obstruct iris region, we can select a small region for iris outer boundary localization as shown in Fig.6. Thus, based on Hough transform, we can adopt fewer edge points to achieve accuracy iris boundary. Fig.7 is the results of iris boundary localization.
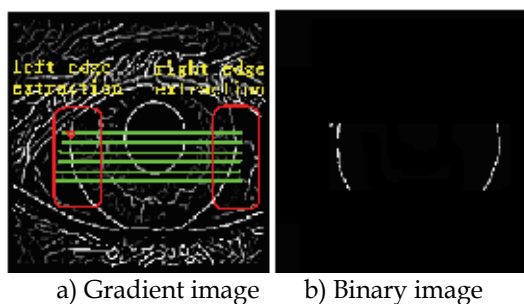


a) Gradient image     b) Binary image

Fig. 6. Binary edges extract principle based on line grey gradient extreme value
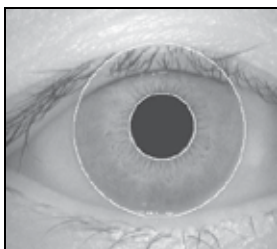
Fig. 7. Iris boundary localization results

Iris boundary localization algorithm is as follows:
**Step 1.** Pupil centre coarse localization;
**Step 2.** Select a small image block and extract edge information based on canny operator;
**Step 3.** Pupil boundary localization based on Hough transform;
**Step 4.** Select a small image block and extract edge information based on line's grey gradient extreme value;
**Step 5.** Iris outer boundary localization based on Hough transforms.

Due to improve localization speed and localization accuracy, taking the advantage of the grey information, we decrease the number of edge points and parameter range down to a small range to locate iris boundary.

### 3.3 Interference detection

There are many images looks like which shown in Fig.8. In iris images captured, interference, such as eyelids, eyelash and facula, will affect iris effective information for iris recognition [Wai-Kin Kong & David Zhang, 2003], we must remove these interferences between pupil boundary and iris outer boundary. From image #1 to image #4, we can see that eyelids cover iris's upper part and lower part usually. So, to detect boundary of eyelid is an important step in accuracy segmentation iris region.



a) Image #1        b) Image #2        c) Image #3        d) Image #4
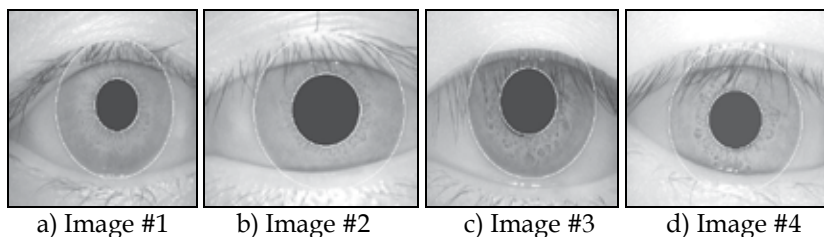
Fig. 8. Eye images

In order to achieve high accuracy segmentation and extract effective iris information in iris region, interference of eyelids, eyelash and faculas should be detected at a high accuracy rate [W. Kong & D.Zhang, 2001; Tian Qi-chuan, 2006]. Interferences detection is the other important aspect in segmenting iris region. Some image processing technologies are used to improve robustness of the algorithm to remove the interference of eyelash, light spots and image contrast. For every eyelid, using three-line detection can approach the eyelid. In order to get effective iris features, self-adaptive algorithms of detecting eyelid and eyelash and faculas are introduced in iris segmentation.

In some literature, four kinds of methods are introduced to remove eyelids as shown in Fig.9. The first method remove eyelids by using arc Hough transform to fit eyelid's

boundary, this method has a performance with high accuracy and low speed. The second method locate eyelids by line eyelids boundary localization to remove the affection of eyelids in iris recognition, this method has good real-time performance. The third method thinks iris regions near outer boundary has fewer distinguishable features and disturbed easily by eyelids, so a ring-band region can be removed and a ring-band near pupil's boundary can perform enough distinguishable information for iris high accuracy recognition. The fourth method is to locate eyelids based on multi-line's detection, and this method has lower compute cost and high eyelids localization accuracy.

a) Arc detection method    b) Line detection method

c) Circular method    d) Straight line to fit eyelid boundary
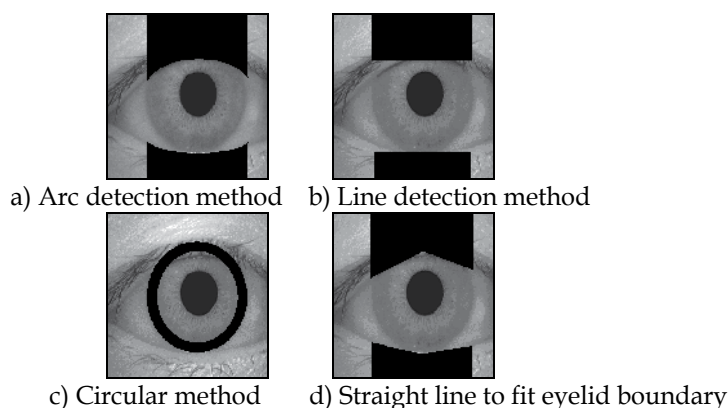
Fig. 9. Several principles for eyelids removing

Enough effective information will be helpful to iris recognition, so we select the fourth method mentioned above to locate iris eyelids. Firstly, horizontal edge can be extracted from eye image as shown in Fig.10, then radon transform is used in line localization to indicate eyelids outlines responding to four parts divided based on pupil centre position as shown in Fig.11. From the localization results we can see that the retained iris region has very high signal noise rate.

Fig. 10. Binary edge image for eyelids detection

a) iris boundaries    b) Straight line detection    c) Iris region

Fig. 11. Eyelids detection principle

Eyelid detection algorithm as follows:

**Step 1.**  select small image block based on iris boundary's parameters;

**Step 2.**  extract edge based on gradient operator and divide the selected image block into four parts;

**Step 3.**  locate four lines as eyelid's boundaries based on radon transform.

## 4. Iris region normalization

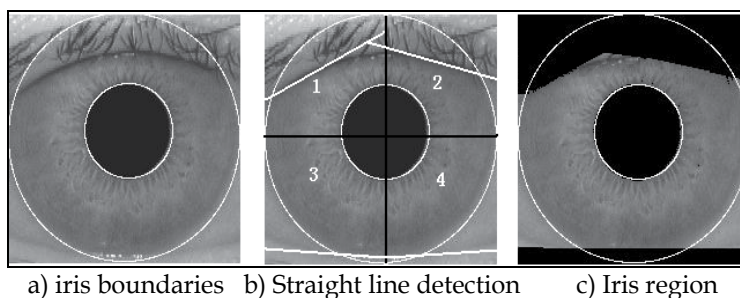The dimensional inconsistencies between eye images are mainly due to the stretching of the iris caused by pupil dilation from varying levels of illumination. Other cases of inconsistency include, varying imaging distance, rotation of the camera, head tilt, and rotation of the eye within the eye socket. The normalisation process will produce iris regions, which have the same constant dimensions, so that two photographs of the same iris under different conditions will have characteristic features at the same spatial location. Here we normalize iris circular region into rectangular region, the procedure is named as iris region normalization.

As we all known, another point of note is that the pupil region is not always concentric within the iris region, and is usually slightly nasal. Even we have achieved the parameters of iris boundaries, iris normalization that how to transform different resolution image into the same resolution rectangular region still is a problem. Usually, we can realize iris normalization by sampling M along with angle direction and sampling N along with radial direction. Fig.12 is an iris plastic model used to normalize iris region, we can indicate whole iris region by using the grey information of these pixels determined by coordinates combines of inter boundary and outer boundary [Libor Masek, 2001].



a) Iris boundary circles                 b) Rectangle region of iris ring
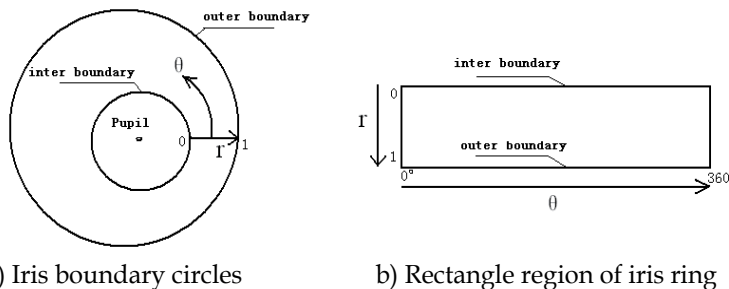
Fig. 12. Rubber sheet model

The homogenous rubber sheet model remaps each point within the iris region to a pair of polar coordinates $(r, \theta)$ where $r$ is on the interval [0, 1] and $\theta$ is angle $(0°, 360°)$. $I(x, y)$ is Cartesian coordinate of iris images, and $I(r, \theta)$ is corresponding polar coordinate. $(x_p, y_p)$ is the unit of inner boundary in Cartesian coordinate, $(x_l, y_l)$ is that of outer boundary, then coordinate transform is defined as follow:

$$I(x(r,\theta), y(r,\theta)) \rightarrow I(r,\theta) \tag{4}$$

$$\begin{cases} x(r,\theta) = (1-r)x_p(\theta) + rx_l(\theta) \\ y(r,\theta) = (1-r)y_p(\theta) + ry_l(\theta) \end{cases} \tag{5}$$

In above equation, $r = \dfrac{i}{M+1}, i = 1,2,...,M$ , $\theta = \dfrac{j}{N} \cdot 360°, j = 1,2,...,N$ ，$M$ is sample rate along with angle direction and $N$ is sample rate along with radial direction [C.H.Daouk, 2002]. If we choose small $M$ and $N$ , then low iris template size will be achieved [Raul Sanchez-Reillo,2010]. After boundary location and coordinate standardization, iris is showed as rectangular region $(i,j)$ .

If the sample number along circle is set to $N$ , the sample number along radius is set to $M$ , and then iris region can be transformed into a rectangle of $M \times N$ pixels. The details of normalization algorithm as follows:

**Step 1.** Achieve the parameters of $(x_p, y_p, r_p)$ and $(x_o, y_o, R_l)$ based on iris boundary localization on iris image $I(x,y)$ ;

**Step 2.** Calculate the distance between pupil centre and iris centre, and achieve connection direction angle;

$$
\begin{cases}
\phi = \arctan \dfrac{y_p - y_o}{x_p - x_o} \\
\Delta r = \sqrt{(x_p - x_o)^2 + (y_p - y_o)^2}
\end{cases}
\tag{6}
$$

**Step 3.** Select the centre of pupil as pole, then every point of iris inter boundary has the same formula $r(\theta) = r_p$ in polar coordinates, the sample point's position along iris outer boundary will be achieved as follows:

$$
\begin{cases}
\theta = j \times \pi / 180 \\
R(\theta) = \Delta r \cos(\pi - \theta - \phi) \\
\quad + \sqrt{R_l^2 - \Delta r^2 + (\Delta r \cos(\pi - \theta - \phi))^2}
\end{cases}
\tag{7}
$$

where $j = 1,2,\cdots,N$ .

**Step 4.** Every pixel's grey information of normalization iris region can be achieved using those grey of $(x,y)$ positions confirmed as follows:

$$
\begin{cases}
Rp = (1 - \dfrac{i}{M+1}) \times r(\theta) + \dfrac{i}{M+1} \times R(\theta) \\
x = Xp + Rp \cos(\theta) \\
y = Yp - Rp \sin(\theta) \\
Normalrize\_Iris(i,j) = I(x,y)
\end{cases}
\tag{8}
$$

where, $i = 1,2,\cdots,M$ , $j = 1,2,\cdots,N$ ，$\theta = j \times \pi / 180$ .

After removing the first line and the last line, we will achieve iris normalization region of $M \times N$ . Fig.13 (a) is an iris image, (b) is sample image, and (c) is the normalization result. From Fig.13, we can see that interference caused by eyelash and eyelid change iris texture, so we must label these interference so that we can eliminate these interference in pattern match. Fig.14 shows the label results according interference in iris normalization region.
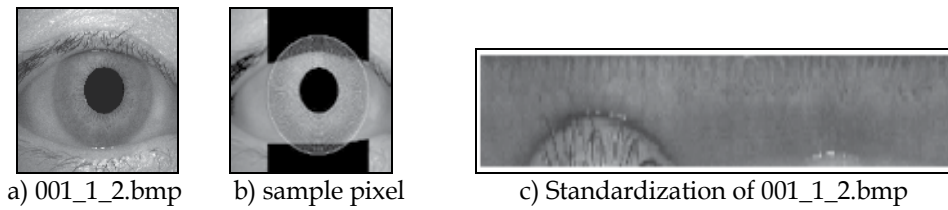
a) 001_1_2.bmp          b) sample pixel          c) Standardization of 001_1_2.bmp

Fig. 13. Standardized rectangular region



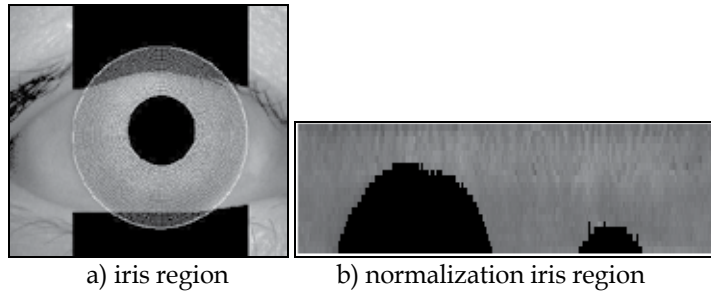a) iris region          b) normalization iris region

Fig. 14. Interference label in Iris normalization region: iris image after boundary localization and interference localization (left) and the corresponding normalization iris region with the labeled interference (right)

## 5. Iris feature extraction based on local binary pattern analysis

### 5.1 Iris feature extraction based on local binary pattern

Iris recognition algorithms main include Gabor filter method, local zero-crossing wavelet, independence complements analysis, and so on [Li Ma, 2004; Kwanghyuk Bae, 2003; Tian Qi-chuan, 2006; Seung-In Noh, 2002]. The binary ordinal of iris texture has become iris recognition frame. Iris binary template is stored as personal identify feature reference template in the future. From these algorithms, we learn of that iris recognition adopts local features to indicate iris pattern [Zhenan Sun, 2004], here, we introduce a new algorithm based on local binary pattern analysis for iris recognition.

Local Binary Pattern (LBP) is an easy-to-compute, robust local texture descriptor, and it has been shown to be promising in the computer vision field, including industrial inspection, motion analysis, and face recognition. In this paper, we show that LBP can solve iris feature extraction according the inherent intensity-related texture problem, is robust to some illumination and interference, and has potential for pattern recognitions [W.W.Boles, 1998; Devrim Unay, 2007; T. Ahonen, 2006].

For instance, in iris region, image intensity smoothly varies across an image. This intensity inhomogeneity, or socalled bias field, can significantly degrade the performance of some recognition algorithms. Because the bias field is locally smooth, we argue that it should not change the local structure. Furthermore, in iris acquisition inter- and intra-user misalignment of the images is a known problem. This misalignment problem may limit the application of automated match on iris images. In this case, rotation invariant descriptors may prevent some of those limitations.

LBP is a grayscale invariant local texture operator with powerful discrimination and low computational complexity. An LBP operator thresholds a neighborhood by the gray value of

its center ($g_c$) and represents the result as a binary code that describes the local texture pattern. The operator ($LBP_{P,R}$) is derived based on a symmetric neighbor set of $P$ members $g_p(p = 0, \cdots, P-1)$ within a circular radius of $R$.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{9}$$

where

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{10}$$

Fig. 15 illustrates the computation of $LBP_{8,1}$ for a single pixel in a rectangular $3 \times 3$ neighborhood. Binary feature is more robust than image magnitude character.
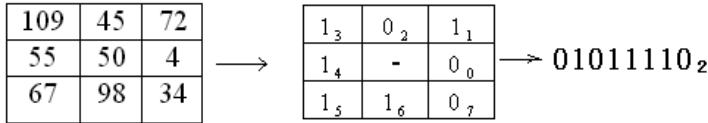
| 109 | 45 | 72 |
|-----|----|----|
| 55  | 50 | 4  |
| 67  | 98 | 34 |

$\longrightarrow$

| $1_3$ | $0_2$ | $1_1$ |
|-------|-------|-------|
| $1_4$ | -     | $0_0$ |
| $1_5$ | $1_6$ | $0_7$ |

$\longrightarrow 01011110_2$

Fig. 15. Example of computing $LBP_{8,1}$: a pixel neighbourhood (left), its thresholded version (middle), and the corresponding binary LBP pattern with the computed LBP code (right)

In the general definition, LBP is defined in a circular symmetric neighborhood that requires interpolation of intensity values for exact computation. In order to keep computation simple, in this study we decided to use the two rectangular neighborhoods as shown in Fig.16.



$LBP_{8,1}$ $\qquad$ $LBP_{16,2}$
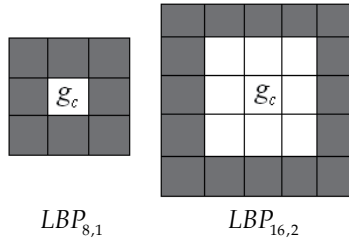
Fig. 16. The rectangular neighbourhoods of LBP used. Gray-shaded rectangles refer to the pixels belonging to the corresponding neightborhood

Rotation invariant patterns: The $LBP_{P,R}$ operator can produce $2^P$ different output values from $P$ neighbor pixels. As $g_0$ is always assigned to be the gray value of neighbor to the right of $g_c$, rotation will result in a different $LBP_{P,R}$ value for the same binary pattern. Because iris region normalization have transform iris circular region into rectangular region, there aren't the effect of rotation.

So, Iris feature extraction algorithm based on local binary pattern can be written as follows:

**Step 1.** According to iris normalization region $I_{M \times N}$, we can move $LBP_{P,R}$ to location $I(i,j)$ in $I_{M \times N}$, and then we can achieve LBP codes at position $(i,j)$, where $I(i,j) \in I_{M \times N}$, $i = 1, 2, \cdots, M, j = 1, 2, \cdots N$.

**Step 2.** We can construct iris feature template based on these LBP codes.



Fig. 17. Feature extraction based on LBP :( left-upper), iris normalization region (right-upper), and feature extraction results based on LBP (below)

## 5.2 Match score calculation

Because iris features belong to binary features, we can calculate match score by using vector similarity: if $A$ and $B$ are feature template vectors, then the similarity of $A$ and $B$ can be calculated as follows:

$$Similarity(A,B) = \frac{\langle A,B \rangle}{\|A\|\|B\|} \tag{11}$$

According to iris rotation, we can eliminate the effect of iris rotation by using shift operation on the binary pattern $A$ several times and assign the $Similarity(A^{shift},B)$ that is the largest in similarity scores under different shift case:

$$Similarity(A^{shift},B) = \max_{shift \in \{8i|i=1,2\cdots\}} \bigcup \frac{\langle A^{shift},B \rangle}{\|A\|\|B\|} \tag{12}$$

## 5.3 Iris feature selection

Feature selection can be used to improve classification performance [Linlin Shen, 2005].The eyelid, eyelash and facula can affect iris character, so features of this area aren't true features and they should be removed after boundary location. The threshold segmentation method is used to eliminate eyelash and facula, and radon transform based on arc and line is used to detect eyelid. The purpose of interfere detection is to achieve iris region accurately. While we transform iris region into rectangular region, we also label the interference points in rectangular region. Those features of the interference part labelled should be removed in calculating pattern similarity score. Only those stable features is helpful to iris classify, however, how to know whether features are stable or not still is a research problem.

We think stable features is those features which are captured every time, exist in many iris image, and aren't affected by interference, so we want to train classifier by selecting stable features from intra serial iris images.

**Step 1.** Select training data;

**Step 2.** Calculate intra similarity of iris, the largest similarity score is defined to the similarity score of two iris feature pattern alignment under different shift cases, then label these same features of templates as stable features and label other features as unstable features;

**Step 3.** These templates with stable labels and unstable labels can be regarded as identity reference template.

In Fig.18, (a1) and (a2) are two eye images of the same person, and they are captured in different condition. Fig.18 (b1) and (b2) are normalization regions of (a1) and (a2), (c1) and (c2) are feature extraction results, and (d) is the feature selection result. In Fig.18 (c1), (c2) and (d), red pixels indicate '1' features, green pixels indicate '0' features, and black pixels in Fig.18 (d) indicate unstable features. Those unstable features should be removed when we calculate matching score.


a1) 001_1_1.bmp       a2) 001_1_2.bmp


b1) normalization region of 001_1_1.bmp    b2) normalization region of 001_1_2.bmp


c1) feature codes of 001_1_1.bmp      c2) feature codes of 001_1_2.bmp


d) feature selection result
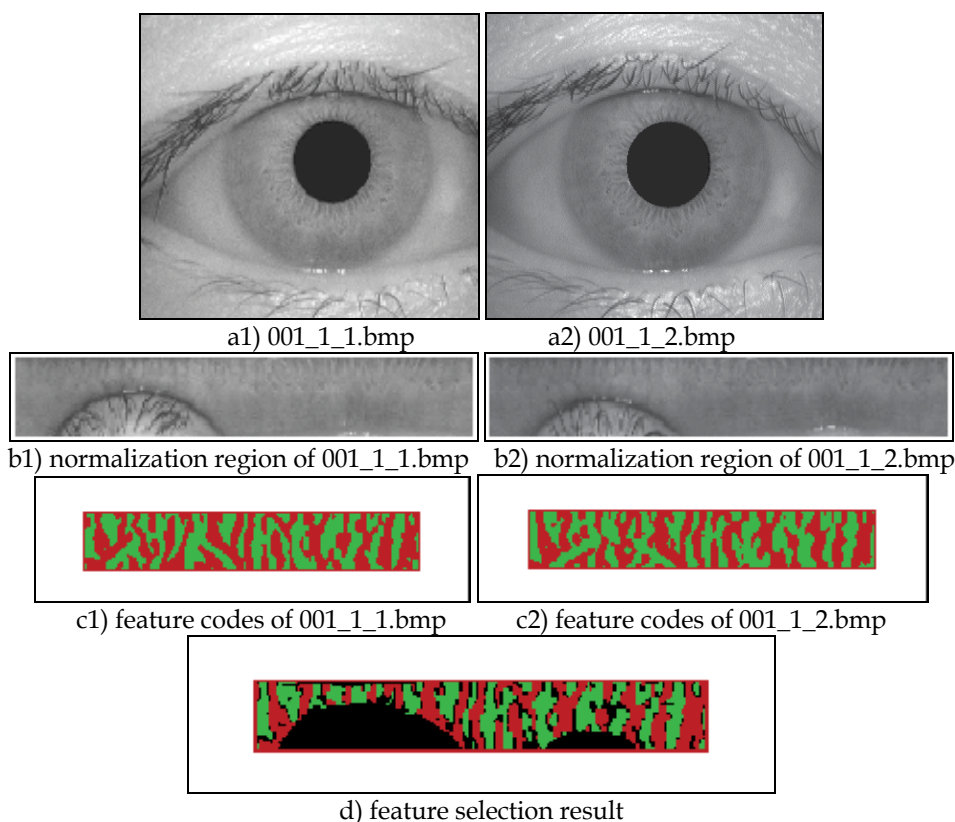
Fig. 18. Stable feature selection results

Feature selection is a method for improving pattern classification accuracy, in this paper, feature selection is to select stable and effectiveness features by comparing intra images of the same iris. In Fig.18, It can be seen that unstable features mostly appear at boundary of texture shape change. Those regions whose gray value changes mildly are less affected by illumination, therefore stable features are low frequency signal practically. Points on boundary, denote unstable characters, are easy to modify the sign of LBP detection. So selection key feature from multi-images is to describe image pattern based on stable low frequency character. Therefore, stable features should be found in images that have large difference, based on pattern mapping, complicated pattern could be mapped in the space that represented as some key features. It can enhance template description ability by making high dimension expression into low dimension.

### 5.4 Decision

Iris recognition used to for identity verification, after capturing iris image and extracting features, recognition system based on iris pattern can recognise user's identity using his/her iris by comparing similarity with a classification threshold.

Classification threshold can be achieved depend on classification performance on training data and demand on recognition real-time performance and recognition correct rate.

If whole reference templates are $W = \{w_1, w_2, \cdots, w_c\}$, *Threshold* is classification threshold, then the steps of iris recognition procedure can be written as follows:

**Step 1.** Extraction features $P$ of a person iris images based on LBP;

**Step 2.** Calculation similarity degree $Similarity(P, T_{w_i})$ between $P$ and reference identity templates $T_{w_i}$;

**Step 3.** If $Similarity(P, T_{w_i}) > Threshold$, then the identity of the person can be confirmed and he/she belongs to $w_i$; otherwise $i = i + 1$, go to step2 until $i > c$.

## 6. Experimental results

Under this condition, the difference of similarity distribution between same iris pattern and the different ones can be converted into classifying as the same pattern and the difference pattern. The unrecognized images have 108 pattern 540 images totally, so there are 540 samples in same pattern, and there are 28890 samples in difference pattern [Chinese Academy of Sciences-institute of automation, 2003].

Fig.19 is the comparison of iris classification results based on LBP and feature selection and only based on LBP without feature selection. We can learn that there are very low error classification rate in Fig.19 (b1) and (b2). Tab.1 shows the performance of iris recognition. The distribution of similarity and error rate show the algorithm has good recognition performance.

Performance of classification algorithm includes FAR (False Accept Rate), FRR (False Reject Rate), EER (Equal Error Rate, it is defined to the value of FAR and FRR when FAR=FRR) and Decide-ability $D$ defined as formula [John Dangman, 1993 & 2002; Shinyoung Lim, 2001]:

$$D = \frac{\left| \mu_{same} - \mu_{difference} \right|}{\sqrt{(\sigma_{same}^2 + \sigma_{difference}^2)/2}} \tag{13}$$

$D$ denotes classified quality of training patterns, in which $\mu_{same}$, $\mu_{difference}$, $\sigma_{same}^2$, $\sigma_{difference}^2$ are the means and variances of same class pattern and different class patterns respectively.

| Performance | LBP Without feature selection | LBP with feature selection |
|---|---|---|
| EER | 0.5% | 0.45% |
| D | 7.1 | 8.9 |

Table 1. Performance of LBP

## 7. Conclusion

We introduce a whole iris recognition system. Especially, when using only one image for iris recognition, the registered iris features are susceptible to illumination, contrast and other factors, resulting in a large number of trustless feature points in feature template. A method

based on LBP features extraction and selection from multiple images is presented, in which stable features are selected to describe the iris identity while the unreliable feature points are labelled in enrolment template. Research show LBP is robust to bias field and rotation. Overall performance of the iris recognition system is satisfied.



| a1) The similarity distribution | a2) The distribution of error rate |



| b1) The distribution of similarity | b2) The distribution of error rate |
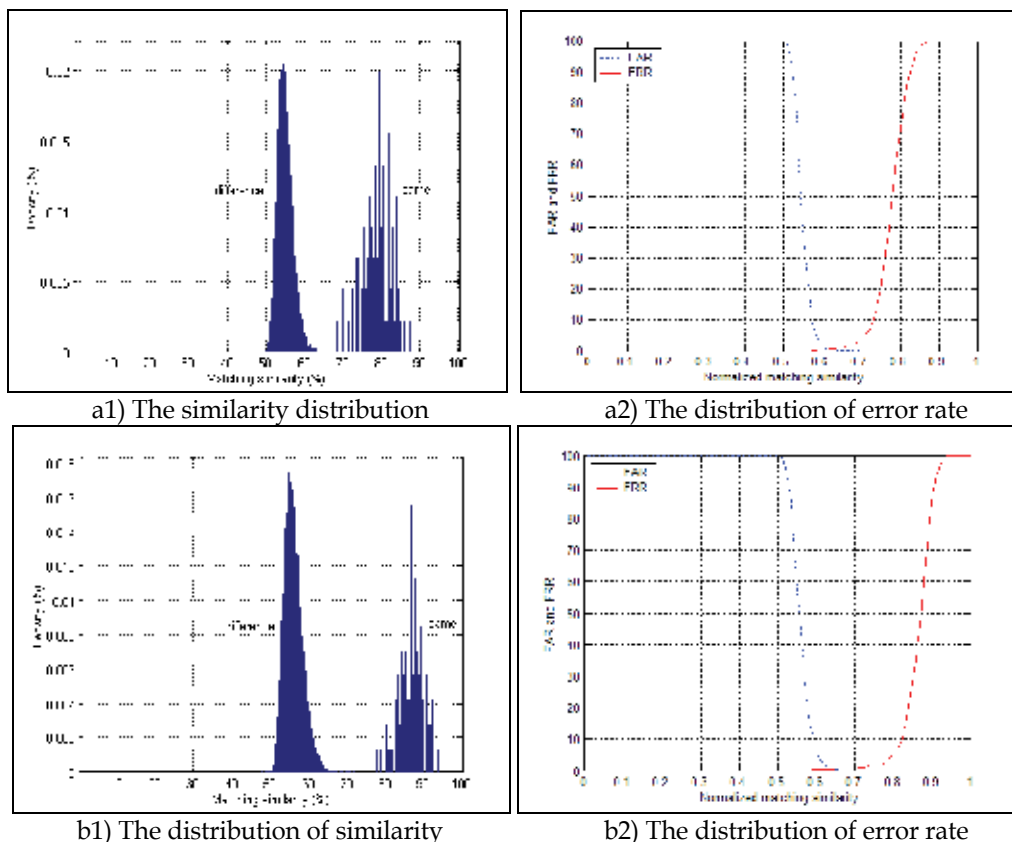
Fig. 19. Experimental results, (a1) and (a2) is iris classification results without feature selection, (b1) and (b2) is iris classification results with feature selection

## 8. Acknowledgment

## 9. References

Richard P W, Jane C A, et al. A machine-vision system for iris recognition,*Machine Vision and Application (S0932-8092),* 1996,9(1): 1-8

John Daugman. High confidence visual recognition of persons by a test of statistical independence, *IEEE Transactions on pattern Analysis and Machine Intelligence (S0162-8828)*, 1993, 15 (11): 1148-1161

Richard P W. Iris recognition: an emerging biometric technology, *Proceeding of IEEE (S0018-9219)*, 1997, 85(9): 1348-1363

Tian Qi-chuan, Pan Quan, et al. Fast iris boundary localization algorithm supervised by pupil center, *Journal of System Simulation*,2006,18(7):1777-1780

Wai-Kin Kong, David Zhang. Detecting eyelash and reflection for accurate iris segmentation, *International Journal of Pattern Recognition and Artificial Intelligence*,2003,17(6): 1025-1034

Tian Qi-chuan, Pan Quan, et al. Approach of noise detecting and processing in iris recognition, *Computer Engineering*, 2006,32(2):172-174

W. Kong, D.Zhang. Accurate iris segmentation based on novel reflection and eyelash detection model, *Proceeding of 2001 international Symposium on intelligent multimedia, Video and Speech Processing*, Hong Kong, 2001

Libor Masek. Recognition of human iris patterns for biometric identification. *http://www.csse.uwa.edu.au/~pk/studentprojects/libor/*

C.H.Daouk, L.A.El-Esber, et al. Iris recognition, *IEEE ISSPIT2002*:558-562

Raul Sanchez-Reillo, Carmen Sanchez-Avila. Iris recognition with low template size, *AVBPA 2001*,LNCS 2091:324-329

Li Ma, Tieniu Tan, Yunhong Wang, et al.Local intensity variation analysis for iris recognition, *Pattern recognition,* 2004,37: 1287-1298

Kwanghyuk Bae, Seungin Noh, et al. Iris feature extraction using independent component analysis, *AVBPA 2003*,LNCS 2688:838-844

Tian Qi-chuan, Pan Quan, et al. Iris feature extraction algorithm based on local zero-crossing detection, *Journal of Electronics & Information Technology.*2006,28(8): 1452-1460

Seung-In Noh, Kwanghuk Pae, et al. Multiresolution independent component analysis for iris identification, *The 2002 International Technical Conference on Circuits/Systems, Computers and Communications,* Phuket, Tailand, July 2002

Zhenan Sun, Tieniu Tan, et al. Robust encoding of local ordinal measures: a general framework of iris recognition, *ECCV workshop on Biometric Authentication* 2004

W.W.Boles and B.Boashash. A human identification technique using images of the iris and wavelet transform, *IEEE Transactions on Signal Processing,* 1998, 46(4): 1185-1188

Devrim Unay, Ahmet Ekin, et al. Robustness of Local Binary Patterns in Brain MR Image Analysis, *Proceedings of the 29th Annual International Conference of the IEEE EMBS,* Lyon, France,August 23-26, 2007,2098-2101

T. Ahonen, A. Hadid, et al. Face description with local binary patterns: application to face recognition, *IEEE Ttrans on Pattern Analysis and Machine Intelligence,* 2006(28):2037-2041

Linlin Shen, Li Bai, et al. Gabor feature selection for face recognition using improved adaBoost learning, *IWBRS2005,* 2005, LNCS 3781:39-49

John Dangman. High confidence visual recognition of person by a test of statistical independence, *IEEE Trans Pattern Anal Machine Intelligence,* 1993, 15 (11): 1148-1161.

John Dangman. The importance of being random: statistical principles of iris recognition, *Pattern recognition,* 2002,1-13

Shinyoung Lim, Kwanyong Lee, et al. Efficient iris recongition through improvement of feature vector and classifier, *ETRI Journal,*2001,23(2) :61-70

Chinese Academy of Sciences-institute of automation. Database of 756 grayscale eye images (version1.0). *http://www.sinobiometrics.com,*2003

# The State-of-the-Art in Iris Biometric Cryptosystems

Christian Rathgeb and Andreas Uhl

*Multimedia Signal Processing and Security Lab (WaveLab),*
*Department of Computer Sciences, University of Salzburg*
*A-5020 Salzburg, Austria*

## 1. Introduction

In 1984 a photographer named Steve McCurry traveled to Pakistan in order to document the ordeal of Afghanistan's refugees, orphaned during the Soviet Union's bombing of Afghanistan. In the refugee camp Nasir Bagh, which was a sea of tents, he took a photograph of a young girl approximately at the age of 13. The portrait by Steve McCurry turned out to be one of those images that sears the heart, and in June 1985 it ran on the cover of National Geographic. The girl's sea green eyes have captivated the world since then and because no one knew her name she became known as the "Afghan girl".

In January 2002, 17 year later, a team from National Geographic Television brought McCurry back to Pakistan to search for the girl with green eyes. When they showed her picture around Nasir Bagh, the still standing refugee camp, there were a number of women who came forward and identified themselves erroneously as the famous Afghan girl. In addition, after being shown the 1985 photo, a handful of young men falsely claimed the Afghan girl as their wife. The team was able to finally confirm her identity using the iris feature analysis of the Federal Bureau of Investigation (FBI), which matched her iris patterns to those of the photograph with almost full certainty (Braun, 2003). Her name was Sharbat Gula, then around the age of 30, and she had not been photographed since. The revealment of Sharbat Gula's identity manifested the strength of iris recognition technologies. Figure 1 (a) shows the original image of her which was printed on the cover of National Geographic in 1985 and another portrait taken in 2002 which was used for identification.

Iris biometrics refers to high confidence recognition of a person's identity by mathematical analysis of the random patterns that are visible within the iris of an eye from some distance (Daugman, 2004). Figure 1 (b) shows a good-quality NIR infrared image of an human eye captured by an iris recognition device. In contrast to other biometric characteristics, such as fingerprints (Maltoni et al., 2009), the iris is a protected internal organ whose random texture is complex, unique, and very stable throughout life. Because the randomness of iris patterns has very high dimensionality, recognition decisions are made with confidence levels, high enough to support rapid and reliable exhaustive searches through national-sized databases.

Until now iris recognition has been successfully applied in diverse access control systems managing large-scale user database. For instance, in the UK project IRIS (Iris Recognition Immigration System), over a million frequent travelers have registered with the system for automated border-crossing using iris recognition. IRIS is in operation on different UK

<table>
<tr><td align="center">(a)</td><td align="center">(b)</td></tr>
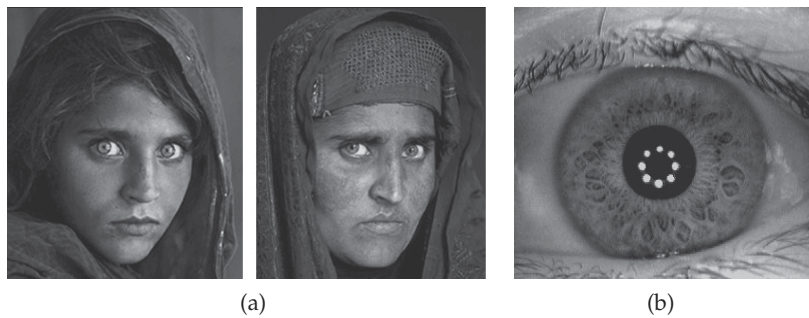</table>

Fig. 1. (a) Sharbat Gula at the age of approximately 13 and 30 (taken from Daugman (2011))
(b) Sample image of a person's iris (taken from CASIAv3-Interval iris database).

airports including London Heathrow and Gatwick, Manchester and Birmingham. While the registration process usually takes between 5 and 10 minutes enrolled passengers do not even need to assert their identity. They just look at the camera in the automated lanes crossing an IRIS barrier in about 20 seconds. Until now several different large-scale iris recognition systems have been successfully deployed.

However, the broad use of biometric technologies have raised many concerns. From the privacy perspective most concerns arise from the storage and misuse of biometric data (Cimato et al., 2009). Besides the fact that users share biometric traits rather reluctantly biometric applications are often considered as a threat to privacy (Jain et al., 2006). These concerns are well-justified since physiological biometric traits are irrevocable in the sense that these cannot be modified during the lifetime of a data subject. In case biometric traits are compromised these become useless and biometric authentication based on these traits must not be considered secure anymore. A rather recent field of research which is referred to as Biometric Cryptosystems (Uludag et al., 2004) is expected to increase the confidence in biometric authentication systems as this technology offers novel solutions to biometric template protection (Jain, Flynn & Ross, 2008) and, thus, preserves the privacy of biometric traits. Approaches to biometric cryptosystems have been proposed for different biometric characteristics (including behavioral modalities) where the best performing systems are based on iris (Cavoukian & Stoianov, 2009a). As iris biometric cryptosystems have rather recently emerged a systematic classification and in-depth discussion of existing approaches is presented in this chapter. Furthermore, custom implementations of existing systems are presented and evaluated on open databases. Based on the experimental study the reader is provided with a in-depth discussion of the state-of-the-art in iris biometric cryptosystems, which completes this work.

The remainder of this chapter is organized as follows: in Sect. 2 the fundamentals of iris recognition are briefly summarized. Subsequently, biometric template protection is motived and template protection schemes are categorized in Sect. 3. In Sect. 4 related work with respect to iris biometric cryptosystems is reviewed. Then custom implementations of key approaches to iris biometric cryptosystems are presented and evaluated in Sect. 5. A comprehensive discussion of iris biometric cryptosystems including advantages and applications, the current state-of-the-art, and open research issues, is presented in Sect. 6. Finally, a summary and a conclusion is given in Sect. 7.

## 2. Fundamentals of (iris) biometric recognition

The term biometrics refers to "automated recognition of individuals based on their behavioral and biological characteristics" (ISO/IEC JTC1 SC37). Several physiological as well as behavioral biometric characteristics have been used (Jain, Flynn & Ross, 2008) such as fingerprints, iris, face, hand, voice, gait, etc., depending on types of applications. Biometric traits are acquired applying adequate sensors and distinctive features are extracted to form a biometric template in the enrollment process. During verification (authentication process) or identification (identification can be handled as a sequence of verifications and screenings) the system processes another biometric measurement which is compared against the stored template(s) yielding acceptance or rejection.

Several metrics exist when measuring the performance of biometric systems. Widely used factors include False Rejection Rate (FRR), False Acceptance Rate (FAR), and Equal Error Rate (EER) (Jain et al., 2004). While the FRR defines the "proportion of verification transactions with truthful claims of identity that are incorrectly rejected", the FAR defines the "proportion of verification transactions with wrongful claims of identity that are incorrectly confirmed" (ISO/IEC FDIS 19795-1). The Genuine Acceptance Rate (GAR) is defined as, GAR = 1 - FRR. As score distributions overlap, FAR and FRR intersect at a certain point, defining the EER of the system. According to intra- and inter-class accumulations generated by biometric algorithms, FRRs and FARs are adjusted by varying system thresholds. In general decreasing the FRR ($\hat{=}$ increasing the GAR) increases the FAR and vice versa.

### 2.1 Iris recognition

Among all biometric characteristics the pattern of an iris texture is believed to be the most distinguishable among different people (Bowyer et al., 2007). The iris is the annular area between the pupil and the sclera of the eye. Breakthrough work to create iris recognition algorithms was proposed by J. G. Daugman, University of Cambridge Computer Laboratory. Daugman's algorithms (Daugman, 2004) for which he holds key patents form the basis of the vast majority of today's commercially dispread iris recognition systems. According to these algorithms generic iris recognition systems consist of four stages: (1) image acquisition, (2) iris image preprocessing, (3) iris texture feature extraction, and (4) feature matching.

With respect to the image acquisition good-quality images are necessary to provide a robust iris recognition system. Hence, one disadvantage of iris recognition systems is the fact that users have to cooperate fully with the system. At preprocessing the pupil and the outer boundary of the iris are detected. An example of this process is illustrated in Figure 2 (a)-(b). Subsequently, the vast majority of iris recognition algorithms un-wrappes the iris ring to a normalized rectangular iris texture, shown in Figure 2 (c). To complete the preprocessing the contrast of the resulting iris texture is enhanced applying histogram stretching methods. Based on the preprocessed iris texture, which is shown in Figure 2 (d) feature extraction is applied. Again, most iris recognition algorithms follow the approach of Daugman by extracting a binary feature vector, which is commonly referred to as iris-code. While Daugman suggests to apply 2D-Gabor filters in the feature extraction stage plenty of different methods have been proposed (for further details see Bowyer et al. (2007)). An example of an iris-code is shown in Figure 2 (e). In most matching methods iris-codes are compared by applying the bit-wise XOR-operator to count miss-matching bits such that the Hamming distance indicates the grade of dissimilarity (small values indicate high similarity). In order to compensate against head tilts template alignment is achieved by applying circular shifts in both directions where the minimal Hamming distance between two iris-codes refers to an optimal alignment.
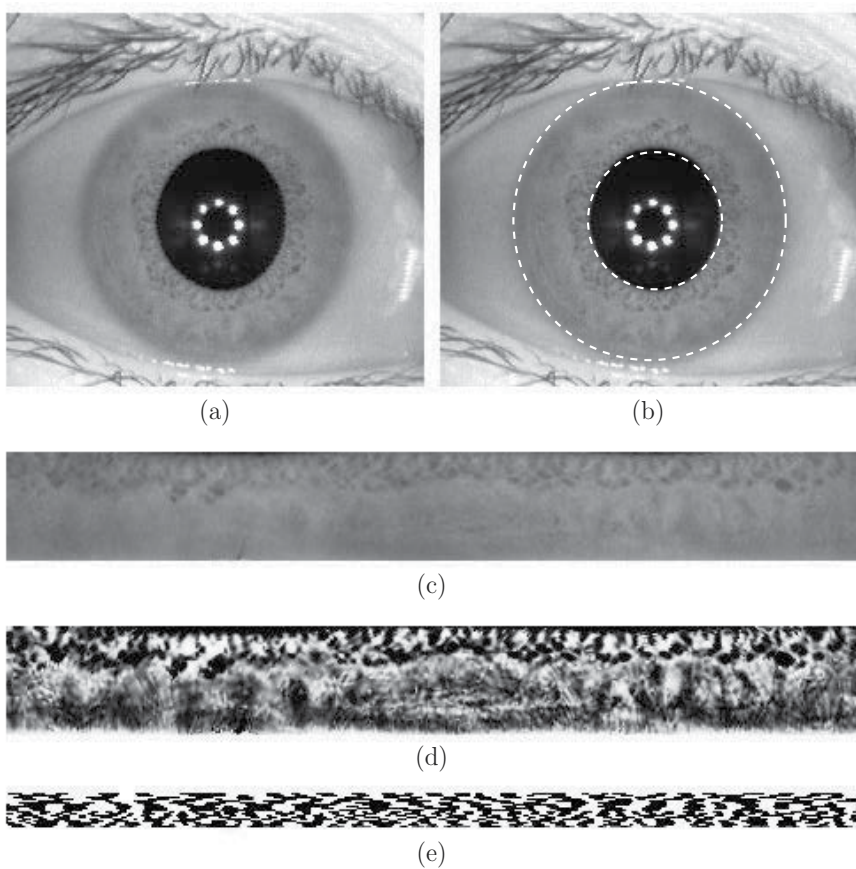
Fig. 2. Common processing chain in iris recognition: (a) image of eye (b) detection of pupil and iris (c) unrolled iris texture (d) preprocessed iris texture (e) sample iris-code.

Hence, the matching of iris-codes can be performed in an efficient process, which can be parallelized easily. In contrast to other biometric systems based on different modalities which require a more complex matching procedure thousands of comparisons can be done within one second. With respect to biometric recognition systems operating in identification mode iris recognition algorithms are capable of handling large-scale databases. In addition, potential occlusions originating from eye lids or eye lashes are masked out during matching by storing a bit-mask generated in the preprocessing step.

## 3. Biometric template protection

Biometric cryptosystems are designed to securely bind a digital key to a biometric or generate a digital key from a biometric (Cavoukian & Stoianov, 2009a). Biometric cryptosystems release cryptographic keys which are associated with the biometric traits of registered users. Hence, biometric cryptosystems offer solutions to secure biometric-based key management as well as biometric template protection. Since authentication is performed indirectly by verifying key validities the system does not need to store the original biometric templates. In addition, most
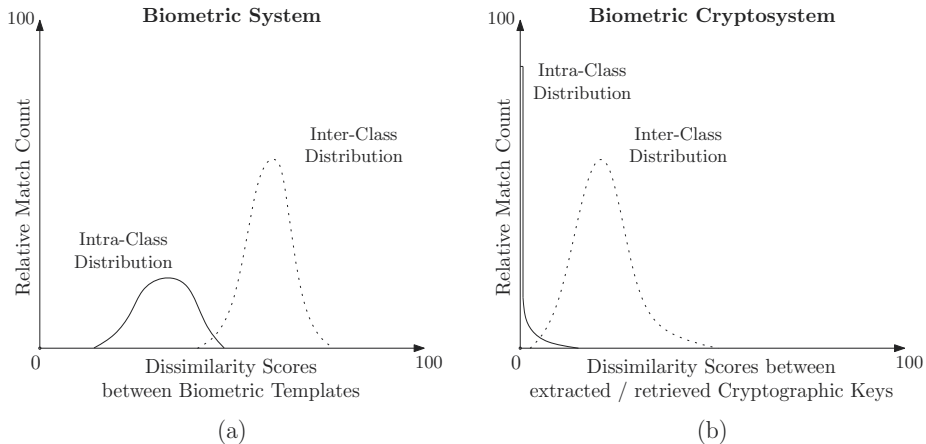
Fig. 3. Performance measurement: (a) generic biometric system (b) biometric cryptosystem in which the return of hundred percent correct keys indicates genuine users.

biometric cryptosystems provide mechanisms to update these keys at any time so that users are able to apply different keys at different applications.

In the context of biometric cryptosystems the meanings of the aforementioned biometric performance metrics change. Threshold-based authentication is eliminated since acceptance requires the generation or retrieval of a hundred percent correct key. The fundamental difference within performance measurements regarding generic biometric systems and biometric cryptosystems is illustrated in Figure 3 (a)-(b). The FRR of a biometric cryptosystem defines the percentage of incorrect keys returned to genuine users (again, GAR = 1 - FRR). By analogy, the FAR defines the percentage of correct keys returned to non-genuine users. Compared to existing biometric systems, biometric cryptosystems tend to reveal noticeably inferior performance (Uludag et al., 2004). This is because within biometric cryptosystem the enrolled template is not seen and, therefore, can not be adjusted for the direct comparison with a given biometric sample. In addition, biometric recognition systems are capable of setting more precise thresholds to adjust the tolerance of the system.

The majority of biometric cryptosystems require the storage of biometric dependent public information which is referred to as helper data (Jain, Nandakumar & Nagar, 2008) (biometric cryptosystems are often referred to as helper data-based methods). Due to the natural variance in biometric measurements it is not possible for most biometric traits to extract a cryptographic key directly. Additionally, the application of helper data provides revocability of the generated keys. The stored helper data, which must not reveal any significant information about the original biometric signal, is applied to extract a key. The comparison of biometric templates is performed indirectly by verifying the validity of keys, so that the output of the authentication process is either a key or a failure message. The verification of keys represents a biometric comparison in the cryptographic domain (Jain et al., 2005). Hence, biometric cryptosystems can be applied as a means of biometric template protection (Jain, Nandakumar & Nagar, 2008). Based on how helper data are derived, biometric cryptosystems are further classified as key-binding or key-generation systems as shown in Figure 4 (a)-(b).

### 3.1 Key-generation and key-binding

Within a key-binding scheme helper data is obtained by binding a chosen cryptographic key to biometric features. As a result of the binding process a fusion of the secret key and the
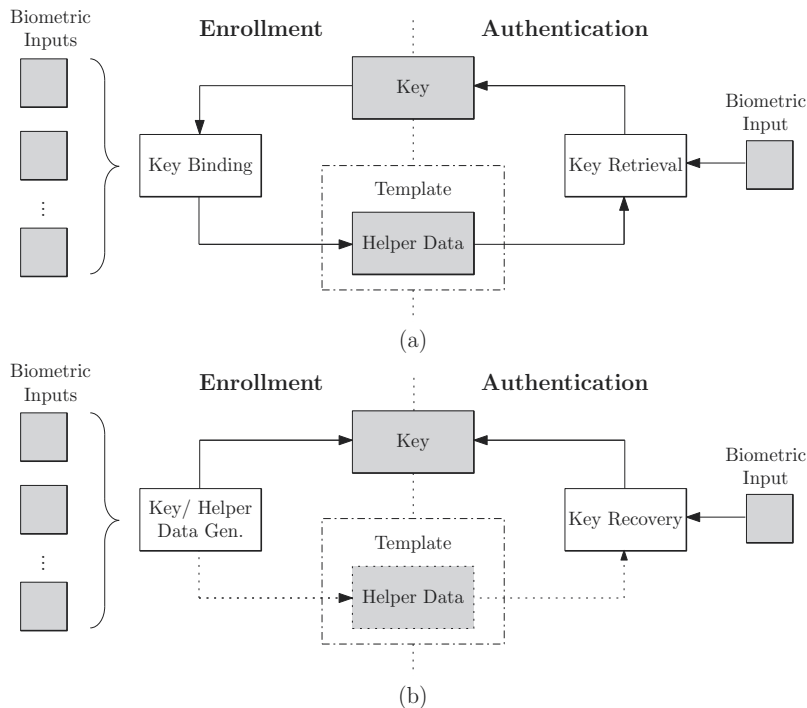
Fig. 4. Key-Binding and Key-Generation: (a) the basic concept of a key-binding scheme (b) the basic concept of a key-generation scheme.

biometric template is stored, which does neither reveal any information about the key nor about the original biometric data. Applying an appropriate key retrieval algorithm, keys are extracted out of the stored helper data during biometric authentication (Uludag et al., 2004). The cryptographic key is independent of biometric features so that the key is updateable while an update of the key usually requires re-enrollment in order to generate new helper data. The general operation mode of a key-binding scheme is illustrated in Figure 4 (a).

In a key-generation scheme the helper data is derived only from the biometric template so that the cryptographic key is directly generated from the helper data and a given biometric sample (Jain, Nandakumar & Nagar, 2008). While the storage of helper data is not obligatory the majority of proposed key-generation schemes do store helper data. If key-generation schemes extract keys without the use of any helper data these keys can not be changed in case of compromise, unless the key-generation algorithm is undergone a change. This means, stored helper data allows updating cryptographic keys. Key generation schemes in which helper data are applied are also called "fuzzy extractors" or "secure sketches" as described in (Dodis et al., 2004) (for both primitives, formalisms are defined). A fuzzy extractor reliably extracts a uniformly random string from a biometric input while public information is used to reconstruct that string from another biometric measure. In contrast, in a secure sketch public helper data is applied to recover the original biometric template from another biometric input. In Figure 4 (b) the basic concept of a generic key-generation scheme is illustrated.

Several approaches to biometric cryptosystems can be used as both, key-generation schemes and key-binding schemes (e.g. Juels & Sudan (2002); Juels & Wattenberg (1999)). Hybrid approaches which make use of both of these basic concepts (e.g. Boult et al. (2007)) have been
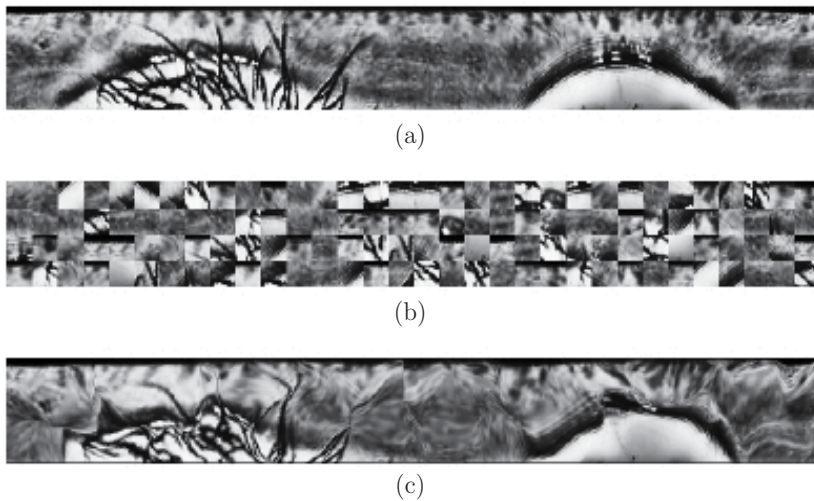
(a)



(b)



(c)

Fig. 5. Example of cancellable iris biometrics: (a) original iris texture. (b) transformed iris texture based on block permutation. (c) transformed iris texture based on surface folding.

proposed as well. Furthermore, schemes which declare diverse goals such as enhancing the security of any kind of existing secret (e.g. Monrose et al. (1999)) have been introduced. In contrast to key-binding and key-generation schemes so-called key-release schemes represent a loose coupling of biometric authentication and key-release (Uludag et al., 2004). While the loose coupling of biometrics and the cryptographic system allows to exchange both components easily this loose coupling emerges as a great drawback as well, since it implies the separate storage of biometric templates and keys and, thus, offers more vulnerabilities to conduct attacks. Key-release schemes are hardly appropriate for high security applications and not usually considered a biometric cryptosystem at all.

### 3.2 Cancellable biometrics

Cancellable biometrics consist of intentional, repeatable distortions of biometric signals based on transforms which provide a matching of biometric templates in the transformed domain (Ratha et al., 2001). Focusing on iris biometrics several different transforms have been proposed (e.g. Hämmerle-Uhl et al. (2009); Zuo et al. (2008)), in addition generic approaches which could be applied to iris have been presented (e.g. BioHashing in Teoh et al. (2004)). These transforms are designed in a way that it should be impossible to recover the original biometric data. An example of generating cancellable iris biometrics is shown in Figure 5. Additionally, the correlation of several transformed templates should not reveal any information about the original biometrics. If the transformed biometric data is compromised, transform parameters are changed, which means, the biometric template is updated. To prevent impostors from tracking users by cross-matching databases it is suggested to apply different transforms for different applications. Approaches to cancellable biometrics represent solutions to biometric template protection, too. In contrast to biometric cryptosystems cancellable biometrics do not associate cryptographic keys with biometric data.

### 3.3 Privacy aspects

Most concerns against biometric technologies arise from the abuse of personal data as well as the permanent tracking and observation of activities (Cimato et al., 2009). As previously
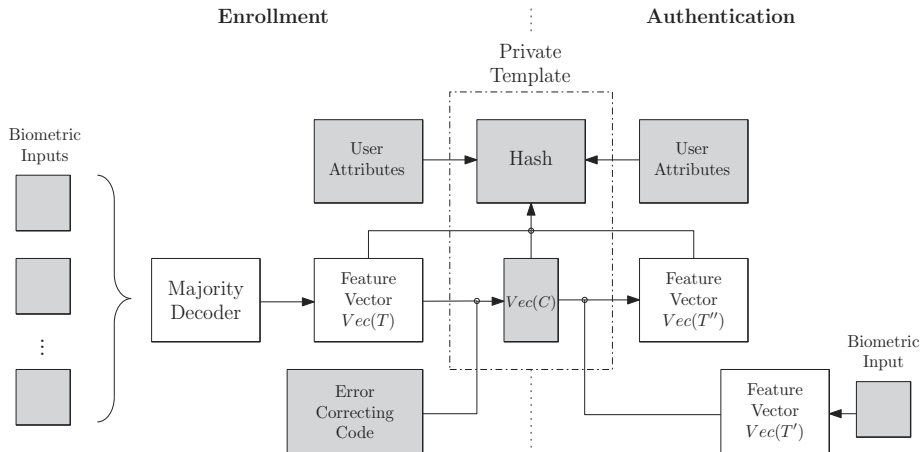
Fig. 6. Private template scheme: the basic operation mode of the private template scheme in which the biometric template itself serves as cryptographic key.

mentioned, in case raw biometric traits are compromised these become useless and biometric authentication based on these traits must not be considered secure anymore. Biometric cryptosystems (as well as cancellable biometrics) are expected to increase the confidence in biometric authentication systems. This is because these technologies offer solutions to biometric template protection (Jain, Nandakumar & Nagar, 2008) and, thus, preserve the privacy of biometric traits. The fundamental feature within both technologies is that comparisons of biometric templates are performed in the encrypted domain (Uludag et al., 2004). Compared to template encryption techniques, where biometric templates are exposed during each authentication, here biometric templates are permanently secured. Furthermore, different versions of secured biometric templates can be applied in different applications (Ratha et al., 2001) which prevents from the tracking of users. In case of compromise the reconstruction of original biometric data is hardly feasible for impostors while protected biometric templates are easily updated. Additionally, biometric cryptosystems provide techniques to biometric dependent key-release.

## 4. Iris biometric cryptosystems

Biometric cryptosystems have been designed for diverse physiological and behavioral biometric characteristics (further details can be found in Cavoukian & Stoianov (2009a)). In the following subchapters key concepts to biometric cryptosystems which have been applied to iris biometrics are discussed in detail.

### 4.1 Private template scheme
The first to propose an iris biometric key-generation scheme were Davida et al. (Davida et al., 1998; 1999) in their "private template" scheme, in which the biometric template itself (or a hash value of it) serves as a cryptographic key. The basic operation mode of a private template scheme, which requires the storage of helper data, is illustrated in Figure 6. In the private template scheme helper data are error correction check bits which are applied to correct faulty bits of a given iris-code. In the enrollment process $M$ 2048-bit iris-codes are generated which are put through a majority decoder to reduce the Hamming distance between iris-codes. This majority decoder computes the vector $Vec(V) = (V_1, V_2, ..., V_n)$ for a $n$-bit code vector, denoted
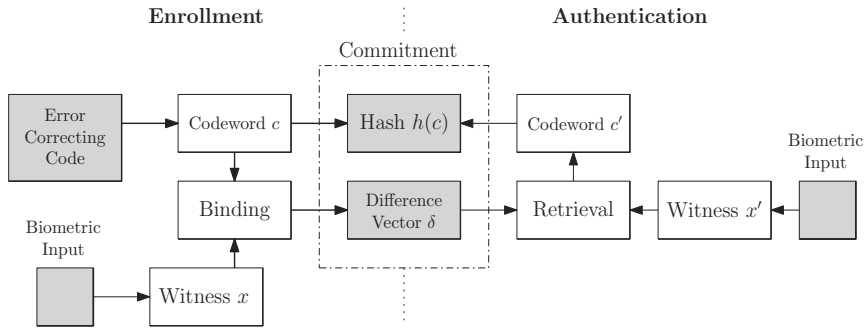
Fig. 7. Fuzzy commitment scheme: the concept of the fuzzy commitment scheme in which a key, prepared with error correction codes, is bound to a binary feature vector.

by $Vec(v_i) = (v_{i,1}, v_{i,2}, ..., v_{i,n})$, where $V_j = majority(v_{1,j}, v_{2,j}, ..., v_{M,j})$. The common metric for $V_j$ is the majority of 0's and 1's of bit $j$ from each of the $M$ vectors. A majority decoded iris-code $T$, denoted by $Vec(T)$, is concatenated with check digits $Vec(C)$, to generate $Vec(T)||Vec(C)$. The check digits $Vec(C)$ are part of an error correction code. Then a hash value Hash(Name, Attr, $Vec(T)||Vec(C)$) is generated, where Name is the user's name, Attr are public attributes of the user and Hash($\cdot$) is a hash function. Finally, an authorization officer signs this hash resulting in Sig(Hash(Name, Attr, $Vec(T)||Vec(C)$)). During authentication several iris-codes are captured and majority decoded resulting in $Vec(T')$. With the use of $Vec(C)$ which is stored as part of the template (helper data) the corrected template $Vec(T'')$ is constructed. In the end, Hash(Name, Attr, $Vec(T'')||Vec(C)$) is calculated and Sig(Hash(Name, Attr, $Vec(T'')||Vec(C)$)) is checked. Experimental results are omitted and it is commonly expected that the proposed system reveals poor performance due to the fact that the authors restrict to the assumption that only 10% of bits of an iris-code change among different iris images of a single data subject. But in general, average intra-class distances of iris-codes lie within 20-30%. Additionally, implementations of the proposed majority decoding technique (e.g. in Yang & Verbauwhede (2007)) were not found to decrease intra-class distances to that extent.

### 4.2 Fuzzy commitment scheme

Juels and Wattenberg (Juels & Wattenberg, 1999) combined techniques from the area of error correcting codes and cryptography to achieve a type of cryptographic primitive entitled "fuzzy commitment" scheme. A fuzzy commitment scheme consists of a function $F$, used to commit a codeword $c \in C$ and a witness $x \in \{0,1\}^n$. The set $C$ is a set of error correcting codewords $c$ of length $n$ and $x$ represents a bit stream of length $n$, termed witness (biometric data). The difference vector of $c$ and $x$, $\delta \in \{0,1\}^n$ where $x = c + \delta$, and a hash value $h(c)$ are stored as the commitment termed $F(c, x)$ (secure biometric template). Each $x'$, which is sufficiently "close" to $x$, according to an appropriate metric, should be able to reconstruct $c$ using the difference vector $\delta$ to translate $x'$ in the direction of $x$. A hash of the result is tested against $h(c)$. With respect to biometric key-binding the system acquires a witness $x$ at enrollment, selects a codeword $c \in C$, calculates the commitment $F(c, x)$ ($\delta$ and $h(c)$) and stores it in a database. At the time of authentication, a witness $x'$ is acquired and the system checks whether $x'$ yields a successful decommitment. Figure 7 shows the basic operation mode of a fuzzy commitment scheme.

The fuzzy commitment scheme was applied to iris-codes by Hao et al. (Hao et al., 2006). In their scheme 2048-bit iris-codes are applied to bind and retrieve 140-bit cryptographic keys

prepared with Hadamard and Reed-Solomon error correction codes. Hadamard codes are applied to eliminate bit errors originating from the natural variance and Reed-Solomon codes are applied to correct burst errors resulting from distortions. The system was tested with 700 iris images of 70 subjects achieving a GAR of 99.53% and a zero FAR. These are rather impressive results which were not achieved until then. In order to provide a more accurate error correction decoding in an iris-based fuzzy commitment scheme, which gets close to a theoretical bound obtained by Bringer et al. (Bringer et al., 2007; 2008), the authors apply two-dimensional iterative min-sum decoding. Within their approach a matrix is created where lines as well as columns are formed by two different binary Reed-Muller codes. Thereby a more efficient decoding is available. Adapting the proposed scheme to the standard iris recognition algorithm of Daugman a GAR of 94.38% is achieved for the binding of 40-bit cryptographic keys. Due to the fact that Bringer et al. apply their scheme to diverse data sets a more significant performance evaluation than that of Hao et al. (Hao et al., 2006) is provided. Rathgeb and Uhl (Rathgeb & Uhl, 2009b) provide a systematic approach to the construction of fuzzy commitment schemes based on iris biometrics. After analyzing the error distribution in iris-codes of different iris recognition algorithms, Reed-Solomon and Hadamard codes are applied, similar to Hao et al. (Hao et al., 2006). Experimental results provide a GAR of 95.08% and 93.43% for adopting the fuzzy commitment approach to two different iris recognition algorithms. In other further work (Rathgeb & Uhl, 2009a) the authors apply a context-based reliable component selection in order to extract cryptographic keys from iris-codes which are then bound to Hadamard codewords resulting in a GAR of 93.47% at zero FAR. Besides, different techniques to improve the performance of iris based fuzzy commitment schemes have been proposed (Rathgeb & Uhl, 2010a; Zhang et al., 2009).

### 4.3 Fuzzy vault scheme

One of the most popular biometric cryptosystems called "fuzzy vault" was introduced by Juels and Sudan (Juels & Sudan, 2002). The key idea of the fuzzy vault scheme is to use an unordered set $A$ to lock a secret key $k$, yielding a vault, denoted by $V_A$. If another set $B$ overlaps largely with $A$, $k$ can be reconstructed, which means the vault $V_A$ is unlocked. The vault is created applying polynomial encoding and error correction. During the enrollment phase a polynom $p$ is selected which encodes the key $k$ in some way (e.g. the coefficients of $p$ are formed by $k$), denoted by $p \leftarrow k$. Then the elements of $A$ are projected onto the polynom $p$, i.e. $p(A)$ is calculated. Additionally, so-called chaff points are added in order to obscure genuine points of the polynom. The set of all points, called $R$, forms the template. To achieve a successful authentication another set $B$ needs to overlap with $A$ sufficiently. If this is the case it is possible to locate many points in $R$ that lie on $p$. Applying error correction codes $p$ can be reconstructed and, hence, $k$. The components of a fuzzy vault scheme are illustrated in Figure 8. The security of the whole scheme lies in the infeasibility of the polynomial reconstruction and the number of applied chaff points. In contrast to the aforementioned fuzzy commitment scheme the main advantage of this approach is the feature of order invariance, i.e. to be able to cope with unordered data. For example, the minutiae points of a captured fingerprint are not necessarily ordered from one measurement to another with respect to specific directions due to fingerprint displacement, rotations and contrast changes. If features are formed by relative positions, unordered sets of minutiae points will still be able to reconstruct the secret.

Apart from fingerprints, which is the most apart biometric characteristic for this scheme (e.g. in Clancy et al. (2003); Nandakumar et al. (2007)) iris biometrics have been applied in fuzzy vault schemes by Lee et al. (Lee, Bae, Lee, Park & Kim, 2007). Since iris features are usually
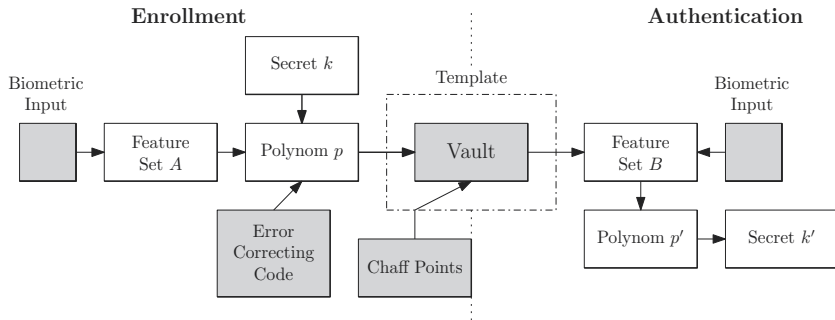
Fig. 8. Fuzzy vault scheme: the basic operation mode of the fuzzy vault scheme in which a unordered set of biometric features is mapped on to a secret polynom.

ordered, in order to obtain an unordered set of features, independent component analysis is applied obtaining a GAR of 99.225% at a zero FAR. Wu et al. (Wu et al., 2008a;b) proposed a fuzzy vault based on iris biometrics as well. After image acquisition and preprocessing the iris texture is divided into 64 blocks where for each block the mean gray scale value is calculated resulting in 256 features which are normalized to integers to reduce noise. At the same time, a Reed-Solomon code is generated and subsequently the feature vector is translated to a cipher key using a hash function. The authors report a FAR of 0.0% and a GAR of approximately 94.45% for a total number of over 100 persons. Reddy and Babu (Reddy & Babu, 2008) enhance the security of a classic fuzzy vault scheme based on iris biometrics by adding a password with which the vault as well as the secret key is hardened. In experiments a system which exhibits a GAR of 92% and a FAR of 0.03% is hardened, resulting in a GAR of 90.2% and a FAR of 0.0%. However, if passwords are compromised the systems security decreases to that of a standard one, thus the FAR of 0.0% was calculated under unrealistic preconditions (Rathgeb & Uhl, 2010b). A multi-biometric fuzzy vault based on fingerprint and iris was proposed by Nandakumar and Jain (Nandakumar & Jain, 2008). The authors demonstrate that a combination of biometric modalities leads to better recognition performance and higher security. A GAR of 98.2% at a FAR of $\sim$ 0.01%, while the corresponding GAR values of the iris and fingerprint fuzzy vaults are 88.0% and 78.8%, respectively.

## 5. Implementation of iris biometric cryptosystems

In oder to provide a technical insight to the implementation iris biometric cryptosystems different iris biometric feature extraction algorithms are applied to different variations of iris-based fuzzy commitment schemes. The construction of these schemes is described in detail and the resulting systems are evaluated on a comprehensive data set.

### 5.1 Biometric databases

Experiments are carried out using the CASIAv3-Interval iris database[1] as well as on the IIT Delhi iris database v1[2], two public available iris datasets. Both databases consist of good quality NIR illuminated indoor images, sample images of both databases are shown in Figure

---

[1] The Center of Biometrics and Security Research, CASIA Iris Image Database, URL:
http://www.idealtest.org
[2] The IIT Delhi Iris Database version 1.0, URL:
http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm

| Data Set | Persons | Classes | Images | Resolution |
|---|---|---|---|---|
| CASIAv3-Interval | 250 | 396 | 2639 | 320×280 |
| IITDv1 | 224 | 448 | 2240 | 320×240 |
| Total | 474 | 844 | 4879 | – |

Table 1. Databases applied in experimental evaluations.
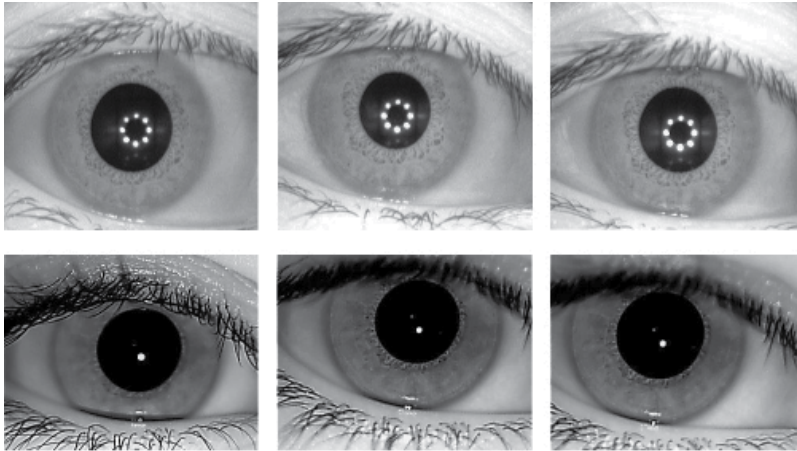


Fig. 9. Sample images of single classes of the CASIAv3-Interval database (above) and the IITDv1 database (below).

9. These datasets are fused in order to obtain one comprehensive test set. The resulting test set consists of over 800 classes as shown in Table 1 allowing a comprehensive evaluation of the proposed systems.

### 5.2 Preprocessing and feature extraction

In the preprocessing step the pupil and the iris of a given sample image are located applying Canny edge detection and Hough circle detection. More advanced iris detection techniques are not considered, however, as the same detection is applied for all experimental evaluations obtained results retain their significance. Once the pupil and iris circles are localized, the area between them is transformed to a normalized rectangular texture of $512 \times 64$ pixel, according to the "rubbersheet" approach by Daugman (Daugman, 2004). As a final step, lighting across the texture is normalized using block-wise brightness estimation. An example of a preprocessed iris image is shown in Figure 2 (e).

In the feature extraction stage we employ custom implementations of two different algorithms used to extract binary iris-codes. The first one was proposed by Ma et al. (Ma et al., 2004). Within this approach the texture is divided into 10 stripes to obtain 5 one-dimensional signals, each one averaged from the pixels of 5 adjacent rows, hence, the upper $512 \times 50$ pixel of preprocessed iris textures are analyzed. A dyadic wavelet transform is then performed on each of the resulting 10 signals, and two fixed subbands are selected from each transform resulting in a total number of 20 subbands. In each subband all local minima and maxima above a adequate threshold are located, and a bit-code alternating between 0 and 1 at each extreme point is extracted. Utilizing 512 bits per signal, the final code comprises a total number of $512 \times 20 = 10240$ bits.

| Algorithm | $p$ | $\sigma$ | DoF (bit) | EER (%) |
|-----------|-----|----------|-----------|---------|
| Ma et al. | 0.4965 | 0.0143 | 1232 | 0.4154 |
| Log-Gabor | 0.4958 | 0.0202 | 612 | 0.6446 |

$p$ ... mean Hamming distance
$\sigma$ ... standard deviation
DoF ... degrees of freedom

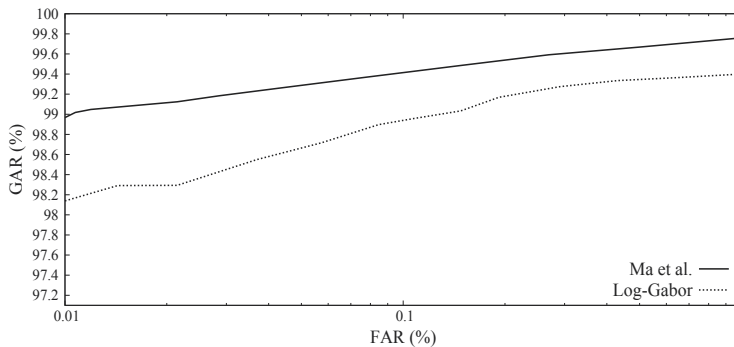Table 2. Benchmark Values of applied Feature Extraction Algorithms.



Fig. 10. Receiver operation characteristic curves for the algorithm of Ma et al. and the Log-Gabor feature extraction.

The second feature extraction method follows an implementation by Masek[3] in which filters obtained from a Log-Gabor function are applied. Here a row-wise convolution with a complex Log-Gabor filter is performed on the texture pixels. The phase angle of the resulting complex value for each pixel is discretized into 2 bits. To have a code comparable to the first algorithm, we use the same texture size and row-averaging into 10 signals prior to applying the one-dimensional Log-Gabor filter. The 2 bits of phase information are used to generate a binary code, which therefore is again $512 \times 20 = 10240$ bit. This algorithm is somewhat similar to Daugman's use of Log-Gabor filters, but it works only on rows as opposed to the 2-dimensional filters used by Daugman.

A major issue regarding biometric cryptosystems is the entropy of biometric data. If cryptographic keys are associated with biometric features which suffer from low entropy these are easily compromised (e.g. by performing false acceptance attacks). In fact it has been shown that the iris exhibits enough reliable information to bind or extract cryptographic keys, which are sufficiently long to be applied in generic cryptosystems (Cavoukian & Stoianov, 2009a). A common way of measuring the entropy of iris biometric systems was proposed in Daugman (2003). By calculating the mean $p$ and standard deviation $\sigma$ of the binomial distribution of iris-code Hamming distances the entropy of the iris recognition algorithm, which is referred to as "degrees of freedom", is defined as $p \cdot (1 - p)/\sigma^2$. For both algorithms these magnitudes are summarized in Table 2 including the equal error rates (EERs) for the entire dataset. As can be seen both algorithms provide enough entropy to bind and retrieve at least 128 bit cryptographic keys. The receiver operation characteristic (ROC) curve of both algorithms are plotted in Figure 10. For the algorithm of Ma et al. and Masek a GAR of 98.98% and 98.18% is obtained at a FAR of 0.01%, respectively. While both recognition systems obtain EERs below

---

[3] L. Masek: Recognition of Human Iris Patterns for Biometric Identification, University of Western Australia, 2003, URL: http://www.csse.uwa.edu.au/~pk/studentprojects/libor/sourcecode.html
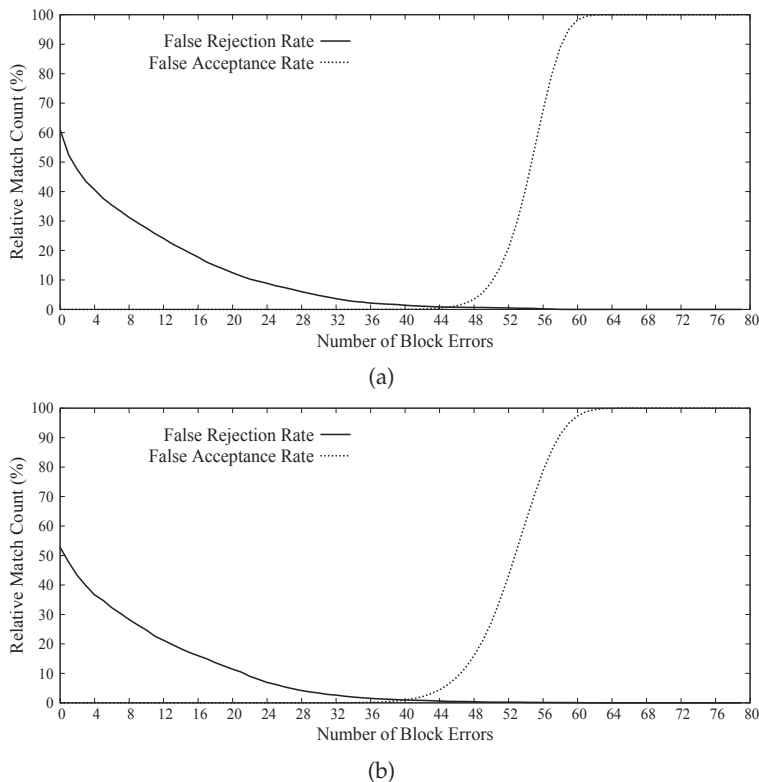
Fig. 11. False rejection rate and false acceptance rates for the fuzzy commitment scheme of Hao et al. for the feature extraction of (a) Ma et al. and (b) the Log-Gabor algorithm.

1% the recognition performance is expected to decrease for the according fuzzy commitment schemes (Uludag et al., 2004).

### 5.3 Fuzzy commitment schemes

The first fuzzy commitment scheme follows the approach of Hao et al. (Hao et al., 2006). In the original proposal a 140-bit cryptographic key is encoded with Hadamard and Reed-Solomon codes. While Hadamard codes are applied to correct natural variance between iris-codes Reed-Solomon codes handle remaining burst errors (resulting from distortions such as eyelids or eyelashes). For the applied algorithm of Ma et al. and the Log-Gabor feature extraction we found that the application of Hadamard codewords of 128-bit and a Reed-Solomon code $RS(16, 80)$ reveals the best experimental results for the binding of 128-bit cryptographic keys (Rathgeb & Uhl, 2009b). At key-binding, a $16 \cdot 8 = 128$ bit cryptographic key $R$ is first prepared with a $RS(16, 80)$ Reed-Solomon code. The Reed-Solomon error correction code operates on block level and is capable of correcting $(80 - 16)/2 = 32$ block errors. Then the 80 8-bit blocks are Hadamard encoded. In a Hadamard code codewords of length $n$ are mapped to codewords of length $2^{n-1}$ in which up to 25% of bit errors can be corrected. Hence, 80 8-bit codewords are mapped to 80 128-bit codewords resulting in a 10240-bit bit stream which is bound with the iris-code by XORing both. Additionally, a hash of the original key $h(R)$ is stored as second part of the commitment. At authentication key retrieval is performed by XORing an extracted iris-code with the first part of the commitment. The resulting bit

| Algorithm | GAR (%) | FAR (%) | Corrected Blocks |
|-----------|---------|---------|------------------|
| Ma et al. | 96.35 | 0.0095 | 32 |
| Log-Gabor | 95.21 | 0.0098 | 27 |

Table 3. Summarized experimental results for the fuzzy commitment scheme of Hao et al.
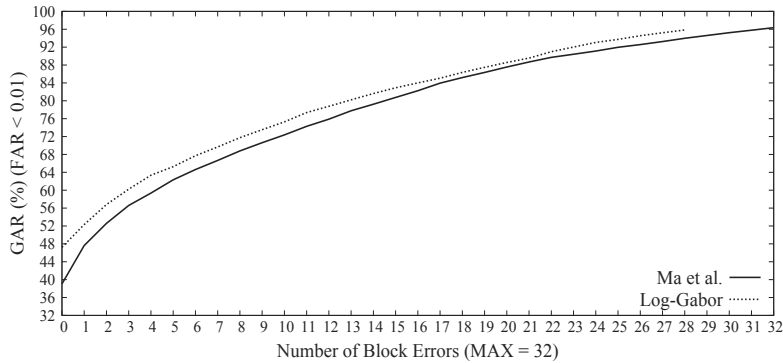


Fig. 12. Genuine acceptance rate for the fuzzy commitment scheme of Hao et al. for the feature extraction of Ma et al. and the Log-Gabor algorithm.

stream is decoded applying Hadamard decoding and Reed-Solomon decoding afterwards. The resulting key $R'$ is then hashed and if $h(R') = h(R)$ the correct key $R$ is released. Otherwise an error message is returned.

The second fuzzy commitment scheme was proposed by Bringer et al. (Bringer et al., 2008). Motivated by their observation that the system in Hao et al. (2006) does not hold the reported performance rates on data sets captured under unfavorable conditions a more effective error correction decoding is suggested. The proposed technique which is referred to as Min-Sum decoding presumes that iris-codes of 2048 bits are arranged in a two-dimensional manner. In the original system a 40-bit key $R$ is encoded with a two-dimensional Reed-Muller code such that each 64-bit line represents a codeword and each 32-bit column represents a codeword, too. To obtain the helper data $P$ the iris-code is XORed with the two-dimensional Reed-Muller code. It is shown that by applying a row-wise and column-wise Min-Sum decoding the recognition performance comes near to practical boundaries. In order to adopt the system to the applied feature extractions 8192 bits of iris-codes are arranged in 64 lines of 128 bits (best experimental results are achieved for this configuration). To generate the commitment a 56-bit cryptographic key $R$ is used to generate the error correction matrix. Since Reed-Muller codes are generated using Hadamard matrices and each line and each column of the resulting two-dimensional code has to be a codeword, $2^n + 1$ codewords define a total number of $2^{n+1}$ codewords. Due to the structure of the error correction code $2^{7 \cdot 8} = 2^{56}$ possible configurations of the $128 \times 64 = 8192$-bit error correction code exist. At authentication a given iris-code is XORed with the commitment and the iterative Min-Sum decoding is applied until the correct key $R$ is retrieved or a predefined threshold is reached.

With respect to iris biometrics cryptosystems these variations of the fuzzy commitment scheme represent the best performing systems in literature (Cavoukian & Stoianov, 2009a).
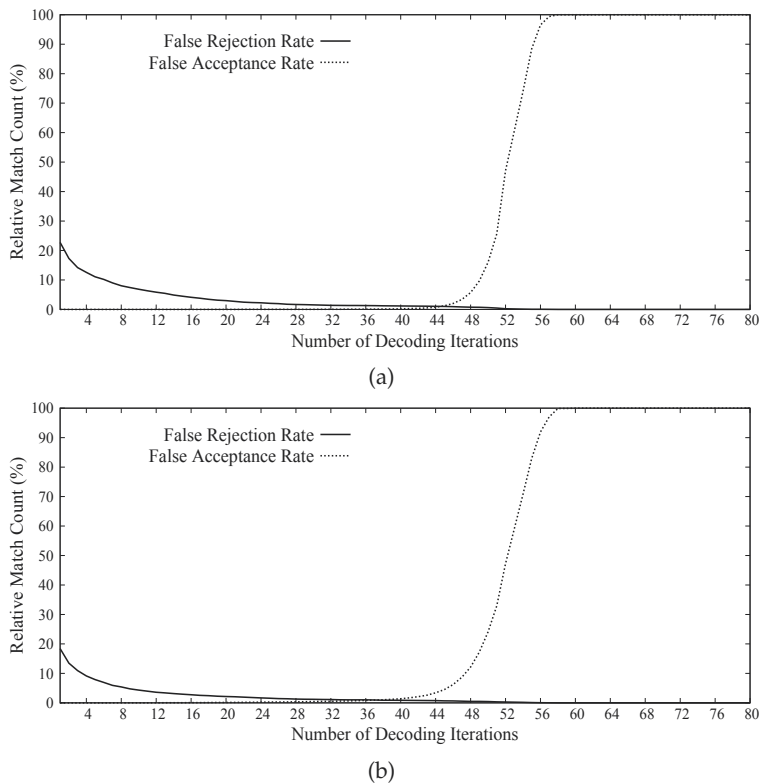
(a)



(b)

Fig. 13. False rejection rate and false acceptance rates for the fuzzy commitment scheme of Bringer et al. for the feature extraction of (a) Ma et al. and (b) the Log-Gabor algorithm.

## 5.4 Performance evaluation

According to the fuzzy commitment scheme of Hao et al. the FRR and FAR for the algorithm of Ma et al. is plotted in Figure 11 (a) according to the number of corrected block errors after Hadamard decoding. In contrast to generic biometric systems only discrete thresholds can be set in order to distinguish between genuine and non-genuine persons. The characteristics of the FRR and FAR for the algorithm of Ma et al. is rather similar to that of the Log-Gabor feature extraction which is plotted in Figure 11 (b). Block-level error correction is necessary for both feature extraction methods in order to correct burst errors. As previously mentioned, for both algorithms the maximal number of block errors that can be handled by the Reed-Solomon code is 32, which suffices in both cases. In Figure 12 the GARs for both feature extraction methods are plotted according to the number of corrected block errors where the according FARs are required to be less than 0.01%. For the algorithm of Ma et al. a GAR of 96.35% and a FAR of 0.0095% is obtained where the full error correction capacity is exploited. With respect to the Log-Gabor feature extraction a GAR of 95.21% and a FAR of 0.0098% are achieved where 27 block errors are corrected, respectively. Table 3 summarizes obtained performance rates for both iris biometric feature extraction methods. Like in the original iris recognition systems the algorithm of Ma et al. performs better than the Log-Gabor feature extraction. However, as it was expected for both methods accuracy decreases. This is because error correction is designed to correct random noise while iris-codes do not exhibit a uniform distribution of mismatching bits (distinct parts of iris-code comprise more reliable bits than

| Algorithm | GAR (%) | FAR (%) | Decoding Iterations |
|-----------|---------|---------|---------------------|
| Ma et al. | 96.99 | 0.01 | 20 |
| Log-Gabor | 93.06 | 0.01 | 6 |

Table 4. Summarized experimental results for the fuzzy commitment scheme of Bringer et al.
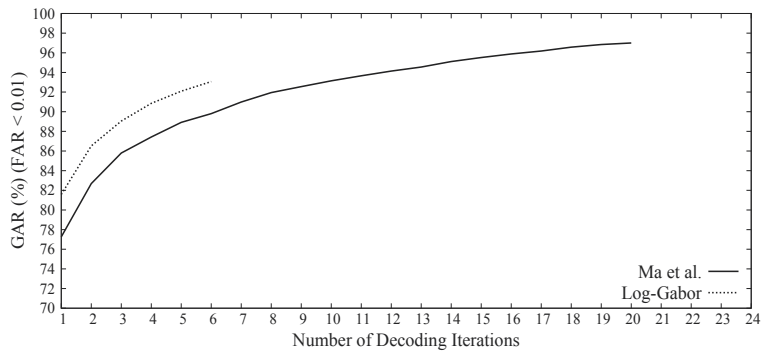


Fig. 14. Genuine acceptance rate for the fuzzy commitment scheme of Bringer et al. or the feature extraction of Ma et al. and the Log-Gabor algorithm.

others (Rathgeb et al., 2010)) and, in addition, decision thresholds can not be set as precise as in generic biometric systems. Furthermore, the resulting fuzzy commitment schemes show worse performance rates than those reported in Hao et al. (2006), which is because those results were achieved for a rather small test set of iris images captured under ideal conditions. Therefore, the achieved results in this work are more significant as these are obtained from different test sets for different feature extraction methods.

In the second fuzzy commitment scheme which follows the approach in Bringer et al. (2008) iterative decoding of rows and columns of two-dimensional iris-codes is performed. Figure 13 (a)-(b) shows the FRRs and FARs for both feature extraction methods according to the number of decoding iteration, necessary to retrieve the correct key. Again, the characteristics of FRRs and FARs are rather similar for both algorithms. In Figure 14 the GARs for both feature extraction methods are plotted according to the number of decoding iterations where the according FARs are required to be less than 0.01%. For the algorithm of Ma et al. and the Log-Gabor feature extraction a GAR of 96.99% and 93.06% are obtained according to a FAR of 0.01%, respectively. Table 4 summarizes obtained performance rates for the fuzzy commitment scheme of Bringer et al. for both iris biometric feature extraction methods. For the applied dataset the scheme of Bringer et al. does not show any significant improvement compared to that of Hao et al., although it is believed that the scheme of Bringer et al. works better on non-ideal iris images since error correction is applied iteratively. In other words, in the scheme of Hao et al. error correction capacities may be hit to the limit under non-ideal conditions while in the scheme of Bringer et al. a larger amount of decoding iterations is expected to yield successful key retrieval. However, as a two-dimensional arrangement of error correction codewords is required the according retrieved keys are rather short compared to the approach of Hao et al. In contrast to the first fuzzy commitment scheme results reported in Bringer et al. (2008) coincide with the ones obtained.

For both implementations of iris-based fuzzy commitment schemes obtained performance rates are promising and by all means comparable to those reported in literature. Furthermore,

the systematic construction of these schemes, which does not require any custom-built optimizations, underlines the potential of iris biometric cryptosystems.

## 6. Discussion

After presenting key technologies in the areas of biometric cryptosystems and an implementations of iris-based fuzzy commitment schemes a concluding discussion is done. For this purpose major advantages and potential applications are discussed. An overview of the performance of existing state-of-the-art approaches is given and, finally, open issues and challenges are discussed.

### 6.1 Advantages and applications

Biometric cryptosystems offer several advantages over conventional biometric systems. Major advantages can be summarized as follows:

- Template protection: within biometric cryptosystems the original biometric template is obscured such that a reconstruction is hardly feasible.

- Biometric-dependent key release: biometric cryptosystems provide key release mechanisms based on the presentation of biometric data.

- Pseudonymous biometric authentication: authentication is performed in the encrypted domain and, thus, is pseudonymous.

- Revocability of biometric templates: several instances of secured templates can be generated by binding or generating different keys.

- Increased security: biometric cryptosystems prevent from several traditional types of attacks against biometric systems (e.g. substitution attacks).

- Higher social acceptance: due to the above mentioned security benefits the social acceptance of biometric applications is expected to increase.

These advantages call for several applications. In order to underline the potential of biometric cryptosystems one essential use case is discussed, pseudonymous biometric databases. Biometric cryptosystems meet the requirements of launching pseudonymous biometric databases (Cavoukian & Stoianov, 2009a) since these provide a comparison of biometric templates in the encrypted domain. Stored templates (helper data) do not reveal any information about the original biometric data. Additionally, several differently obscured templates can be used in different applications. At registration the biometric data of the user is employed as input for a biometric cryptosystem. The user is able to register with several applications where different templates are stored in each database (as suggested in Ratha et al. (2001)). Depending on the type of application further user records are linked to the template. These records should be encrypted where decryption could be applied based on a released key. Figure 15 shows the scenario of constructing an pseudonymous biometric database.

Due to the fact that stored helper data does not reveal information about the original biometric data high security in terms of template protection is provided. Since comparison is performed in the encrypted domain biometric templates are not exposed during comparisons (Jain, Nandakumar & Nagar, 2008). This means that the authentication process is fully pseudonymous and, furthermore, activities of users are untraceable because different secured templates are applied in different databases.
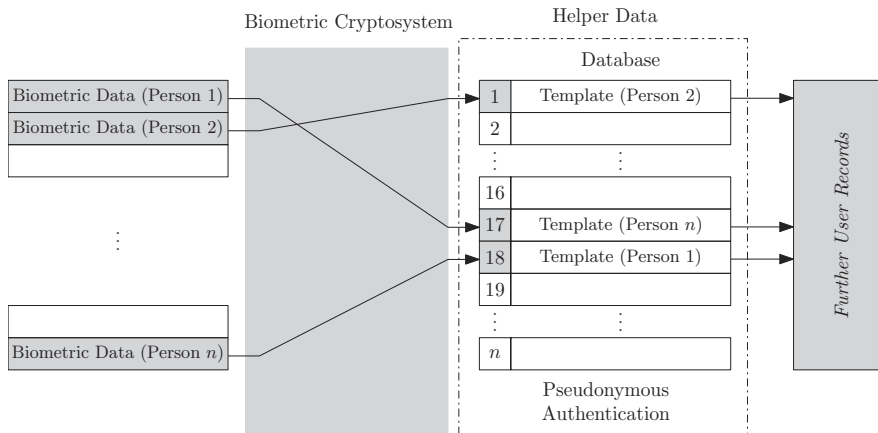
Fig. 15. Pseudonymous databases: users authenticate indirect at a biometric database and access their stored records in a secure way, such that the activities of a user are not traceable.

## 6.2 The state-of-the-art

In early approaches to iris biometric cryptosystems such as the private template scheme (Davida et al., 1998), performance rates were omitted while it has been found that these schemes suffer from serious security vulnerabilities (Uludag et al., 2004). Representing one of the simplest key-binding approaches the fuzzy commitment scheme (Juels & Wattenberg, 1999) has been successfully applied to iris and other biometrics, too. Iris-codes, generated by applying common feature extraction methods, seem to exhibit sufficient information to bind and retrieve cryptographic keys, long enough to be applied in generic cryptosystems. The fuzzy vault scheme (Juels & Sudan, 2002) which requires real-valued feature vectors as input has been applied to iris biometrics as well. The best performing iris-biometric cryptosystems with respect to the applied concept and datasets are summarized in Table 5. Most existing approaches reveal GARs above 95% according to negligible FARs. While the fuzzy commitment scheme represents a well-elaborated approach which has been applied to various feature extraction methods on different data sets (even on non-ideal databases), existing approaches to iris-based fuzzy vaults are evaluated on rather small datasets which does not coincide with high security demands.

With respect to other biometric modalities performance rates of key concepts of biometric cryptosystems are summarized in Table 6. As can be seen iris biometric cryptosystems outperform the majority of these schemes which do not provide practical performance rates as well as sufficiently long keys. Thus, it is believed that the state-of-the-art in biometric cryptosystems in general is headed by iris-based approaches.

## 6.3 Open issues and challenges

With respect to the design goals, biometric cryptosystems offer significant advantages to enhance the privacy and security of biometric systems providing reliable biometric authentication at an high security level. However, several new issues and challenges arise deploying these technologies (Cavoukian & Stoianov, 2009b). One fundamental challenge, regarding both technologies, represents the issue of alignment, which significantly effects recognition performance. Biometric templates are obscured within biometric cryptosystems and, thus, the alignment of these secured templates is highly non-trivial. While focusing on biometric recognition align-invariant approaches have been proposed for several biometric

| Authors | Scheme | GAR / FAR | Data Set | Keybits |
|---|---|---|---|---|
| Hao et al. (2006) | | 99.58 / 0.0 | 70 persons | 140 |
| Bringer et al. (2007) | FCS | 94.38 / 0.0 | ICE 2005 | 40 |
| Rathgeb & Uhl (2010a) | | 95.08 / 0.0 | CASIA v3 | 128 |
| Lee, Choi, Toh, Lee & Kim (2007) | FVS | 99.225 / 0.0 | BERC v1 | 128 |
| Wu et al. (2008a) | | 94.55 / 0.73 | CASIA v1 | 1024 |

FCS ... fuzzy commitment scheme
FVS ... fuzzy vault scheme

Table 5. Experimental results of the best performing Iris-Biometric Cryptosystems.

| Authors | Biometric Modality | GAR / FAR | Data Set | Keybits | Remarks |
|---|---|---|---|---|---|
| Clancy et al. (2003) | Fingerprint | 70-80 / 0.0 | not given | 224 | pre-alignment |
| Nandakumar et al. (2007) | Fingerprint | 96.0 / 0.004 | FVC2002-DB2 | 128 | 2 enroll sam. |
| Feng & Wah (2002) | Online Sig. | 72.0 / 1.2 | 750 persons | 40 | – |
| Vielhauer et al. (2002) | Online Sig. | 92.95 / 0.0 | 10 persons | 24 | – |
| Monrose et al. (2001) | Voice | < 98.0 / 2.0 | 90 persons | $\sim 60$ | – |
| Teoh et al. (2004) | Face | 0.0 / 0.0 | ORL, Faces94 | 80 | non-stolen token |

Table 6. Experimental results of key approaches to Biometric Cryptosystems based on other biometric characteristics.

characteristics, so far, no suggestions have been made to construct align-invariant iris biometric cryptosystems.

The iris has been found to exhibit enough reliable information to bind or extract cryptographic keys at practical performance rates, which are sufficiently long to be applied in generic cryptosystems. Other biometric characteristics such as voice or online-signatures (especially behavioral biometrics) were found to reveal only a small amount of stable information (see Table 6). While some modalities may not be suitable to construct a biometric cryptosystem these can still be applied to improve the security of an existing secret. Additionally, several biometric characteristics can be combined to construct multi-biometric cryptosystems (e.g. Nandakumar & Jain (2008)), which have received only little consideration so far. Thereby security is enhanced and feature vectors can be merged to extract enough reliable data. While for iris biometrics the extraction of a sufficient amount of reliable features seems to be feasible it still remains questionable if these features exhibit enough entropy. In case extracted data do not meet the requirement of high discriminativity the system becomes vulnerable to several attacks. This means, biometric cryptosystems which tend to release keys which suffer from low entropy are easily compromised (e.g. performing false acceptance attacks). Besides the vulnerability of releasing low entropy keys, which may be easily guessed, several other attacks to biometric cryptosystems have been proposed (especially against the fuzzy vault scheme). Therefore, the claimed security of these technologies remains unclear and further improvement to prevent from these attacks is necessary. While some key approaches have already been exposed to fail the security demands more sophisticated security studies for all approaches are required. Due to the sensitivity of biometric key-binding and key-generation systems, sensoring and preprocessing may require improvement, too.

As plenty different approaches to biometric cryptosystems have been proposed a large number of pseudonyms and acronyms have been dispersed across literature such that attempts to represented biometric template protection schemes in unified architectures have

been made (Breebaart et al., 2008). In addition a standardization on biometric template protection is currently under work in the ISO/IEC FCD 24745 (Breebaart et al., 2009).

## 7. Summary and conclusion

Iris recognition has been established as a reliable means of performing access control in various types of applications. Existing algorithms (see Bowyer et al. (2007)) have been well-tested on public datasets meeting the requirements of handling large-scale databases (even in identification mode). However, iris recognition systems still require further improvement with respect to biometric template protection. Biometric templates can be lost, stolen, duplicated, or compromised enabling potential impostors to intrude user accounts and, furthermore, track and observe user activities. Biometric cryptosystems (Uludag et al., 2004), which represent a rather recent field of research offer solutions to biometric template protection as well as biometric-dependent key-release. Within approaches to biometric cryptosystems cryptographic keys are associated with fuzzy biometric data where authentication is performed in a secure manner, indirectly via key validities.

The iris, the sphincter around the pupil of a person's eye, has been found to be the most suitable biometric characteristic to be applied in biometric cryptosystems. In this chapter a comprehensive overview of the state-of-the-art in iris biometric cryptosystems is given. After discussing the fundamentals of iris recognition and biometric cryptosystems existing key concepts are reviewed and implementations of different variations of iris-based fuzzy commitment schemes (Juels & Wattenberg, 1999) are presented. Based on the obtained results, which underline the potential of iris biometric cryptosystems, a concluding discussion is given, including advantages and applications of biometric cryptosystems as well as open issues and challenges.

## 8. Acknowledgments

## 9. References

Boult, T., Scheirer, W. & Woodworth, R. (2007). Revocable fingerprint biotokens: Accuracy and security analysis, *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on* **0**: 1–8.

Bowyer, K. W., Hollingsworth, K. & Flynn, P. J. (2007). Image understanding for iris biometrics: A survey, *Computer Vision and Image Understanding* **110**(2): 281 – 307.

Braun, D. (2003). How they found national geographic's "afgahn girl", *National Geographic* **March 7**.

Breebaart, J., , Yang, B., Buhan-Dulman, I. & Busch, C. (2009). Biometric template protection - the need for open standards, *Datenschutz und Datensicherheit - DuD* **33**: 299–304.

Breebaart, J., Busch, C., Grave, J. & Kindt, E. (2008). A reference architecture for biometric template protection based on pseudo identities, *Proc. of the BIOSIG 2008: Biometrics and Electronic Signatures*, pp. 25–38.

Bringer, J., Chabanne, H., Cohen, G., Kindarji, B. & Zémor, G. (2007). Optimal iris fuzzy sketches, *in Proc. 1st IEEE International Conference on Biometrics: Theory, Applications, and Systems.* pp. 1–6.

Bringer, J., Chabanne, H., Cohen, G., Kindarji, B. & Zémor, G. (2008). Theoretical and practical boundaries of binary secure sketches, *IEEE Transactions on Information Forensics and Security* **3**: 673–683.

Cavoukian, A. & Stoianov, A. (2009a). Biometric encryption, *Encyclopedia of Biometrics*, Springer Verlag.

Cavoukian, A. & Stoianov, A. (2009b). Biometric encryption: The new breed of untraceable biometrics, *Biometrics: fundamentals, theory, and systems*, Wiley.

Cimato, S., Gamassi, M., Piuri, V., Sassi, R. & Scotti, F. (2009). Privacy in biometrics, *Biometrics: fundamentals, theory, and systems*, Wiley.

Clancy, T. C., Kiyavash, N. & Lin, D. J. (2003). Secure smartcard-based fingerprint authentication, *Proc. ACM SIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop* pp. 45–52.

Daugman, J. (2003). The importance of being random: statistical principles of iris recognition, *Pattern Recognition* **36**(2): 279 – 291.

Daugman, J. (2004). How iris recognition works, *IEEE Transactions on Circiuts and Systems for Video Technology* **14**(1): 21–30.

Daugman, J. (2011). How the afghan girl was identified by her iris patterns. http://www.cl.cam.ac.uk/ jgd1000/afghan.html, Retrieved 2011-01-03.

Davida, G., Frankel, Y. & Matt, B. (1998). On enabling secure applications through off-line biometric identification, *Proc. of IEEE, Symp. on Security and Privacy* pp. 148–157.

Davida, G., Frankel, Y. & Matt, B. (1999). On the relation of error correction and cryptography to an off line biometric based identication scheme, *Proc. of WCC99, Workshop on Coding and Cryptography* pp. 129–138.

Dodis, Y., Ostrovsky, R., Reyzin, L. & Smith, A. (2004). Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data, *Proc. Eurocrypt 2004 (LNCS: 3027)* pp. 523–540.

Feng, H. & Wah, C. C. (2002). Private key generation from on-line handwritten signatures, *Information Management and Computer Security* **10**(18): 159–164.

Hämmerle-Uhl, J., Pschernig, E., & A.Uhl (2009). Cancelable iris biometrics using block re-mapping and image warping, *In Proceedings of the Information Security Conference 2009 (ISC'09) LNCS: 5735* pp. 135–142.

Hao, F., Anderson, R. & Daugman, J. (2006). Combining Cryptography with Biometrics Effectively, *IEEE Transactions on Computers* **55**(9): 1081–1088.

Jain, A. K., Flynn, P. J. & Ross, A. A. (2008). *Handbook of Biometrics*, Springer-Verlag.

Jain, A. K., Nandakumar, K. & Nagar, A. (2008). Biometric template security, *EURASIP J. Adv. Signal Process* **2008**: 1–17.

Jain, A. K., Ross, A. & Pankanti, S. (2006). Biometrics: a tool for information security, *IEEE Transactions on Information Forensics and Security* **1**: 125–143.

Jain, A. K., Ross, A. & Prabhakar, S. (2004). An introduction to biometric recognition, *IEEE Trans. on Circuits and Systems for Video Technology* **14**: 4–20.

Jain, A. K., Ross, A. & Uludag, U. (2005). Biometric template security: Challenges and solutions, *in Proceedings of European Signal Processing Conference (EUSIPCO)* .

Juels, A. & Sudan, M. (2002). A fuzzy vault scheme, *Proc. 2002 IEEE International Symp. on Information Theory* p. 408.

Juels, A. & Wattenberg, M. (1999). A fuzzy commitment scheme, *Sixth ACM Conference on Computer and Communications Security* pp. 28–36.

Lee, C., Choi, J., Toh, K., Lee, S. & Kim, J. (2007). Alignment-free cancelable fingerprint templates based on local minutiae information, *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics* **37**(4): 980–992.

Lee, Y. J., Bae, K., Lee, S. J., Park, K. R. & Kim, J. (2007). Biometric key binding: Fuzzy vault based on iris images, *in Proceedings of Second International Conference on Biometrics* pp. 800–808.

Ma, L., Tan, T., Wang, Y. & Zhang, D. (2004). Efficient iris recognition by characterizing key local variations, *IEEE Transactions on Image Processing* **13**(6): 739–750.

Maltoni, D., Maio, D., Jain, A. K. & Prabhakar, S. (2009). *Handbook of Fingerprint Recognition*, 2nd edn, Springer Publishing Company, Incorporated.

Monrose, F., Reiter, M. K., Li, Q. & Wetzel, S. (2001). Using Voice to Generate Cryptographic Keys, *Proc. 2001: A Speaker Odyssey, The Speech Recognition Workshop* . 6 pages.

Monrose, F., Reiter, M. K. & Wetzel, S. (1999). Password hardening based on keystroke dynamics, *Proceedings of sixth ACM Conference on Computer and Communications Security, CCCS* pp. 73–82.

Nandakumar, K. & Jain, A. K. (2008). Multibiometric template security using fuzzy vault, *IEEE 2nd International Conference on Biometrics: Theory, Applications, and Systems, BTAS '08*, pp. 1–6.

Nandakumar, K., Jain, A. K. & Pankanti, S. (2007). Fingerprint-based Fuzzy Vault: Implementation and Performance, *in IEEE Transactions on Information Forensics And Security* **2**: 744–757.

Ratha, N. K., Connell, J. H. & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems, *IBM Systems Journal* **40**: 614–634.

Rathgeb, C. & Uhl, A. (2009a). Context-based texture analysis for secure revocable iris-biometric key generation, *Proc. of the 3rd International Conference on Imaging for Crime Detection and Prevention, ICDP '09*.

Rathgeb, C. & Uhl, A. (2009b). Systematic construction of iris-based fuzzy commitment schemes, *in* M. Tistarelli & M. Nixon (eds), *Proc. of the 3rd International Conference on Biometrics 2009 (ICB'09)*, Vol. 5558 of *LNCS*, Springer Verlag, pp. 940–949.

Rathgeb, C. & Uhl, A. (2010a). Adaptive fuzzy commitment scheme based on iris-code error analysis (second best student paper award), *Proceedings of the 2nd European Workshop on Visual Information Processing (EUVIP'10)*, pp. 41–44.

Rathgeb, C. & Uhl, A. (2010b). Two-factor authentication or how to potentially counterfeit experimental results in biometric systems, *Proc. of the International Conference on Image Analysis and Recognition (ICIAR'10)*, Vol. 6112 of *Springer LNCS*, pp. 296–305.

Rathgeb, C., Uhl, A. & Wild, P. (2010). Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity, *Proceedings of the 4th IEEE International Conference on Biometrics: Theory, Application, and Systems 2010 (IEEE BTAS'10)*, IEEE Press, pp. 1–6.

Reddy, E. & Babu, I. (2008). Performance of Iris Based Hard Fuzzy Vault, *IJCSNS International Journal of Computer Science and Network Security* **8**(1): 297–304.

Teoh, A. B. J., Ngo, D. C. L. & Goh, A. (2004). Personalised cryptographic key generation based on FaceHashing, *Computers And Security* **2004**(23): 606–614.

Uludag, U., Pankanti, S., Prabhakar, S. & Jain, A. K. (2004). Biometric cryptosystems: issues and challenges, *Proceedings of the IEEE* **92**(6): 948–960.

Vielhauer, C., Steinmetz, R. & Mayerhöfer, A. (2002). Biometric hash based on statistical features of online signatures, *ICPR '02: Proceedings of the 16 th International Conference*

*on Pattern Recognition (ICPR'02) Volume 1*, IEEE Computer Society, Washington, DC, USA, p. 10123.

Wu, X., Qi, N., Wang, K. & Zhang, D. (2008a).   A Novel Cryptosystem based on Iris Key Generation, *Fourth International Conference on Natural Computation (ICNC'08)* pp. 53–56.

Wu, X., Qi, N., Wang, K. & Zhang, D. (2008b). An iris cryptosystem for information security, *IIH-MSP '08: Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, IEEE Computer Society, Washington, DC, USA, pp. 1533–1536.

Yang, S. & Verbauwhede, I. (2007).   Secure Iris Verification, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP 2007)*, Vol. 2, pp. II–133–II–136.

Zhang, L., Sun, Z., Tan, T. & Hu, S. (2009).   Robust biometric key extraction based on iris cryptosystem, *In Proceedings of the 3rd International Conference on Biometrics 2009 (ICB'09) LNCS: 5558* pp. 1060–1070.

Zuo, J., Ratha, N. K. & Connel, J. H. (2008). Cancelable iris biometric, *In Proceedings of the 19th International Conference on Pattern Recognition 2008 (ICPR'08)* pp. 1–4.

# Iris Pattern Classification Combining Orientation Recognition

Hironobu Takano and Kiyomi Nakamura
*Toyama Prefectural University*
*Japan*

## 1. Introduction

The importance of personal authentication is gradually increasing with the development of the information society. Biometrics identification technology plays an important role in cyberspace. Unlike other biometrics such as the face or fingerprints, iris recognition has high reliability for personal identification. Iris recognition methods are classified into four categories: the phase-based method (Daugman, 1993), the zero-crossing representation-based method (Boles and Boashash, 1998; Sanchez-Avila and Sanchez-Reillo, 2005), the texture-based method (Wildes, 1997; Ma et al., 2003), and local intensity variation (Ma et al., 2004, a;b). Using the internal CASIA dataset (CBSR, 2005), Ma et al. evaluates the proposed algorithm by comparing the performance of other iris recognition methods proposed by Daugman, Wildes, and Boles and Boashash (Ma et al., 2004, b). The experimental results show the equal error rates (EER) of respective algorithms (Ma, Daugman, Wildes, and Boles) are 0.07%, 0.08%, 1.76%, and 8.13%, respectively. In other studies, the Daugman's method which is a representative algorithm of iris recognition is also evaluated using the subset of internal CASIA dataset (Sun et al., 2006) and the CASIA iris image 1.0 database (Wang et al., 2007), which is available from the CASIA web site. The EERs of Daugman's method reported in Sun et al. and Wang et al. are 0.70% and 0.67%, respectively. These analysis indicate the high accuracy of recognition performance although the EERs of Daugman's method described in these papers are not the same because the iris segmentation method including eyelid and eyelash detection would not be exactly the same.

Iris recognition technology are applied in various fields. Especially, the iris recognition algorithm embedded on a mobile phone requires robustness to rotation changes because capturing the iris pattern by a hand using a camera built in the mobile phone causes the rotation changes. However, the iris recognition methods described above are generally fragile in rotation variation.

We previously proposed a rotation spreading neural network (R-SAN net) that focused on spatial recognition/memory systems (parietal cortex(PG)) in the brain and recognized an object's orientation and shape (Nakamura et al., 1998; Yoshikawa and Nakamura, 2000). The R-SAN net can simultaneously recognize the orientation of the object irrespective of its shape, and the shape irrespective of its orientation. The characteristics of the R-SAN net are to use a two-dimensional input pattern in a polar coordinate system converted from the Cartesian coordinate system. The R-SAN net is suitable for the shape and orientation recognition of concentric circular patterns. The orientation recognition performance of R-SAN net allows the

accurate compensation of the orientation variation. In addition, the R-SAN net has the unique characteristics of orientation recognition. The recognized orientation for unlearned irises was heavily dispersed from the orientation of input iris although the orientation for learned irises was concentrated around input orientation. By combining the orientation recognition characteristics, a novel iris recognition method was developed.

On the other hand, despite the high recognition accuracy, the iris authentication system is vulnerable to deception by fake irises (Matsumoto et al., 2004). Thus, the iris recognition system requires liveness detection for discriminating between live and fake irises. For discriminating between live and fake irises, many liveness detection methods have been proposed earlier, for example, the eye gaze detection method, pupillary reflex method, etc. (Tachibana, 2006; Tsukahara, 2006; Oda, 2000; Kobayashi et al., 2005). The eye gaze detection method constrains the user to move their eyes along with the movement of a marker displayed on a screen. In the liveness detection method using corneal reflection of near-infrared light, an imposter can imitate an iris by painting an artificial corneal reflection image on the iris image. The pupillary reflex method uses the variations in the pupil size with time as a result of flashlight illumination. However, in these methods, biometric data for the identification of an individual and liveness data for classifying live and fake irises are obtained by measuring different physical features. To increase the reliability of liveness detection, we developed a novel liveness detection method using the brightness variation of the iris pattern based on pupillary reflex (Kanematsu et al., 2007).

In this chapter, we introduce the recognition method using the characteristics of orientation recognition for decreasing false acceptance. We also show a novel liveness detection for discriminating the live and fake irises. Section 2 describes the structure of the rotation spreading neural network (R-SAN net). The outline of the real-time iris recognition system using R-SAN net is shown in Section 3. Recognition performance of the iris pattern and orientation are evaluated in Section 4. Section 5 details the iris recognition method which introduces the unlearned iris rejection with the orientation recognition characteristics. The liveness detection method is described in Section 6. Section 7 concludes this chapter.

## 2. Rotation spreading neural network

### 2.1 Structure of the R-SAN net

The structure of the R-SAN net is shown in Fig.1. The R-SAN net consists of orientation and shape recognition systems. In the operation of this net, the input pattern ($300 \times 300$ pixels) is converted to a transformed pattern on the polar coordinates. This transformed pattern is input into the spreading layer, and the spread pattern $V$ is obtained.

In learning, the spread pattern $V_L^{(P)}$ is obtained by using the *P-th* learning input pattern in the spreading layers. The orientation memory matrix $\mathcal{M}_O$ is obtained by associating $V_L^{(P)}$ with the desired outputs of orientation recognition neurons $TO^{(P)}$. The $\mathcal{M}_O$ and $V_L^{(P)}$ are stored in the iris recognition system. In recognition, the output of orientation recognition neurons $YO$ is obtained by multiplying the spread pattern $V_R^{(P)}$ by orientation memory matrix $\mathcal{M}_O$. The orientation is recognized from the output of orientation recognition neurons using the population vector method (Georgopoulus et al., 1982). The shape (iris pattern) is discriminated with the Euclidean distance between the spread patterns obtained in learning and recognition processes.
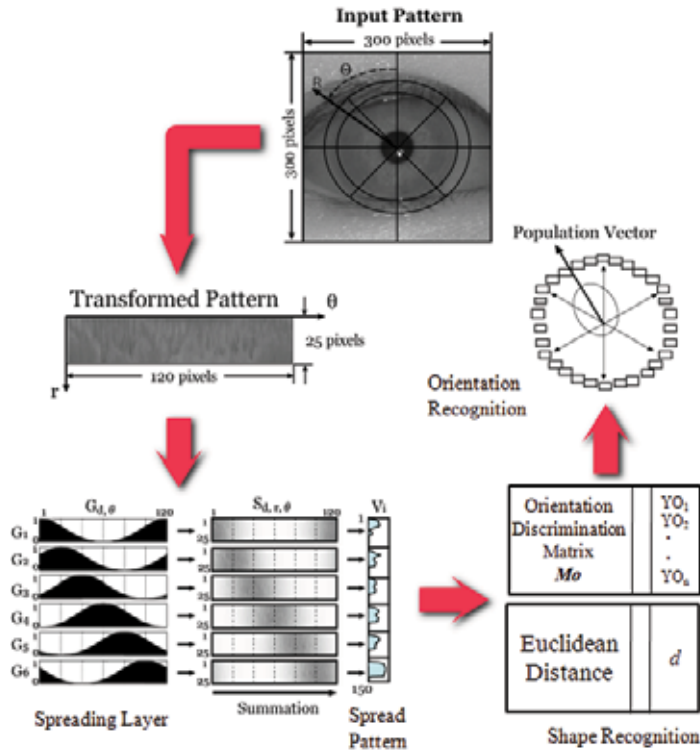
Fig. 1. Structure of the R-SAN net.

## 2.2 Generation of a transformed image

The original image for learning and recognition is a gray scale image of $300 \times 300$ pixels. The transformed image is made by sampling the original image on the polar coordinates $(r, \theta)$ at every 3 degrees in $\theta$ and at equal intervals of 25 pixels in radius $r$ excluding the pupil area. In order to get an accurate value of the transformed $T_{r,\theta}$, the small sampling region is further divided into $3 \times 3$ points and the pixel value of each point is summed. This transformed image is generated by Eq.(1), where $I_{x,y}$ is the pixel value of the original image at $(x, y)$ on the Cartesian coordinates, and $T_{r,\theta}$ is the pixel value of the transformed image at $(r, \theta)$ on the polar coordinates. Example of the original and transformed images is shown in Fig.2 (a) and (b).

$$T_{r,\theta} = \sum_{i=1}^{3} \sum_{j=1}^{3} I_{x,y} \tag{1}$$

$$(x = R \cos \Theta, \; y = R \sin \Theta)$$

$$\left( R = (r-1) + \frac{i}{3}, \; \Theta = \left\{ (\theta - 1) + \frac{j}{3} \right\} \times 3 \right)$$

## 2.3 Spreading layers

The structure of the spreading layers is shown in Fig.3. As shown in Eq.(4), the spread image $S_{d,r,\theta}$ corresponding to the respective spreading weight is obtained by multiplication of the
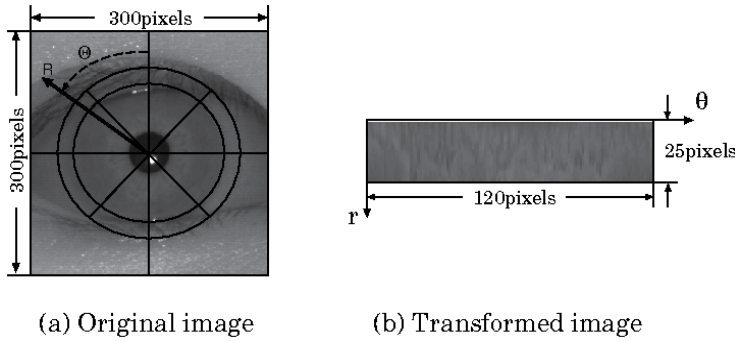
(a) Original image                (b) Transformed image

Fig. 2. Example of original and transformed images.

transformed image $T_{r,\theta}$ with the spreading weight $G_{d,\theta}$, which is the periodic Gaussian curve function predetermined at equal intervals in the $\theta$ direction (Eqs.(2) and (3)). The spread image is summed in the $\theta$ direction and combined to produce the spread pattern vector $V^*$ (Eqs.(5) and (6)).

$$F_S(x) = \exp\left\{-\beta(x - 120n)^2\right\} \tag{2}$$

$$(-60 + 120n < x \leq 60 + 120n, n = 0, \pm 1, \ldots)$$

$$G_{d,\theta} = F_S\left\{20(d-1) - (\theta - 1)\right\} \tag{3}$$

$$(d = 1, 2, \cdots, 6, \ \theta = 1, 2, \cdots, 120)$$

$$S_{d,r,\theta} = T_{r,\theta} \times G_{d,\theta} \tag{4}$$

$$(r = 1, 2, \cdots, 25)$$

$$V_i^* = \sum_{\theta=1}^{120} S_{d,r,\theta} \qquad (i = 25(d-1) + r) \tag{5}$$
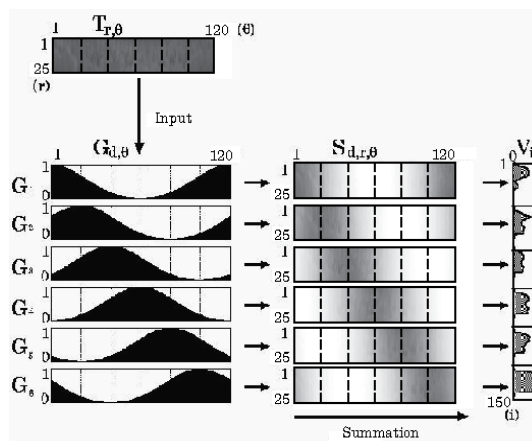
$$V^* = [V_1^*, \cdots, V_{150}^*]^T \tag{6}$$



Fig. 3. Structure of the spreading layers.

To remove the bias of $V^*$ which degrades the recognition performance, the normalized spread pattern vector $V$ is obtained by Eqs.(7) and (8). As a feature vector of the iris pattern, the normalized spread pattern $V$ is used for both learning and recognition.

$$||V^*|| = \sqrt{\sum_{i=1}^{150} V_i^{*2}} \tag{7}$$

$$V = \frac{V^*}{||V^*||} \tag{8}$$

### 2.4 Teaching signal

The teaching signal for orientation recognition is shown in Fig.4. Orientation recognition neurons $YO_i$ ($i = 1, 2, \cdots, 30$) are arranged at equal intervals of orientation. There are six learning signals $KO^{(d)}$ corresponding to the six orientations $d$ to be memorized. The desired outputs of the orientation recognition neurons are broadly tuned to the orientation of an iris pattern and adjusted to the function in Eq.(9). The desired outputs of orientation recognition neurons $TO^{(P)}$ in Eq.(11) are fitted to $KO^{(d)}$ which is the Gaussian curve function defined by Eqs.(9) and (10). Here, $P$ is the learning pattern number and $d$ is learning orientations ($O1 \sim O6$). $\alpha$ is the learning coefficient that defines the tuning width of the teaching signal of orientation recognition neurons. The learning coefficient, $\alpha$, is determined so that an orientation recognition neuron corresponding to the learning orientation outputs a peak value 1.0, and the orientation recognition neurons corresponding to the nearest neighbor learning directions output 0.5. The orientation ($O1 \sim O6$) of iris images rotated every 60 degrees are learned.
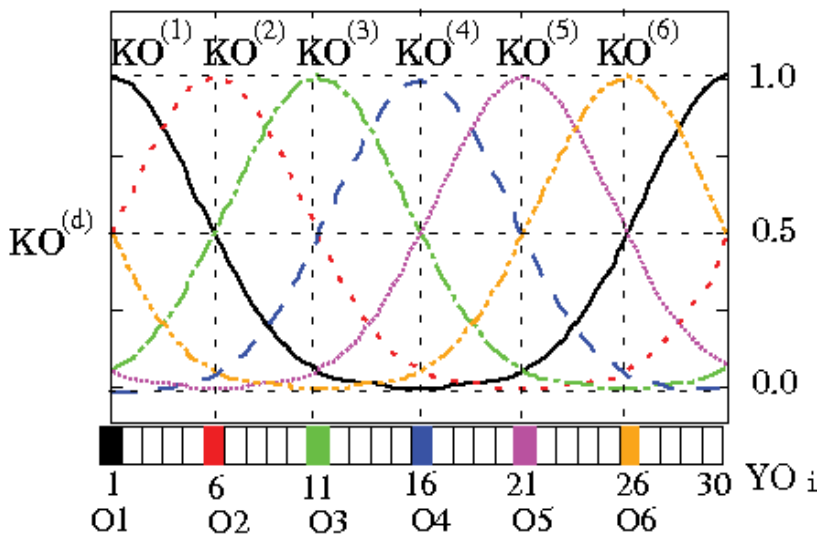


Fig. 4. Teaching signal for orientation recognition.

$$F_O(x) = \exp\left\{-\alpha(x - 30n)^2\right\} \tag{9}$$

$$(-15 + 30n < x \le 15 + 30n, n = 0, \pm 1, \ldots)$$

$$YO_i^{(P)} = KO_i^{(d)} = F_O\left\{5(d-1) - (i-1)\right\} \tag{10}$$

$$(d = 1, 2, \cdots, 6 \,, i = 1, 2, \cdots, 30)$$

$$\boldsymbol{TO}^{(P)} = \boldsymbol{KO}^{(d)} = [KO_1^{(d)}, KO_2^{(d)}, \cdots, KO_{30}^{(d)}]^T \tag{11}$$

### 2.5 Population vector method

The orientation of the iris pattern is indicated by the angle of a population vector $\phi$. The $\phi$ is defined as an ensemble of vectors of the orientation recognition neurons $\boldsymbol{YO} = [YO_1, \cdots, YO_{30}]^T$ where each vector points to the neuron's optimally tuned orientation and has a length in proportion to the neuron's output (Georgopoulus et al., 1982). The arrangement of orientation neurons and the orientation population vector are shown in Fig.5. This assumes that the neurons in the parietal cortex recognize the axis orientation of an object by population coding, as seen in neurophysiological studies. Each orientation recognition neuron $YO_i$ has a respective representative orientation $\psi_i$ that characterizes the best orientation for the optimal response in Eq.(12). The population vector orientation $\phi$ is calculated by the vectorial summation of 30 orientation neurons ($YO_1, \cdots, YO_{30}$) by Eqs.(13) and (14).

$$\psi_i = \frac{2\pi}{30} \times (i - 1) \quad [\text{rad}] \tag{12}$$

$$(i = 1, 2, \cdots, 30)$$

$$x = \sum_{i=1}^{30} YO_i \cos \psi_i$$
$$y = \sum_{i=1}^{30} YO_i \sin \psi_i \tag{13}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{14}$$

### 2.6 Learning process

The R-SAN net uses generalized inverse learning for orientation recognition (Amari, 1978). The spread pattern $V_{\text{L}}^{(P)}$ is obtained from the *P-th* learning input pattern in the spreading layers. The orientation memory matrix $\boldsymbol{\mathcal{M}}_{\text{O}}$ is obtained by associating $V_{\text{L}}^{(P)}$ with the desired outputs of orientation recognition neurons $\boldsymbol{TO}^{(P)}$ by Eq.(17). The number of learning patterns is given by multiplying the number of learning irises by the number of learning orientations for each iris. For example, when the number of learning irises is 10 and the number of learning orientations is 6, it is 10 (irises) $\times$ 6 (orientations) = 60 (patterns). For iris pattern learning, the
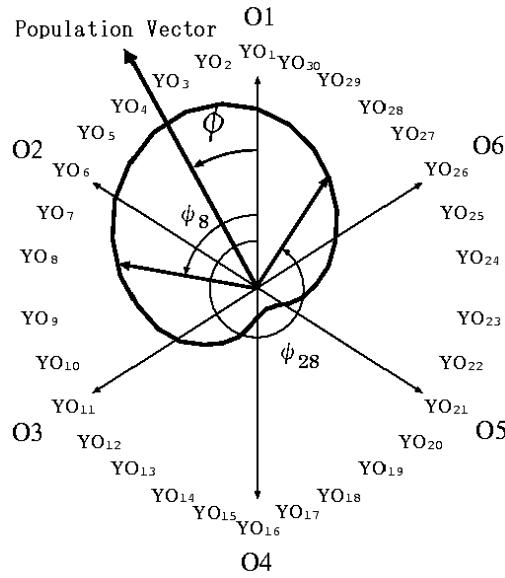
Fig. 5. Arrangement of orientation recognition neurons in population vector method.

spread patterns $V_{\mathrm{L}}^{(P)}$ for the respective irises are registered in the iris recognition system.

$$\mathcal{X} = [V_{\mathrm{L}}^{(1)}, V_{\mathrm{L}}^{(2)}, \cdots, V_{\mathrm{L}}^{(60)}] \tag{15}$$

$$\mathcal{X}^{\dagger} = (\mathcal{X}^T \mathcal{X})^{-1} \mathcal{X}^T \tag{16}$$

$$\mathcal{M}_{\mathrm{O}} = \mathcal{TO}\mathcal{X}^{\dagger} \tag{17}$$

$$\mathcal{TO} = [TO^{(1)}, TO^{(2)}, \cdots, TO^{(60)}]$$

### 2.7 Recognition process

In recognition, the spread iris pattern $V_{\mathrm{R}}$ used in recognition is generated from an input iris image. For orientation recognition, the output of orientation recognition neurons $YO = [YO_1, \cdots, YO_{30}]^T$ is obtained by multiplying the spread pattern $V_{\mathrm{R}}$ by orientation memory matrix $\mathcal{M}_{\mathrm{O}}$ in Eq.(18). The orientation of the input iris pattern is recognized from the output of orientation recognition neurons using the population vector method. This method provides the orientation of the iris pattern by synthesizing the continuous spectra of the outputs of the orientation recognition neurons.

$$YO = \mathcal{M}_{\mathrm{O}} V_{\mathrm{R}} \tag{18}$$

The shape (iris pattern) is discriminated with the Euclidean distance between the spread patterns obtained in learning and recognition processes. The value of Euclidean distance $d$ in Eq.(19) has the range of $0 \le d \le 2$, because the norm of spread pattern $||V||$ are normalized as "1". In Eq.(19), $V_{\mathrm{L}}$ and $V_{\mathrm{R}}$ correspond to the normalized spread pattern of $0°$ during learning and during recognition, respectively. When it has the Euclidean value of "0", resemblance is the highest.

$$d = ||V_{\mathrm{L}} - V_{\mathrm{R}}|| \tag{19}$$

## 3. Real-time iris recognition system

### 3.1 System configuration

The configuration of the real-time iris recognition system is shown in Fig.6. The system consists of a near-infrared CCD camera, PC, near-infrared lighting, and flash-light generation equipment. The PC has an image input board to acquire the iris images and DLL (Dynamic Link Library) software for processing iris images. The flash-light generation equipment emits a flash using an external trigger signal synchronized with the arbitrary input image frame. The near-infrared lighting provides clear iris images for the CCD camera. A fixed-size pupil image can be obtained by utilizing the pupillary reflex caused by illuminating the same eye. The pupillary reflex is also used for liveness detection. The input iris images captured by the CCD camera are gray-scaled images of $640 \times 480$ pixels every 1/30 sec.
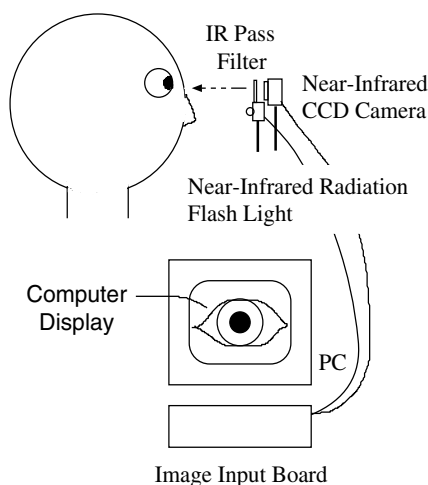


Fig. 6. Configuration of iris recognition system.

### 3.2 Outline of the iris recognition system

The flowchart of the iris recognition system is shown in Fig.7. First, the template matching method with a partial eye template detects the pupil position from the eye image continuously taken by the near-infrared CCD camera (Miyazaki et al., 2007). After detection of the pupil location, variations of pupil size (diameter) due to pupillary reflex are measured by calculating the average distance between the center of pupil compensated by the labeling and least squared fitting method and pixels on the circumference. The iris size (diameter) is measured by edge detection using the Prewitt filter. Although the measurement sizes of the iris and pupil in the image change with the magnification of a lens and the distance between the camera and eye, the iris and pupil sizes (diameter) can be accurately measured using the characteristics of the fixed (almost equal) iris size without individual differences. The variation of pupil size (diameter) during pupillary reflex induced by using a weak LED-flashlight is shown in Fig.8. When the pupil diameter normalized by the measured iris diameter is between 2.9 mm and 3 mm as shown in Fig.8, the eye image is taken as a standard image of the iris and pupil. A $300 \times 300$ pixel image including the iris pattern is extracted from this eye image, which is zoomed using the linear interpolation. Thus, the pupil size
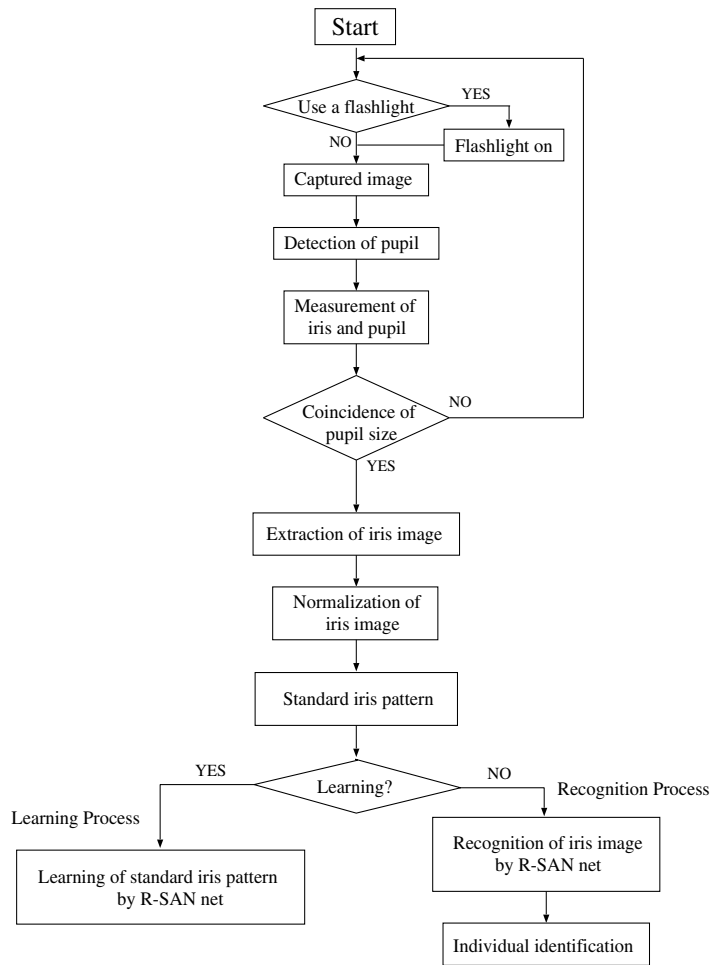
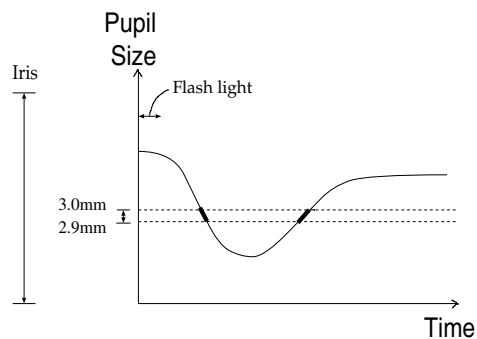Fig. 7. Flowchart of the iris recognition system.



Fig. 8. Variation of pupil diameter caused by pupillary reflex.

between 2.9 mm and 3 mm on the normalized image becomes 50 pixels on the PC screen. The iris size on the screen is also zoomed in the same manner as pupil size normalization. The inset image at the upper left of Fig.9 is a normalized iris image. This normalized image is used as the standard iris pattern which is the input image of the R-SAN net. In the R-SAN net, the standard iris pattern is converted to the transformed pattern on the polar coordinates. In the spreading layer of the R-SAN net, the spread pattern $V$ is obtained by multiplying the transformed pattern by the spreading weight. In learning process, the spread pattern $V_L$ and orientation memory matrix $\mathcal{M}_O$ are stored in the iris recognition system. In recognition process, the orientation angle is obtained by the population vector method using the outputs of orientation recognition neurons. The iris pattern is recognized with the Euclidean distance between the spread patterns obtained in learning and recognition processes.
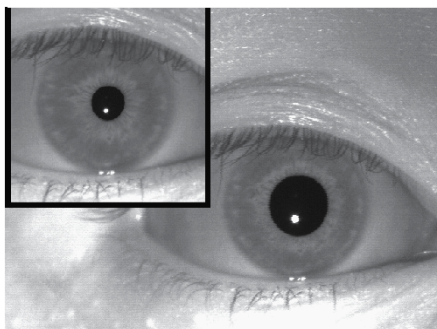


Fig. 9. Size normalization of an iris image.

## 4. Iris recognition experiment

The characteristics of orientation and shape recognition for learned and unlearned irises were investigated with iris images captured under usual indoor lighting conditions (illuminance was 300 lx). The 38 iris images of 19 subjects (2 images for each subject) were used for recognition experiments. The iris images in the learning and recognition tests were at orientation $0°$. One iris image obtained from 19 subjects was used for training. The orientation recognition test was examined using 19 iris images (another iris image of the learned iris and other 18 unlearned irises). We tried 19 sets of recognition tests by changing the learning and recognition iris images one by one. Recognition results were thus obtained for 361 trials consisting of 19 trials for learned subjects and 342 for unlearned subjects. In shape (iris pattern) recognition test, 18 unlearned iris images among 19 subjects were recognized for each learned subject. Thus, 361 recognition trials consisting of 19 trials for learned irises and 342 trials for unlearned irises were examined. The iris pattern recognition was evaluated using the false rejection rate (FRR) and false acceptance rate (FAR). When the output of Euclidean distance for learned iris is higher than the decision threshold, we considered that the person was rejected and calculated the false rejection rate by counting the trials of false rejection. On the other hand, when the output of Euclidean distance calculated for unlearned iris was lower than the decision threshold, we considered that the imposters were accepted incorrectly. We calculated the false acceptance rate by counting the trials of false acceptance.

### 4.1 Orientation recognition performance

The orientation recognition result for learned and unlearned iris images was shown in Fig.10. The horizontal axis is the input iris number, and the vertical axis is the recognized orientation angle. The average ± standard deviation of recognized orientation for learned and unlearned irises were $0.82 \pm 2.77[°]$ and $2.01 \pm 62.87[°]$, respectively. As shown in Fig.10, the recognized orientation of learned irises distributed around 0 degree. However, the recognized orientation of unlearned irises was heavily dispersed (SD was very large). The histogram of recognized orientation angle for learned and unlearned irises was shown in Fig.11. The horizontal axis is the absolute error of recognized orientation angle, and the vertical axis is the percentage of iris image included in each bin. The black and white bars show the distribution of the absolute error of recognized orientation for learned and unlearned irises, respectively. The absolute error of recognized orientation for learned irises was less than 5 degrees. However, 87% absolute error of recognized orientation for unlearned irises distributed more than 10 degrees.
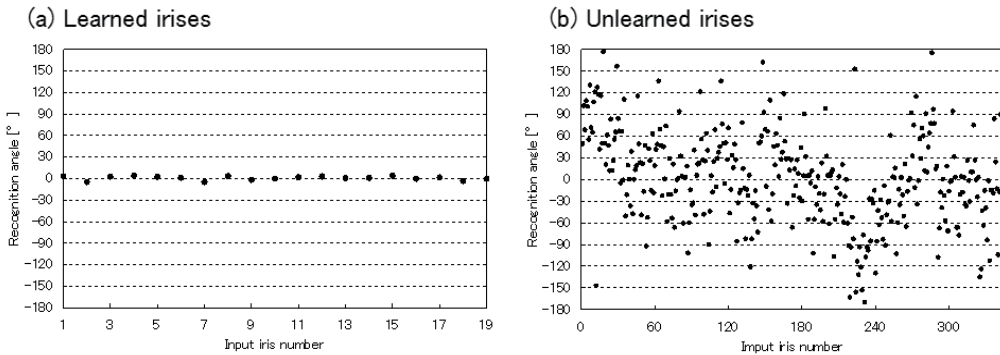


Fig. 10. Orientation recognition result for (a) learned and (b) unlearned irises.

### 4.2 Shape recognition performance

Shape recognition performance was evaluated using equal error rate (EER) determined by finding the point where false acceptance rate intersects the false rejection rate. The result of shape recognition is shown in Fig.12. The horizontal axis is the decision threshold for discriminating between registered persons and imposters. The vertical axis is the FRR and FAR. Circle and dashed line show the FRR. Square and solid line show the FAR. The equal error rate was 2.02% when the decision threshold of Euclidean distance was 0.33. At the FARs of 1% and 0.1%, the FRRs were 4.01% and 9.58%, respectively.

## 5. Unlearned iris rejection with recognized orientation

The orientation recognition performances indicated the R-SAN net had fairly good orientation recognition characteristics for learned irises. On the other hand, the orientation angle of unlearned irises was hardly recognized because the distribution of recognized orientation angle was widely dispersive. Thus, the R-SAN net can recognize the orientation of only learned irises. Using the difference of orientation recognition characteristics between learned and unlearned irises, the unlearned iris would be removed before iris discrimination with Euclidean distance. Before iris recognition using Euclidean distance calculated with spread
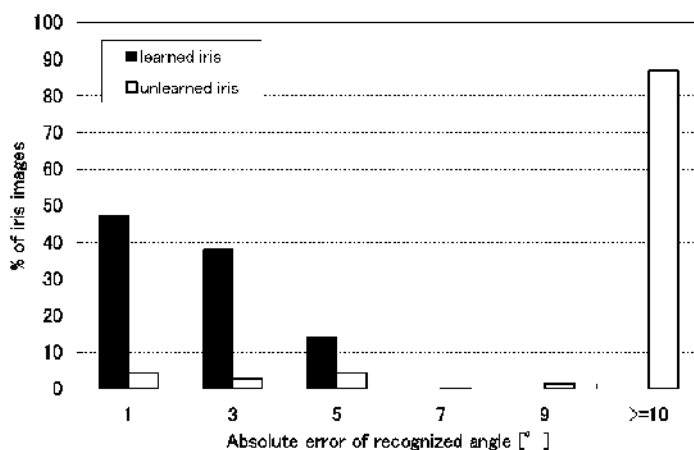
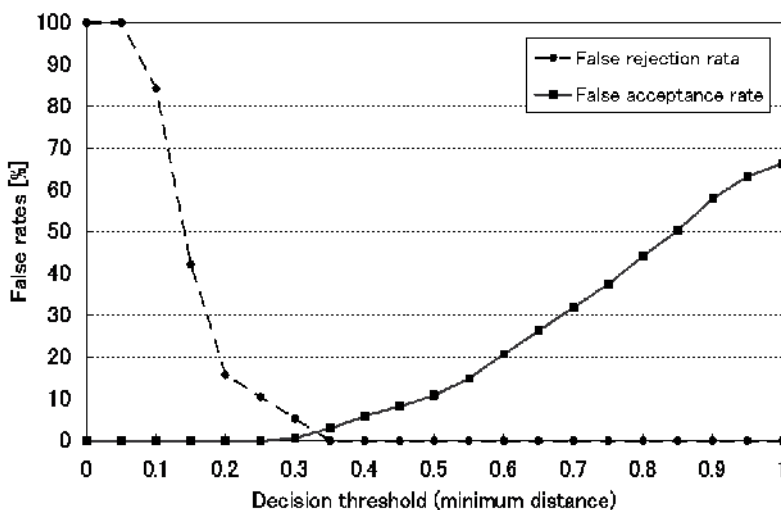Fig. 11. Histogram of absolute error of recognized orientations for learned and unlearned iris images.



Fig. 12. False acceptance and rejection rates by Euclidean distance.

patterns, the unregistered irises are rejected using the average and standard deviation of recognized orientation angle $(\theta_{av}, \sigma_o)$ for learned irises obtained by the R-SAN net. The input iris was determined as imposter if the recognized orientation is greater than $\theta_{av} + 2.1\sigma_o$ or less than $\theta_{av} - 2.1\sigma_o$. Note that all of learned irises are not rejected with the orientation discrimination because the recognized orientation for learned irises were within $\theta_{av} \pm 2.1\sigma_o$. The shape (iris pattern) recognition performance obtained by new recognition method was shown in Fig.13. The iris images used for learning and recognition are the same as Section 4. This result indicated the FAR drastically decreased. The equal error rate was 0.79% at the decision threshold of 0.34. At the FAR of 1% and 0.1%, the FRR was 0% and 6.95%,

respectively. The unregistered iris rejection by the recognized orientation is very effective to improve the shape recognition performance.
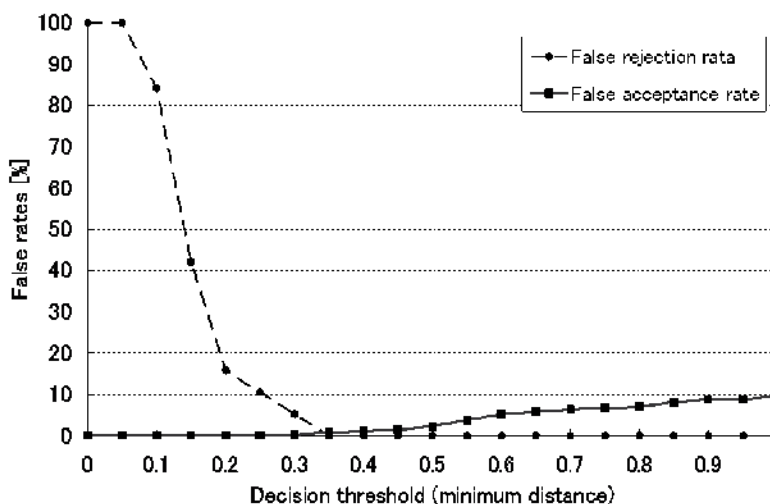


Fig. 13. False acceptance and rejection rates obtained by new iris discrimination method with the characteristics of orientation recognition.

## 6. Liveness detection using iris pattern

For discriminating between live and fake irises, a new liveness detection method that acquires both liveness and biometric data from the iris portion is introduced. In this method, a variation in the averaged pixel value (brightness) of the iris portion is used as the liveness data. The average brightness is calculated using the pixel values in a predetermined region of the iris portion. The variation in the brightness of the iris portion is caused by a pupillary reflex induced by a weak LED-flashlight. By measuring the variation in the average brightness of the iris portion, the fake iris would be rejected because the fake iris has a constant pattern. However, the brightness varies due to changes in the ambient lighting condition even if the iris pattern does not show a change. Thus, the brightness of the eye image captured by a camera is normalized to prevent a variation in the brightness of the iris image caused by ambient lighting variation (Takano et al., 2007).

In order to discriminate between live and fake irises, it was necessary to investigate the variation rates for live and fake irises. The experiment was performed with 16 images of fake irises, i.e. paper-printed iris images. The brightness variation rates of the live and fake irises are shown in Fig. 14. Trial numbers 1 to 80 shown as triangles represent the brightness variation for fake irises, while trial numbers 81 to 160 shown as circles represent the brightness variation for live irises. The variation rates of the averaged brightness obtained from live and fake irises were 10.6% and 1.5%, respectively. In addition, the maximum brightness variation rate for fake iris images was less than 2.5%. The large difference between the brightness variation rates of live and fake irises provides an anti-deception countermeasure. The decision threshold $L_{th}$ for classification of live and fake irises is obtained by using the brightness variation rates of live and fake irises in Eq.(20). $AV_l$ and $AV_f$ are the average brightness

variation rate of the live and fake irises, respectively. $SD_l$ and $SD_f$ are the standard deviation of brightness variation rate for the live and fake irises, respectively. The decision threshold of the brightness variation rate determined from the present experiment was 3.7%.

$$L_{th} = \frac{AV_l - AV_f}{SD_l + SD_f} SD_f + AV_f \qquad (20)$$
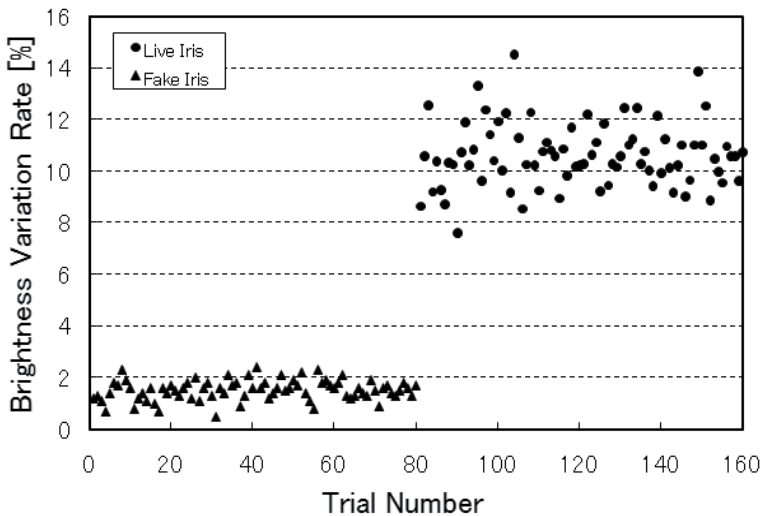


Fig. 14. The characteristics of brightness variation rates obtained from live and fake irises.

## 7. Conclusions

In this chapter, we showed the orientation and shape recognition performance of the R-SAN net for learned and unlearned irises. The orientation of learned irises can be correctly recognized. On the other hand, the orientation of unlearned irises cannot be recognized because the recognized orientation is heavily dispersed from the orientation of input iris. In the shape recognition with unlearned iris rejection, the equal error rate was 0.79% at decision threshold of 0.34.

The R-SAN net has the unique characteristics of the orientation recognition. The orientation of only learned iris were recognized correctly. However, the recognized orientation of unlearned irises were heavily scattered. By introducing the unlearned iris discrimination with the recognized orientation, new recognition method was developed. The experimental result of new recognition method showed that the false acceptance rate drastically decreased. The unregistered iris rejection method using recognized orientation provided the effective improvement of the iris recognition performance.

The highly reliable liveness detection method by using the average brightness variation of an iris portion based on the pupillary reflex was evaluated with live and fake irises. The averaged brightness variation rate of fake irises was extremely small compared with that of live irises. From the experimental results, the live and fake irises were discriminated with the decision criterion of 3.7% brightness variation rate.

In future work, we will test individual recognition with many more samples of iris patterns. We will also implement the R-SAN net as the security system of the mobile phone, and optimize the orientation and iris pattern recognition algorithms to reduce the computational cost.

## 8. References

Daugman, J. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 15, No. 11, pp. 1148-1161.

Boles, W. & Boashash, B. (1998). A human identification technique using images of the iris and wavelet transform. *IEEE Trans. Signal Process.*, Vol. 46, No. 4, pp. 1185-1188.

Sanchez-Avila, C. & Sanchez-Reillo, R. (2005). Two different approaches for iris recognition using Gabor filters and multiscale zero-crossig representation. *Pattern Recognition*, Vol. 38, No. 2, pp. 231-240.

Wildes, R. P. (1997). Iris recognition: an emerging biometric technology. *Proc. IEEE*, Vol. 85, No. 9, pp. 1348-1363.

Ma, L.; Tan, T.; Wang, Y. & Zhang, D. (2003). Personal identification based on iris texture analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 25, No. 12, pp. 1519-1533.

Ma, L.; Tan, T.; Wang, Y. & Zhang, D. (2004). Local intensity variation analysis for iris recognition. *Pattern Recognition*, Vol. 37, No. 6, pp. 1287-1298.

Ma, L.; Tan, T.; Wang, Y. & Zhang, D. (2004). Efficient iris recognition by characterizing key local variations. *IEEE Trans. Image Process.*, Vol. 13, No. 6, pp. 739-750.

Center for Biometrics and Security Research (2005). CASIA-IrisV4, 09.04.2011, Available from http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp

Sun, Z.; Tan, T. & Qiu, X. (2006). Graph matching iris image blocks with local binary pattern, In: *Advances in Biometrics*, LNCS 3832, pp. 366-372.

Wang, F.; Yao, X. & Han, J. (2007). Minimax probability machine multialgorithmic fusion for iris recognition. *Information Technology Journal*, Vol. 6, No. 7, pp. 1043-1049.

Nakamura, K.; Miyamoto, S. & Morisada, K. (1998). Characteristics of spatial spreading associative neural network in simultaneous recognition of object orientation and shape. *IEICE Trans. Inf. & Syst.*, Vol. J81-D-II, No. 6, pp. 1194-1204.

Yoshikawa, T. & Nakamura, K. (2000). Evaluation of recognition ability and inside parameters for spreading associative neural network. *IEICE Trans. Inf. & Syst.*, Vol. J83-D-II, No. 5, pp. 1332-1343.

Matsumoto, T.; Hirabayashi, M. & Sato, K. (2004). A vulnerability evaluation of iris matching (Part 3). *Proc. The 2004 Symposium on Cryptography and Information Security (SCIS 2004)*, pp. 701-706.

Tachibana, M. (2006). Injustice detection system for iris recognition. *Jpn. Kokai Tokkyo Koho*, No. JP2006-85226 A.

Tsukahara, S. (2006). Iris recognition device. *Jpn. Kokai Tokkyo Koho*, No. JP2006-136450 A.

Oda, T. (2006). Iris code generation system and iris recognition system. *Jpn. Kokai Tokkyo Koho*, No. JP2000-33080 A.

Kobayashi, H.; Takano, H. & Nakamura, K. (2005). Real-time iris recognition system not influenced by ambient light change using a rotation spreading neural network. *IEICE Technical Report*, No. NC2004-163, pp. 155-160.

Kanematsu, M.; Takano, H. & Nakamura, K. (2007). Highly reliable liveness detection method for iris recognition. *Proc. SICE 2007*, pp. 361-364.

Georgopoulus, A. P.; Kalaska, J. F.; Caminiti, R. & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci.*, Vol. 2, pp. 1527-1537.

Amari, S. (1978). *A Mathematical Principle of Neural Networks*, Sangyo Publishing, Tokyo.

Miyazaki, S.; Takano, H. & Nakamura, K. (2007). Suitable checkpoints of features surrounding the eye for eye tracking using template matching. *Proc. SICE 2007*, pp. 356-360.

Takano, H.; Kobayashi, H. & Nakamura, K. (2007). Rotation invariant iris recognition method adaptive to ambient lighting variation. *IEICE Trans. Inf. & Syst.*, Vol. E90-D, No. 6, pp. 955-962.

# Part 4

# Other Biometrics

# Gabor-Based RCM Features for Ear Recognition

Ali Pour Yazdanpanah[1] and Karim Faez[2]
*[1]Department of Electrical Engineering, Islamic Azad University, Najaf Abad Branch*
*[2]Department of Electrical Engineering, Amirkabir University of Technology*
*Iran*

## 1. Introduction

Ear biometrics has received deficient attention compared to the more popular techniques of face, eye, or fingerprint recognition. The ear as a biometric is no longer in its infancy and it has shown encouraging progress so far. ears have played an important role in forensic science for many years, especially in the United States, where an ear classification system based on manual measurements was developed by (Iannarelli, 1989). In recent years, biometrics recognition technology has been widely investigated and developed. Human ear, as a new biometric, not only extends existing biometrics, but also has its own characteristics which are different from others. Iannarelli has shown that human ear is one of the representative human biometrics with uniqueness and stability (Iannarelli, 1989). Since ear as a major feature for human identification was firstly measured in 1890 by Alphonse Bertillon, so-called ear prints have been used in the forensic science for a long time (Bertillon, 1890). Ears have certain advantages over the more established biometrics; as Bertillon pointed out, they have a rich and stable structure that does not suffer from the changes of ages, skin-color, cosmetics, and hairstyles. Also the ear does not suffer from changes in facial expression, and is firmly fixed in the middle of the side of the head so that the background is more predictable than is the case for face recognition which usually requires the face to be captured against a controlled background. The ear is large compared with the iris, retina, and fingerprint and therefore is more easily captured at a distance.

We presented gabor-based region covariance matrix as an efficient feature for ear recognition. In this method, we construct a region covariance matrix by using gabor features, illumination intensity component, and pixel location, and use it as an efficient and robust ear descriptor for recognizing peoples. The feasibility of the proposed method has been successfully tested on ear recognition using two USTB databases, specifically used total 488 ear images corresponding to 137 persons. The effectiveness of the proposed method is shown in terms of the comparative performance against some popular ear recognition methods.

This chapter is organized as follows. In section 2, related works are presented. In section 3, region covariance matrix (RCM) and the method for fast RCM computation are presented. In section 4, the proposed method presented in detail. In section 5, ear image databases are introduced. In section 6, experimental results are shown and commented. The chapter concludes in section 7.

## 2. Related works

Ear recognition depends heavily on the particular choice of features that used in ear biometric systems. The Principal Component Analysis method (PCA) is a classical statistical characteristic extracts method. The PCA (Xu, 1994; Abdi & Williams, 2010) transformation is based on second order statistics, which is commonly used in biometric systems. With second order methods, a description with minimum reconstruction error of the data is found using the information contained in the covariance matrix of the data. It is assumed that all the information of Gaussian variables (zero mean) is contained in the covariance matrix. The Independent Component Analysis (ICA) is another popular feature extraction method. ICA (Comon, 1994; Stone, 2005) provides a linear representation that minimizes the statistical dependencies among its components, which is based on higher order statistics of the data. These dependencies among higher order features could be eliminated by isolating independent components. It is a statistical method for transforming an observed multidimensional random vector into components that are statistically independent from each other as much as possible. The ability of the ICA to handle higher-order statistics in addition to the second order statistics is useful in achieving an effective separation of feature space for given data. The higher order features are capable of capturing invariant features of natural images. In (Zhang & Mu, 2008), PCA and ICA methods with RBFN classifier is presented. In these two methods, PCA and ICA are used to extract features and RBFN is used as classifier. In this chapter, these two methods denote by PCA+RBFN, and ICA+RBFN respectively.

Hmax+SVM is another popular feature extraction method for ear recognition. Hmax model is motivated by a quantitative model of visual cortex, and SVMs are classifiers which have demonstrated high generalization capabilities in many different tasks, including the object recognition problem. This method (Yaqubi et al., 2008) combines these two techniques for the robust Ear recognition problem. With Hmax, a new set of features has been introduced for human identification, each element of this set is a complex feature obtained by combining position- and scale- tolerant edge detectors over neighboring positions and multiple orientations. This system's architecture is motivated by a quantitative model of visual cortex (Riesenhuber & Poggio, 1999).

Another feature extraction method for ear recognition is presented by (Guo & Xu, 2008). This method called Local Similarity Binary Pattern (LSBP). Local Similarity Binary Pattern considers both the connectivity and similarity information in representation. LSBP histogram captures the information of connectivity and similarity, such as lines and connective area. In this method, in order to enhance efficient representation, histograms not only encode local information but also spatial information by image decomposition. Because of the special characteristics of ear images, the connectivity and similarity of intensity plays a significant role in ear recognition, which can be encoded by Local Similarity Binary Pattern.

## 3. RCM

### 3.1 Covariance matrix as a region descriptor
The covariance matrix is a symmetric matrix. Covariance matrix diagonal entries represent the variance of each feature and their non-diagonal entries represent their correlations.

Using covariance matrices as the descriptors of the region has many advantages. The covariance matrix presents a natural way of fusing multiple features without normalizing features or using blending weights. It embodies the information embedded within the histograms as well as the information that can be derived from the appearance models. In general, for each region, a single covariance matrix is enough to match with that region in different views and poses. The noise corrupting individual samples are mostly filtered out with the average filter during covariance computation process. Due to the equal size of the covariance matrix of any region, we can compare any two regions without being restricted to a constant window size. If the raw features such as, image gradients and orientations, are extracted according to the scale difference, It has also scale invariance property over the regions in different images.

As given above, covariance matrix can be invariant to rotations. However, if information regarding the orientation of the points are embedded within the feature vector, it is possible to detect rotational discrepancies. We also want to mention that the covariance is invariant to the mean changes such as identical shifting of color values. This can be an advantageous property when objects are tracked under different illumination conditions. Region covariance matrix (RCM) presented by (Tuzel et al., 2006). RCM is a covariance matrix of many image statistics computed within a region.

We define $I$ as an one dimensional unit normalized intensity image. The method can be generalized to other type of images, which can be a 2D intensity image, or 3D color image or multi spectral. Assume F be the $W \times H \times d$ dimensional feature image extracted from $I$

$$F(x,y) = \phi(I,x,y) \tag{1}$$

Where the function $\phi$ can be any mapping function such as color, image gradients $I_x, I_{xx}, \dots$, edge magnitude, edge orientation, filter responses, etc. this pixel-wise mapping list can be extended by including higher order derivatives, radial distances, texture scores, angels, and temporal frame differences in case a video data is available.

For a given rectangular window $R$, let $\{f_k\}_{k=1\dots n}$ be the d-dimensional feature vectors inside $R$.

Each feature vector $f_k$ introduces a pixel (x, y) within that window. Since we extract the mutual covariance of the features, the windows can actually be any shape not necessarily rectangles. Basically, covariance is a statistical measure of how much two variables vary together. Covariance can be a negative, positive or zero number, conditional upon what is the relation between two features (Forsyth & Ponce, 2002). If the features increase together, the covariance is positive. If one feature increases and the other decreases, the covariance is negative, and if the two features are independent, the covariance is zero. We introduce each window $R$ with a covariance matrix of the features.

$$
\begin{aligned}
C_R &= \begin{pmatrix} C_R(1,1) & \dots & C_R(1,d) \\ \vdots & \ddots & \\ C_R(d,1) & & C_R(d,d) \end{pmatrix} \\
&= \frac{1}{n-1}\sum_{k=1}^{n}(f_k - \mu)(f_k - \mu)^T
\end{aligned}
\tag{2}
$$

Where $\mu$ is the mean vector of the corresponding features for the points within the region $R$. The diagonal coefficients represent the variance of the corresponding features. For example, the $j_{th}$ diagonal element represents the variance for the $j_{th}$ feature. The non-diagonal elements represent the covariance between two different features.

The feature vectors can be constructed using different type of mapping functions like pixel coordinates, color intensity, gradient, etc.

$$f_k = \begin{bmatrix} x\ y\ I(x,y)\ I_x(x,y) \ldots \end{bmatrix} \tag{3}$$

or they can be constructed using the polar coordinates

$$f_k = \begin{bmatrix} r(x',y')\ \theta(x',y')\ I(x,y)\ I_x(x,y) \ldots \end{bmatrix} \tag{4}$$

where

$$(x',y') = (x - x_0, y - y_0) \tag{5}$$

are the relative coordinates with respect to window center $(x_0, y_0)$, and

$$r(x',y') = \sqrt{(x'^2 + y'^2)} \tag{6}$$

is the distance from $(x_0, y_0)$ and

$$\theta(x',y') = \arctan(\frac{y'}{x'}) \tag{7}$$

is the orientation component. For human detection problem, (Tuzel et al., 2007) introduced the mapping function as

$$f_k = \begin{bmatrix} x\ y\ |I_x|\ |I_y|\ \sqrt{I_x^2 + I_y^2}\ |I_{xx}|\ |I_{yy}|\ \theta(x,y) \end{bmatrix} \tag{8}$$

Where $|.|$ denotes the absolute operator. First- and second-order gradients and pixel location were used in this function to construct RCM. The other form of feature mapping function which is introduced by (Tuzel et al., 2006) for gray level images is

$$f_k = \begin{bmatrix} x\ y\ I(x,y)\ |I_x|\ |I_y|\ |I_{xx}|\ |I_{yy}| \end{bmatrix} \tag{9}$$

Three other kinds of feature mapping functions are introduced by (Tuzel et al., 2007; Pang et al., 2008).

$$f_k = \begin{bmatrix} x\ y\ |I_x|\ |I_y|\ |I_{xx}|\ |I_{yy}|\ \theta(x,y) \end{bmatrix} \tag{10}$$

$$f_k = \begin{bmatrix} x\ y\ |I_x|\ |I_y|\ |I_{xx}|\ |I_{yy}| \end{bmatrix} \tag{11}$$

$$f_k = \begin{bmatrix} x\ y\ I(x,y)\ |I_x|\ |I_y|\ |I_{xx}|\ |I_{yy}|\ \theta(x,y) \end{bmatrix} \tag{12}$$

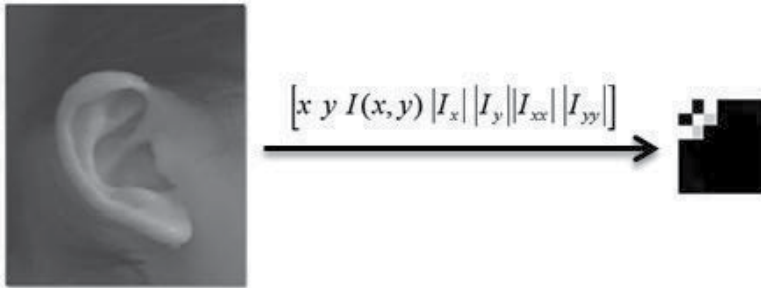Figure 1, denotes a sample covariance matrix for a given image.



Fig. 1. Covariance matrix provided for these seven features

Despite RCM advantages, computation of the covariance matrices for all rectangular regions within an image is computationally prohibitive using the routine methods. Several applications such as detection, segmentation, and recognition require computation and comparison of covariance matrices of regions. However, routine methods disregard the fact that there exist a high number of overlaps between those regions and the statistical moments extracted for such overlapping areas can be utilized to enhance the computational speed.

### 3.2 Fast covariance computation using integral images

Instead of repeating the summation operator for each possible window as described by (Veksler, 2003 ; Porikli, 2005), we can calculate the sum of the values within rectangular windows in linear time. For each rectangular window we need a constant number of operations to calculate the sums over specific rectangles many times. First, we should define the cumulative image function. Each element of this function is equal to the sum of all values to the left and above of the pixel including the value of the pixel itself. We can calculate the cumulative image for every pixel with four arithmetic operations per pixel. Then we should calculate the sum of image function in a rectangle. This operation can be computed with another four arithmetic operations with some modifications at the border. Therefore by using a linear amount of computation, the sum of image function over any rectangle can be calculated in linear time.

Integral images are intermediate image representations used for fast calculation of region sums (Viola & Jones, 2001). Later Porikli (Porikli, 2005) was extended this idea for fast calculation of region covariances. He presented that the covariances can be obtained by a few arithmetic operations with a series of integral images.

We can rewrite (i, j)-th element in covariance matrix which introduces in (2) as

$$C_R(i,j) = \frac{1}{n-1}\sum_{k=1}^{n}(f_k(i)-\mu(i))(f_k(j)-\mu(j)) \tag{13}$$

By expanding the mean we have

$$C_R(i,j) = \frac{1}{n-1}\left[\sum_{k=1}^{n}f_k(i)f_k(j) - \frac{1}{n}\sum_{k=1}^{n}f_k(i)\sum_{k=1}^{n}f_k(j)\right] \tag{14}$$

To compute region R (rectangular region) covariance, we need to calculate the sum of each feature dimension $f(i)_{i=1..n}$ as well as the sum of multiplication of any two feature dimensions $f(i)f(j)_{i,j=1..n}$. In this stage, we can use a series of integral images to compute these sums with a few arithmetic operations.

For each feature dimension $f(i)$ and multiplication of any two feature dimensions $f(i)f(j)$ we should construct integral images. Finally, we have $d+d^2$ integral images. Define p as the $W \times H \times d$ tensor of the integral images along each feature dimensions.

$$P(x',y',i) = \sum_{x<x',y<y'} F(x,y,i) \tag{15}$$

And define $Q$ as the $W \times H \times d \times d$ tensor of the second order integral images.

$$Q(x',y',i,j) = \sum_{x<x',y<y'} F(x,y,i)F(x,y,j) \tag{16}$$

$P_{x,y}$ is the $d$ dimensional vector and $Q_{x,y}$ is the $d \times d$ dimensional matrix.

$$P_{x,y} = \begin{bmatrix} P(x,y,1) \dots P(x,y,d) \end{bmatrix}$$

$$\tag{17}$$

$$Q_{x,y} = \begin{pmatrix} Q(x,y,1,1) & \dots & Q(x,y,1,d) \\ & \vdots & \\ Q(x,y,d,1) & \dots & Q(x,y,d,d) \end{pmatrix}$$

If we have the rectangular region as $R(x',y';x'',y'')$ shown in figure 2, the covariance of the region that bounded by $(1,1)$ and $(x'',y'')$ is

$$C_{R_{(1,1;x'',y'')}} = \frac{1}{n-1} \left[ Q_{x'',y''} - \frac{1}{n} P_{x'',y''} P_{x'',y''}^T \right] \tag{18}$$

Where $n = x'' \times y''$. In the same way, the covariance of the region $R(x',y':x'',y'')$ is

$$C_{R_{(x',y';x'',y'')}} = \frac{1}{n-1} \Big[ Q_{x'',y''} + Q_{x',y'} - Q_{x'',y'} - Q_{x',y''}$$

$$- \frac{1}{n} (P_{x'',y''} + P_{x',y'} - P_{x',y''} - P_{x'',y'})$$ $$\tag{19}$$

$$\cdot (P_{x'',y''} + P_{x',y'} - P_{x',y''} - P_{x'',y'})^T \Big]$$

Where $n = (x'' - x') \times (y'' - y')$. Therefore, by using the integral images, the covariance of each rectangular region can be computed in $O(d^2)$ time. In our method we used integral image based covariance computation as a fast approach for RCM computation of the given features.
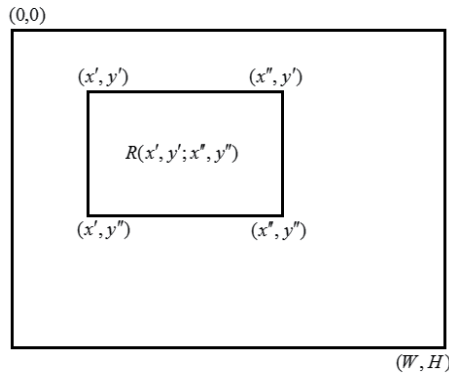
Fig. 2. Rectangular region R

### 3.3 Covariance matrix distance calculation

Since RCMs lie on connected Riemannian manifold, the Euclidean distance is not proper for our features, for instant, this space is not closed under multiplication with negative scalars. We use the distance measure presented in (Forstner & Moonen, 1999) to compute the distance/dissimilarity of the covariance matrices.

$$\rho(C_1,C_2) = \sqrt{\sum_{i=1}^{d} \ln^2 \lambda_i(C_1,C_2)} \tag{20}$$

where $\lambda_1(C_1,C_2),\dots,\lambda_d(C_1,C_2)$ are generalized eigenvalues of $C_1,C_2$ and computed from

$$\lambda_i C_1 x_i = C_2 x_i \qquad i = 1\dots d \tag{21}$$

where $x_i \neq 0$ are the generalized eigenvectors.

## 4. Gabor-based region covariance matrix

### 4.1 Gabor features extraction

The RCM-based methods with feature mapping functions (9),(10) have great success in people detection, object tracking, and texture classification (Tuzel et al., 2006; Tuzel et al., 2007). However our experimental results showed that the recognition rates of these methods are very low when being applied to ear recognition which is a very difficult task from the classification point of view. We construct effective features for RCM by using Gabor features and pixel location and illumination intensity component, to get better result in ear recognition. The biological relevance and computational properties of Gabor wavelets for image analysis have been investigated in (Jones & Palmer, 1987).

The Gabor features of ear images are robust against illumination changes. Gabor representation facilitates recognition without correspondence, because it captures the local structure corresponding to spatial frequency (scale), spatial localization, and orientation selectivity (Schiele & Crowley, 2000).

Daugman (Daugman, 1985) modeled the responses of the visual cortex by Gabor functions because they are similar to the receptive field profiles in the mammalian cortical simple cells. Daugman (Daugman, 1985) enhanced the 2D Gabor functions (a series of local spatial

bandpass filters), which have good spatial localization, orientation selectivity, and frequency selectivity. Lee (Lee, 2003) gave a good description to image representation by using Gabor functions. A Gabor (wavelet, kernel, or filter) function is the product of an elliptical Gaussian envelope and a complex plane wave as

$$\varphi_{\mu,v}(\overline{x}) = \frac{\left\|\overline{k}\right\|^2}{\sigma^2} e^{(-\frac{\left\|\overline{k}\right\|^2 \left\|\overline{x}\right\|^2}{2\sigma^2})} \left[ e^{i\overline{k}\overline{x}} - e^{-\frac{\sigma^2}{2}} \right] \tag{22}$$

Where $\overline{x} = (x,y)$ is the variable in a spatial domain, and $\overline{k}$ is the frequency vector, which determines the scale and direction of Gabor functions $\overline{k} = k_v e^{i\phi_\mu}$, where $k_v = k_{max} / f^v$, with $k_{max} = \pi / 2$. In our application, $f = \sqrt{2}$ and $\phi_\mu = \pi\mu / 8$. The term $\exp(-\sigma^2 / 2)$ is subtracted in order to make the kernel DC-free and, thus, insensitive to illumination. Examples of the real part of Gabor functions used in this chapter are shown in Figure 3. We use Gabor functions with five different scales $(v)$ and eight different orientations $(\mu)$, making a total of 40 Gabor functions. The number of oscillations under the Gaussian envelope is determined by $\sigma = 2\pi$
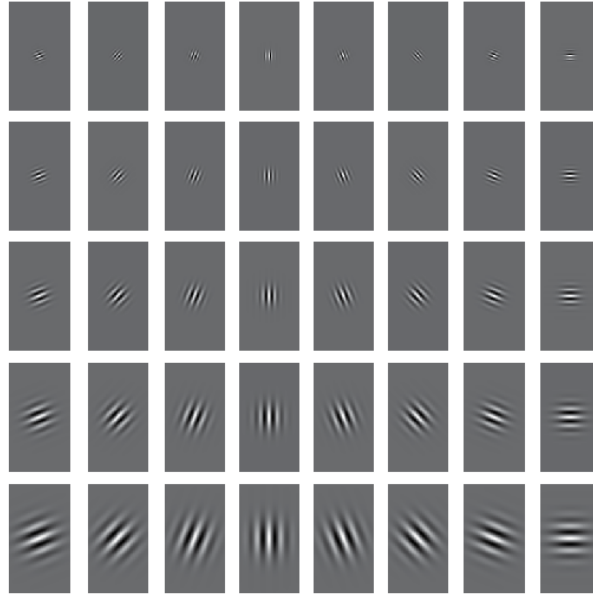


Fig. 3. The real part of gabor function for five different scales and eight different orientations

The gabor kernels family is constructed by taking five scales $(v \in \{0,...,4\})$ and eight orientations $(\mu \in \{0,...,7\})$. The gabor features can be achieved by convolving the gabor kernels with the image $I$

$$g_{\mu,v}(x,y) = \left| I(x,y) * \varphi_{\mu,v}(x,y) \right| \tag{23}$$

Where $|.|$ is a magnitude operator. $g_{\mu,v}(x,y)$ are the gabor representation of an image at orientation $\mu$ and scale $v$. Figure 4 shows the magnitude of gabor representation of an ear image.
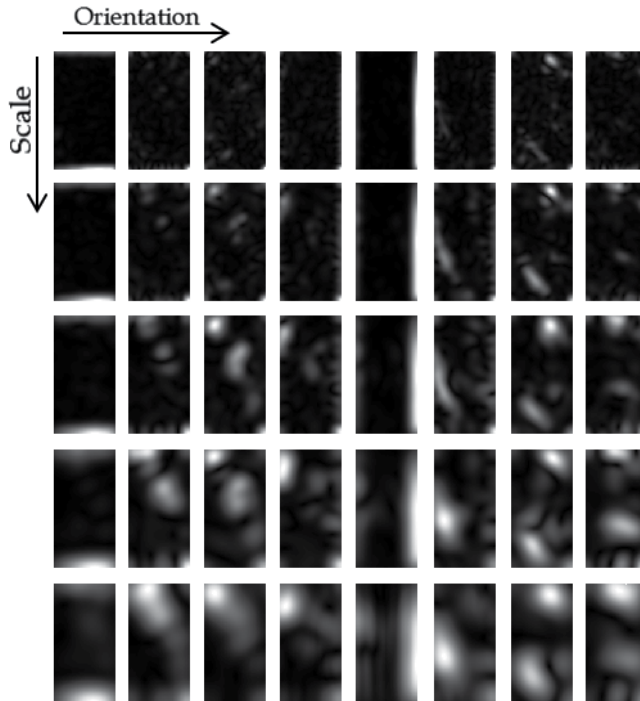
Fig. 4. The magnitude part of gabor representation of an ear image

## 4.2 Gabor based RCM

We propose a new gabor-based feature mapping function to construct effective and robust RCM.

$$f_k = \begin{bmatrix} x & y & I(x,y) & g_{0,0}(x,y) & g_{0,1}(x,y) & \cdots & g_{7,4}(x,y) \end{bmatrix} \tag{24}$$

Where $I(x,y)$ is the pixel illumination intensity and $g_{\mu,v}(x,y)$ are the gabor representation of the ear image. By substituting (24) into (2), we have the gabor-based region covariance matrices in region R $(C_R)$. $C_R$ dimntionality is $43 \times 43$.

In our method, we represent each ear image with five RCMs extracted from five different regions $(C_1,\ldots,C_5)$. First RCM $(C_1)$ defined over the whole ear image, so it gives us a global representation of the ear image. Four other RCMs are defined over part of the ear image, so they give us the part-based representation of the ear image. In order to increase the robustness of our method against illumination variations, we use both global and part-based representations for ear images in our method. Figure 5, denotes these five regions for $C_1, C_2, C_3, C_4, C_5$.

For computing the distance between a gallery RCM and a Probe RCM, we use

$$\rho(G,P) = \min_j \left[ \sum_{i=1}^{5} \rho(C_i^G, C_i^P) - \rho(C_j^G, C_j^P) \right] = \sum_{i=1}^{5} \rho(C_i^G, C_i^P) - \max_j \left[ \rho(C_j^G, C_j^P) \right] \tag{25}$$

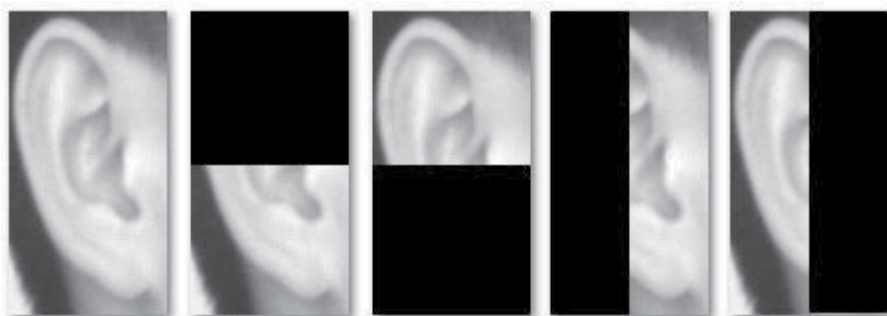Where $C^G$ and $C^P$ are RCMs from gallery and probe sets.

Fig. 5. Five regions for covariance matrices of a sample ear image

Sometimes one local RCM, due to illumination variation or noise, may be affected so much that make its corresponding distance unreliable. That is the reason why we subtracted the most unreliable part in (25) from the summation of all distances between gallery and probe RCMs. We used nearest neighbor classifier with the distance in (25) for our method.

## 5. Databases

Our method tested on two USTB databases (Yuan et al., 2005). Database 1 includes 180 images of human ear corresponding to 60 individual with three images per person. All the images in database 1 acquired under standard condition with a little changes. Figure 6, denotes sample ear images from database 1.
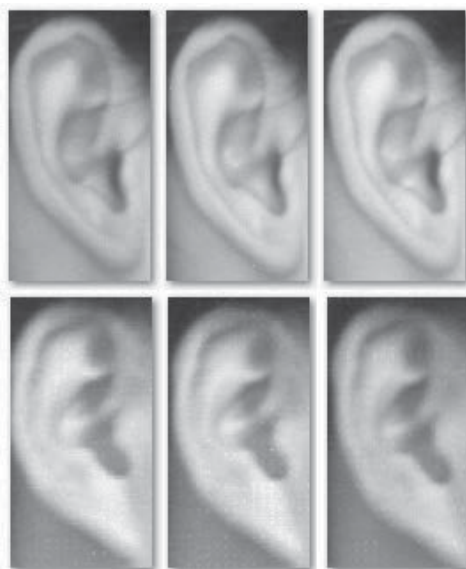


Fig. 6. Sample ear image for two persons from database 1

Database 2 includes 308 images of human ear corresponding to 77 individual with four images per person. All the images in database 1 acquired under illumination variation and ± 30 degree pose variations. Figure 7, shows sample ear images from database 2.
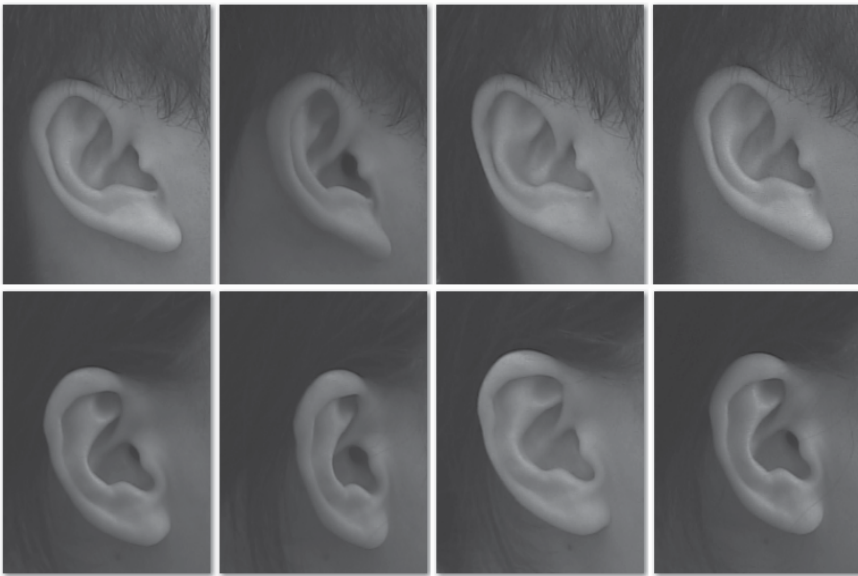
Fig. 7. Sample ear image for two persons from database 2

## 6. Experimental result

We performed our experimental studies comparing various ear reconigtion algorithms including our method with PCA+RBFN method (Zhang & Mu, 2008), ICA+RBFN method (Zhang & Mu, 2008), Hmax+SVM method (Yaqubi et al., 2008), LSBP method (Guo & Xu, 2008), four RCM-based methods (Tuzel et al., 2007; Pang et al., 2008). In order to compare the recognition performance of our method with the above methods, we have used USTB databases (Yuan et al., 2005) in our experiments. In database 1, from a total of 60 persons, two images per person where randomly used for training. There are three different ways of selecting two images for training from three images. In database 2, from a total of 77 persons, three images per person where randomly used for training. There are four different ways of selecting three images for training from four images.

For simplicity, RCM-based methods associated with (9), (10), (11), (12) denote by RCM1, RCM2, RCM3, RCM4 respectively. RCM3 is a subset of RCM1 with lack of intensity component; also RCM2 is a subset of RCM4 with lack of intensity component.

Figures 8 and 9 denote the mean of the recognition rates for database 1 and 2 datasets. From Figures 8 and 9, it can be seen that the recognition performances of four RCM-based methods were worse than other methods, so it can be concluded that the discrimination power, in these RCM-based methods are weak for recognition task. To find out about the intensity parameter $(I(x,y))$ effect on the recognition rate, we compare the result of RCM1 with RCM3 and the result of RCM2 with RCM4. We can conclude that $I(x,y)$ is an important feature in RCMs and it contributes to increasing the recognition performance of RCM-based methods. Thus, we used the illumination intensity component in our mapping function to increase the accuracy of our method.

Table 1 shows the comparision of the standard deviation of recognition performance between all discussed methods on database 1 and 2. From table 1, We can see that the

standard deviation of our method for database 1 are low. Therefore, our method showed better performance than any other methods in database 1. The mean recognition rates of our method in database 1 and 2 are 93.33% and 87.98% respectively. Due to the pose variations in database 2 images, the recognition performance of our method, in terms of average accuracies, outperforms any other methods, except LSBP and ICA methods.
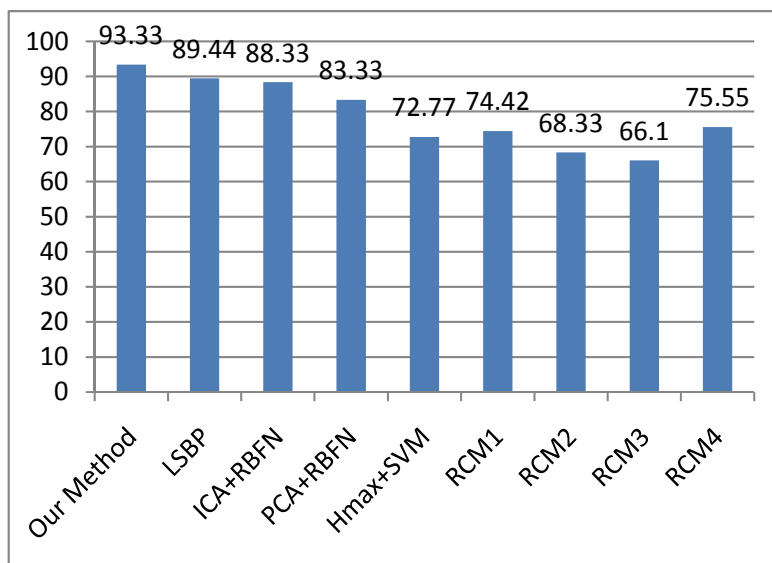


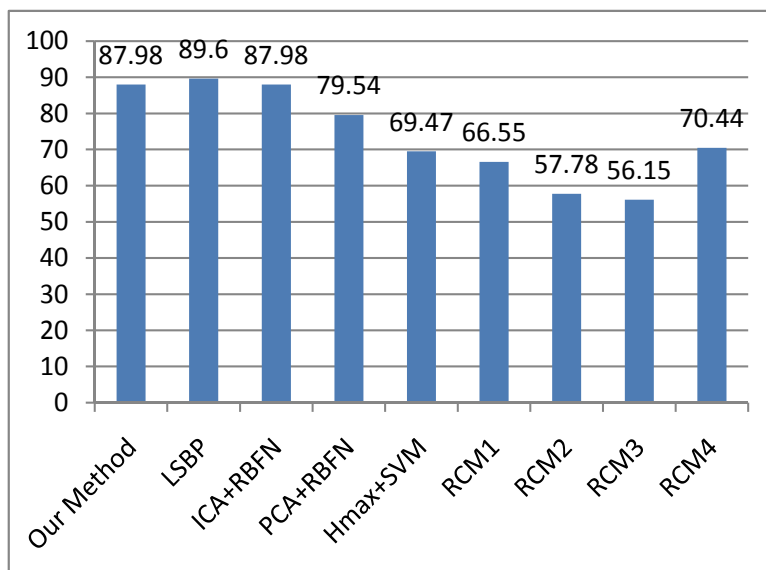Fig. 8.  Mean Recognition rates of different methods on database 1  (%)



Fig. 9.  Mean Recognition rates of different methods on database 2  (%)

| Methods | Standard Deviation | |
|---|---|---|
| | Database 1 | Database 2 |
| Our method | 1.67 | 5.23 |
| LSBP | 1.92 | 4.74 |
| ICA+RBFN | 3.33 | 4.15 |
| PCA+RBFN | 3.53 | 3.07 |
| $H_{max}$+SVM | 1.93 | 2.70 |
| RCM1 | 2.58 | 2.22 |
| RCM2 | 3.33 | 4.80 |
| RCM3 | 2.55 | 2.72 |
| RCM4 | 2.54 | 3.06 |

Table 1. Standard deviations of the recognition rates

Eventually, these results prove that using Gabor features, as main features in constructing RCMs, will improve the discrimination ability for recognizing ear images, and it shows better recognition rate in proportion to previous methods.

## 7. Conclusion

In this chapter, we proposed gabor-based region covariance matrices for ear recognition. In this method we form region covariance matrix by using gabor features, illumination intensity component, and pixel location and utlize it as an efficient ear descriptor. We compared our method with PCA+RBFN method (Zhang & Mu, 2008), ICA+RBFN method (Zhang & Mu, 2008), Hmax+SVM method (Yaqubi et al., 2008), LSBP method (Guo & Xu, 2008), and four RCM-based methods (Tuzel et al., 2007; Pang et al., 2008), using two USTB databases.

Unlike the previous RCM-based methods which have very low recognition rates when being applied to ear recognition, our RCM-based method, which used gabor features as a main feature for constructing RCM, showed better result in ear recognition. Potential results showed that our method achieved improvement, in terms of recognition rate, in proportion to other methods. Our method obtains the average accuracy of 93.33% and 87.98%, respectively, on the databases 1 and 2 for ear recognition.

## 8. References

Abdi, H. & Williams, L. J. (2010), Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2, No. 4, pp. 433–459.

Bertillon, A. (1890). *La Photograhie Judiciaire, avec un appendice sur la classification et l'Identification Anthropometriques*, Gauthier-Villars.

Comon, P. (1994). Independent Component Analysis, A New Concept?. *Signal Processing*. Vol. 36, No. 3, pp. 287-314.

Daugman, J.G., (1985). Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimized by Two-Dimensional Visual Cortical Filters. *Journal of Optical Soc. Am.*, Vol. 2, No. 7, pp. 1160-1169.

Forstner, W., & Moonen B. (1999). A metric for covariance matrices. *In Dept geodesy Geoinform, Stuttgart University.* Stuttgart, Germany: Tech. Rep.

Forsyth, D. A., & Ponce J. (2002). *Computer Vision: A Modern Approach.* Prentice Hall.

Guo, Y., & Xu., Zh. (2008). Ear recognition using a new local matching approach. *Proceedings of IEEE International Conference on Image Processing.*

Iannarelli, A. 1989. *Ear Identification, Forensic Identification Series*, Paramount Publishing Company, Fremont, California.

Jones, J., & Palmer L. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophys.* Vol. 58, No. 6, pp. 1233–1258.

Lee, T.S. (2003). Image Representation Using 2D Gabor Wavelets. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, No. 10, pp. 959-971.

Pang Y., Yuan Y., & Li, X. (2008). Gabor–based region covariance matrices for face recognition. *IEEE Trans. On Circuits and Systems for Video Technology*, Vol. 18, No. 7, pp. 989–993.

Porikli, F. (2005). Integral Histogram: A fast way to extract histograms in Cartesian spaces. *Proceedings of CVPR.* 2005.

Riesenhuber, M. & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.*, Vol. 2, No. 11, pp. 1019 – 25.

Schiele, B., & Crowley J. L. (2000). Recognition without correspondence using multidimensional receptive field histograms. *Int. J. Comput. Vis.* Vol. 36, No. 1, pp. 31–52.

Stone, J. (2005). Independent Component Analysis: A Tutorial Introduction. *The Knowledge Engineering Review archive.* Vol. 20, No. 2, June 2005.

Tuzel, O., Porikli, F., & Meer, P. (2006). Region covariance: A fast descriptor for detection and classification. *Proceedings of Eur. Comput. Vision Conference*, pp. 589–600.

Tuzel, O., Porikli, F., & Meer, P. (2007). Human detection via classification on Riemannian manifolds. *Proceedings of IEEE Comput. Vision Pattern Recog. Conference*, pp. 1–8.

Veksler, O. (2003). Fast variable window for stereo correspondence by integral images. *Proceedings of CVPR.* 2003.

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, HI. Vol. 1, pp. 511–518.

Xu, L. (1994). Theories of Unsupervised Learning, PCA and Its Nonlinear Extension. *Proceedings of IEEE International Conference on Neural Network.* Olando. USA. pp. 1254-1257.

Yaqubi, M., Faez, K., & Motamed, S. (2008). Ear recognition using features inspired by visual cortex and support vector machine technique. *Proceedings of IEEE International Conference on Computer and Communication Engineering*, Kula Lampur, Malaysia.

Yuan, Li, Mu, Zhichun, & Zhengguang, Xu. (2005). Using ear biometrics for personal recognition. *International Workshop on Biometric Recognition Systems.* IWBRS 2005, pp. 221–228.

Zhang, H., & Mu, Zh. (2008). Compound Structure Classifier System For Ear Recognition. *Proceedings of the IEEE International Conference on Automation and Logistics,* Qingdao, China.

# Bi-Modality Anxiety Emotion Recognition with PSO-CSVM

Ruihu Wang[1,2] and Bin Fang[2]
*1Chongqing University of Arts and Sciences,*
*2Chongqing University,*
*China*

## 1. Introduction

In Human Computer Interaction (HCI) community, the ultimate goal which we are striving to achieve is to create a natural, harmony way of bi-directional communication between machine and human. As we well known, many machine intelligence applications are based on machine vision technology, for instance, industrial detection, face recognition, medical image computer aided diagnose (MICAD), fingerprint recognition, and so on. In terms of artificial intelligence, it is built upon digital image processing and video frame analyzing to extract feature information to recognize some interesting objects. Moreover, a high-level machine learning system should be capable of identifying human emotion states and make interactions accordingly. Multimodal human emotion recognition involving facial expression and motion recognition could be applied to intelligent video surveillance system to provide an early warning mechanism in case of potential unsafe action occurring. There is no doubt that the ever increasing various kinds of crimes in our modern society which make the living environment around us even worse, demand for an intelligent and automatic security precautions measures to offer people a more convenient, relaxed living conditions. Meanwhile, automatic intelligent video-based surveillance system has received a lot of interest in the computer vision and human computer interaction community in recent years. CMU's Video Surveillance and Monitoring (VSAM) project [26] and MIT AI Lab's Forest of Sensors project [27] are examples of recent research efforts in this field. As a matter of fact, the safeguard has to keep watch on lots of screens in control center in real application. The video displayed on the screens are captured from cameras which are distributed in various security-sensitive areas such as elevator, airports, railway station or public places. However, because of human's inherent information processing limitation, it is impossible to pick up all of the useful information from the monitors at the same time properly and spontaneously. That means some potentially dangerous information could be missed out. We expect the human computer interaction system is able to take the information-processing burden off the human. An ideal and effective intelligent surveillance system should work automatically without or with minimal human intervention. In addition, we believe that an intelligent surveillance system should be capable of preventing and predict criminal occurring by biometric recognition, rather than identifying the suspect after attacks happened. Figure 1 shows some pictures in which the subjects are under the anxiety or stress emotion state when she interviewed with human resource manager, not criminal nevertheless.

Fig. 1. Anxiety or stress emotion representation

A number of universities and research groups, commercial companies are conducting meaningful projects and have made significant progress in visual surveillance monitoring system. The Safehouse Technology Pty Ltd in Australia developed The Clarity Visual Surveillance System (CVSS)[28], which provides security professionals with the next generation in visual surveillance management. Although state-of-the-art computer vision technology and artificial intelligence algorithm were used, the system did not integrate affective computing methods in it yet. Picard defined "Affective Computing" as: computing that relates to, arises from, or deliberately influences emotions. Affective computing expands human-computer interaction by including emotional communication together with appropriate means of processing affective information [29]. In affective computing research, a variety of input signals like visual, verbal, touch/wearable computing, etc, have been used into emotion communication frequently in most cases. However, visual channel signal, which involves facial expression and gesture motion would be considered only in our surveillance system. It is a typical non-intrusive, intelligent analyzing technology and easily to be implemented. It does not require the subject to comply with some requirement he/she probably won't to meet, for example, iris recognition, fingerprint recognition and wearable device. In a survey of hundreds of US security executives, Buxton[1] pointed out that system which could process the video from the increasing number of cameras were "one of the top items in demand". We believe that visual surveillance based on emotion perception and recognition is the next generation of intelligent monitoring system. More importantly, this system can be capable of supervising individuals in specific environment to carry out data analysis and recommend possible interventions.

## 2. Previous work

Psychological research findings suggest that humans make judgment action depend on the combined visual modalities of face and body more effective than any other channel [18]. Since Paul Ekman and Frisen[3] divided human emotion into six primary categories: happiness, sadness, disgust, angry, surprise, fear, there has been a great amount of research on facial expression recognition. Eckman revealed that the facial expression showed the inside emotional state and the specific sense at that time, which was hardly controlled by human artificially. In automatic facial expression recognition, feature extraction and classifier design are two major procedures. As shown in Table 1, the most frequently being used feature extraction method is Gabor wavelet feature. A subset of these filters is chosen using AdaBoost, which is transmitted for training SVM as a reduced feature representation. A wide range of surveillance systems have been developed for different applications. G. Lu, X. Li, H. Li proposed a surveillance system for nurses to distinguish the neonatal pain expression from non-pain expression automatically using Gabor wavelet transform and support vector machine. They reported 85.29% recognition rate of pain versus non-pain [2]. The Safe-house Technology Pty Ltd in Australia developed Clarity Visual Surveillance System. It employs state-of-the-art computer vision algorithm and artificial intelligence, whereas it did not involve affective computing technology and emotion recognition in it yet.

Bouchrika introduced gait analysis for visual surveillance considering that the identification of the individuals who are suspected of committing crimes is more important than predict when a crime is about to happen [15]. P. Rani, N.Sarkar and J.Adams built an anxiety-recognition system capable of interpreting the information contained in physiological signals processes to predict the probable anxiety state [17]. In [18] H.Kage et.al introduced pattern recognition technologies for video surveillance and physical security. Their system detects the occurrence of a relevant action via image motion analysis. Optical-flow is used to analyze motion and AdaBoost is been for face detection. In [18], H.Gunes and M.Piccardi presented an approach to automatic visual recognition of expressive face and upper-body gestures from video sequences which are suitable for human communicative behavior. They defined some rules for facial expression and body motion recognition respectively to make right decision firstly. Min-max analysis method is used to detect the eyebrows, eyes, mouth and chin by evaluating the topographic gray level relief. Body feature is extracted by using traditional image processing methods such as background subtraction and dilation.

In order to improve the accuracy of classifiers and provide fast, pertinent response, many researchers proposed optimized approaches. Typically Genetic Algorithm or Particle Swarm Optimization is often used for feature selection and parameters optimization [19,20,21,22]. In our work, the key problems are facial expression, head motion and hand gesture analysis. We also care more about run-time consumption which impacts the recognition efficiency.

| Feature Extraction Methods | Classifier design | Optimized Feature Selection |
|---|---|---|
| Gabor wavelet[2,4,5,6,10,16] Optical Flow[7], PCA[12] FLD, Eigenfaces[9], AdaBoost[10,12,13],LBP[13] | Hide Markov Model [11] Support Vector Machine[2,6,7,8], Artificial Neural Network[4,5], Bayesian Network[6,16], AdaBoost[12,13] | AdaSVM[6], MIFS[16] |

Table 1. Overview of methods used in facial expression recognition

## 3. Methods

### 3.1 System framework

A visual automated surveillance system consists of three main phases: face and body motion detection, facial feature and motion feature extraction, classification and recognition, as shown in Figure 2. By means of sensitive facial expression extraction, the subject's intention and prospective behaviour are analyzed to report any suspicious expression or activities to the control central. The system would be able to reduce potential crimes by recognizing suspicious individuals' emotion state ahead of security threats happening.

The front end inputs involve three different parts: (1)static facial image database input; (2)static face image database input, and (3) real face image detection captured by cameras using image sequence analysis. All of these three channels have to be passed through the facial expression feature extraction procedure. AdaBoost, Optical Flow and Gabor feature can be used to extract facial expression feature. In the middle of the system framework, both automatic facial expression recognition and human motion recognition are implemented through classifier learning and test, which involve PSO-SVMs feature extraction and parameters optimization. In our system, we use cascaded SVM architecture combined with particle swarm optimization algorithm to train and test SVM to get an optimized classifier which runs faster in computational-consuming conditions.
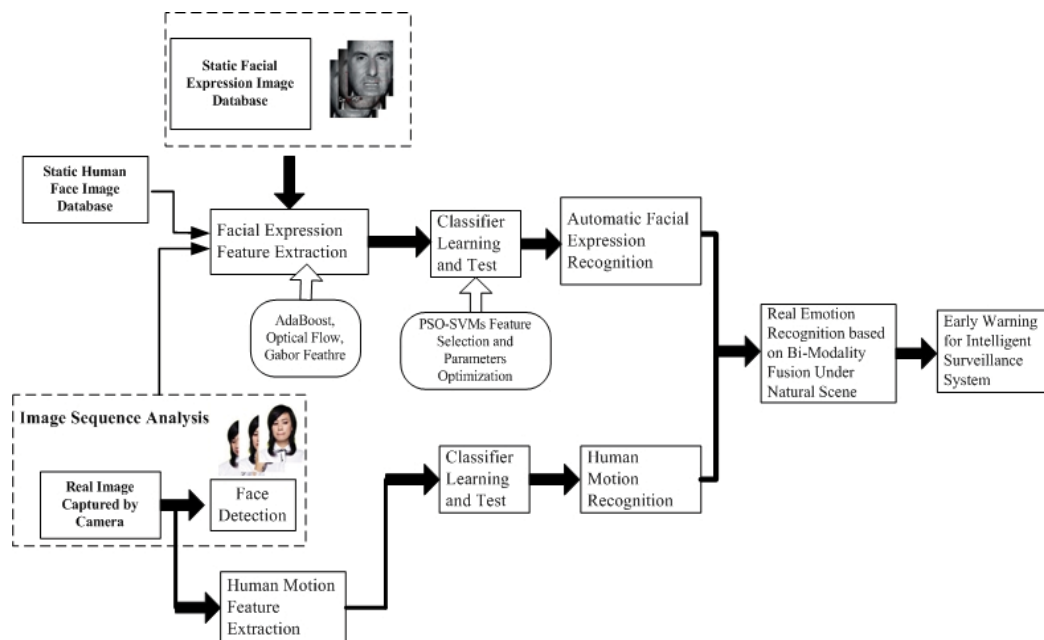
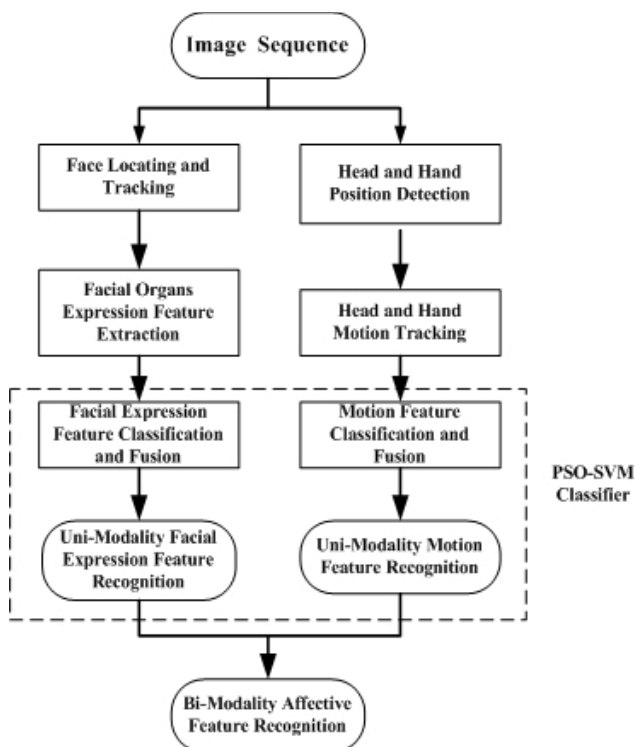Fig. 2. System Framework Schematic of Human Motion Recognition



Fig. 3. Bi-Modality Fusion Affective recognition

The integration of Bi-Modality feature fusion classification for affective state recognition is composed by two uni-modality signals, as shown in Fig 3. One is facial expression feature classification and fusion, the other is motion feature classification and fusion. The essential aspect for computer to understand human expression is to achieve head's rigid motion and face's non rigid motion by which machine can track and analyze the facial expression change. Among this, the head's rigid motion can be denoted by six parameters: three rotation $R(rx, ry, rz)$ parameters: head shifting around $x$, $y$ and $z$ axis, and three translation $\tau(t_x, t_y, t_z)$ parameters: downward and upward, left and right, front and backward. Non rigid motion of Face includes movement of mouse, eye, eyebrow and winkle of forehead.

### 3.2 Cascaded SVM architecture

Support vector machine was developed by Vapnik from the theory of Structural Risk Minimization. However, the classification performance of the practically implemented is often far from the theoretically expected. In order to improve the the classification performance of the real SVM, some researchers attempt to employ ensemble methods, such as conventional Bagging and AdaBoost [29]. However, in [30,31], AdaBoost algorithm are not always expected to improve the performance of SVMs, and even they worsen the performance particularly. This fact is SVM is essentially a stable and strong classifier.

Considering the problem of classifying a set of training vectors belonging to two separate classes,

$$T = \{(x_1, y_1), \cdots, (x_l, y_l)\} \subseteq X \times Y \tag{1}$$

where

$$x_i \in X \subset R^n, y_i \in Y = \{-1, +1\}, i = 1, 2, \cdots, l.$$

SVM can be trained by solving the following optimization problem:

$$\min_w \Phi(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_i \xi_i \tag{2}$$

subject to

$$y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \xi_i, i = 1, 2, \cdots l \tag{3}$$

where $\xi_i > 0$ is the $i$-th slack variable and $C$ is the regularization parameter.

Nonlinear SVM are known to lead to excellent classification accuracies in a wide-range of tasks. It utilizes a set of support vectors to define a boundary between classes, which is dependent on a kernel function. However, as a classifier, SVM is usually slower than neural networks. The reason for this is that the run-time complexity is proportional to the number of support vectors, i.e. to the number of training examples that the SVM algorithm utilizes in the expansion of the decision function [23].

The optimization problem of Nonlinear SVM with kernel function is denoted by

$$\alpha^* = \arg\min_\alpha \frac{1}{2}\sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) - \sum_{j=1}^l \alpha_j, \tag{4}$$

subject to

$$\sum_{i=1}^{l} y_i \alpha_i = 0,$$
$$C \geq \alpha_i \geq 0, i = 1, \cdots, l.$$

(5)

If $\alpha^* = (\alpha_1^*, \cdots, \alpha_l^*)^T$ is a solution, then

$$\omega^* = \sum_{i=1}^{l} y_i \alpha_i^* x_i, b^* = y_j - \sum_{i=1}^{l} y_i \alpha_i^* (x_i \cdot x_j), \forall j \in \{j \mid \alpha_j^* > 0\}$$

(6)

And $(\omega^*, b^*)$ is the optimized solution of (2.3)~(2.4). where $K(x_i \cdot x_j)$ is the kernel function performing the non-linear mapping into feature space. One of the most common used kernel function is Gaussian Radius Basis Function which is with the form,

$$K(x, z) = \exp\left(-\frac{\|x - z\|^2}{\sigma^2}\right)$$

(7)

According to Yang and Honvar [32], the choice of feature used to represent patterns that are presented to a classifier has great impact on several pattern recognition properties, including the accuracy of the learned classification algorithm, the time need for learning a classification function, and the number of examples needed for learning, the cost associated with the features. In addition to feature selection, C.Huang and C. Wang [33] suggested that proper parameters setting can also improve the SVM classification accuracy. The parameters include penalty parameter C and the kernel function parameter $\sigma$ for RBF, which should be optimized before training. D. Iakovidis et.al. proposed a novel intelligent system for the classification of multiclass gene expression data. It is based on a cascading support vector machines scheme and utilize Welch's t-test for the detection of differentially expressed genes. In their system, a 5-block cascading SVMs architecture was used for the 6-class classification problem as shown in Figure 4 [43]. The classification performance was evaluated by adopting a Leave-One-Out (LOO) cross validation approach. LOO is commonly used when the available dataset is small providing an almost unbiased estimate of the generalization ability of a classifier[44] .
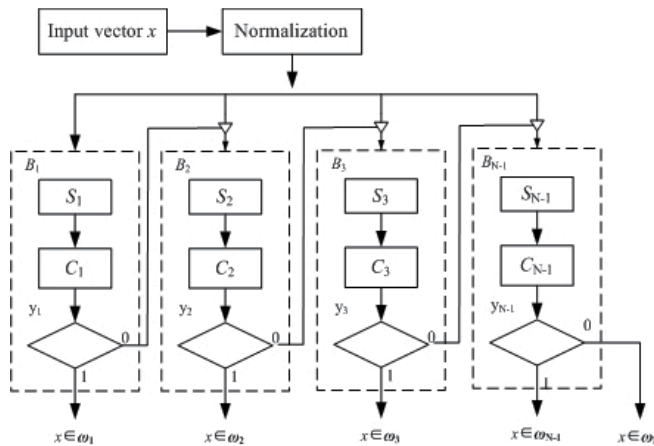


Fig. 4. Serial cascade SVMs for gene expression data classification

In order to achieve optimal feature subset selection and SVM-RBF parameters, Hsu and Lin [34] proposed a Grid algorithm to find the best C and sigma for RBF kernel. However their method has expensive computational complexity and does not perform well. Genetic algorithm is an another alternative tool, which has the potential to generate both the optimal feature subset and SVM parameters at the same time. Huang and Wang [33] conducted some experiments on UCI database using GA-based approach. Their result has better accuracy performance with fewer features than grid algorithm. Compared to Genetic Algorithm, Particle Swarm Optimization has no evolution operators such as crossover and mutation. There are few parameters to adjust. It works well in a wide variety of applications with slight variations [35].

In this paper, we proposed a cascaded SVM structure, as shown in Fig 6 to speed up body gesture classifier performance over conventional SVM-based methods without reducing detection rate too much and the hierarchical architecture of the detector also reduces the complexity of training of the nonlinear SVM classifier.

### 3.3 Standard particle swarm optimization algorithm

Particle Swarm Optimization was introduced firstly by James Kennedy and Russel Eberhart [36]. It is a population-based evolutionary computation search technique. In PSO, each potential solution is assigned a randomized velocity, and the potential solutions, called particles, fly through the problem space by following the current optimum particle. Each particle keeps track of its coordinates in hyperspace which are associated with the best solution (fitness) it has achieved so far. This value is called pbest. Another "best" value is also tracked. The "global" version of the PSO keeps track of the overall best value, and its location, which is called gbest [37]. Each particle is treated as a point in a D-dimension space. The original PSO algorithm is described below:

$$v_{id}^{k+1} = v_{id}^{k} + c_1 r_1 (p_{id} - z_{id}^{k}) + c_2 r_2 (p_{gd} - z_{id}^{k}) \quad (8.1)$$
$$z_{id}^{k+1} = z_{id}^{k} + v_{id}^{k+1} \quad\quad (8.2)$$

$$(8)$$

The ith particle's location vector is represented as $z_i = (z_{i1}, z_{i1}, \cdots z_{iD})$; the velocity is denoted by $v_i = (v_{i1}, v_{i1}, \cdots v_{iD})$ . $p_i = (p_{i1}, p_{i1}, \cdots p_{iD})$ and $p_g = (p_{g1}, p_{g1}, \cdots p_{gD})$ are pbest and gbest respectively. $r_1$ and $r_2$ are two random numbers in the range [0,1].

Equation (8) describes the flying trajectory of a population of particles, how the velocity and the location of a particle is dynamically updated. Equation (8.1) consists of three parts. The first part is the momentum part. The velocity is changed by current value. The second part is the cognitive part which represents the particle's learning capability from its own experience. The third part is the social part which represents the collaboration among all particles [38].

In order to improve the convergence performance of PSO algorithm, Shi and Eberhart proposed a modified particle swarm optimizer. An inertial weight w is brought into the original PSO algorithm. This w plays the role of balancing the global search and local search. It can be a positive constant or even a positive linear or nonlinear function of time [39].

$$v_{id}^{k+1} = w * v_{id}^{k} + c_1 r_1 (p_{id} - z_{id}^{k}) + c_2 r_2 (p_{gd} - z_{id}^{k}) \, (9.1)$$
$$z_{id}^{k+1} = z_{id}^{k} + v_{id}^{k+1} \quad\quad (9.2)$$

$$(9)$$

Simulations have been conducted on this modified PSO to illustrate the impact of this parameter introduced. It was concluded that the PSO with the inertial weight in the range [0.9,1.2] on average will have a better performance. A time decreasing inertial weight can also bring in a significant improvement on the PSO performance, as shown in Equation (10):

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{iter_{\max}} * k \qquad (10)$$

In this paper, PSO is utilized to search the optimal solution of a RBFN-SVM by minimizing the cost function $\Phi$ as the fitness function. Each particle is encoded as a real string representing the kernel centers ($z$) and widths ($\sigma$) as well as linear model coefficients $w$ and $b$. With the movement of the particles in the solution space, the optimal solution with a minimum value of the cost function $\Phi$ will be obtained. Optimizing the kernel centers and widths and the weights of the SVM model relating the feature variables synergistically keeps the model from getting trapped into a local optima and improves the model performance [40]. There are two key factors to determine the optimized hyperparameters when using PSO. One is how to represent the hyper-parameter as the particle's position, namely how to encode. The other problem is how to define the fitness value function which evaluate the goodness of a particle.

In this section, we describe the proposed SVM system for classification task. The aim of this system is to optimize the SVM classifier accuracy by detecting the subset of the best discriminative feature and solving the SVM model selection.
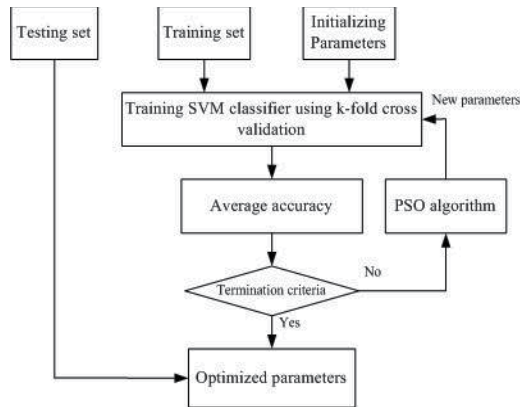


Fig. 5. PSO-SVM based feature selection and parameter optimization architechture scheme

### 3.3.1 PSO setup

The position $p_i \in \Re^{d+2}$ of each particle $P_i$ from the swarm is regarded as a vector encoding.

1.   Feature subset $f$ selection: f is a candidate subset of features, and F is a features set which consists of d available input features.

$$f = \{(x_1, \cdots, x_d) \mid (x_1, \cdots, x_d) \in F \subset \Re^d\}$$

2.   KBNF based SVM parameters optimization, includes $C$ and $\sigma$.

The position vector of each particle can be represented as the other kind of form:

$$p_i = \{(x_{i1},\cdots,x_{id},C_i,\sigma_i)\,|\,(x_{i1},\cdots,x_{id},C_i,\sigma_i) \in \Re^{d+2}\}$$

Let $f(i)$ be the fitness function of the $i$th particle $P_i$. As suggested by Melgani, the choice of the fitness function is important on which PSO evaluates each candidate solution $p_i$ based for designing SVM classifier. They explored a simple SV count as a fitness criterion in the PSO optimization framework.

### 3.3.2 SVM training and classification with PSO
The pseudo code of the proposed method for SVM-PSO classification is given below.

---

Let N be the particle number of particle swarm.

**1 Initialization**
    1.1 Initialize the population of N particles with random  positions and velocities on D dimensions in the solution space.

    1.2 Set the velocity vectors $v_i$ (i=1,2,...N) to zero.

    1.3 For each position $p_i \in \Re^{d+2}$ of the particle $P_i$ (i=1,2,...N) from the swarm, train the SVM classifier and compute the fitness function value.
**2 Particle swarm search**
    2.1 Detect the best global position $p_g$ in the swarm which showing the minimal value of the fitness function value over all explored trajectories.

    2.2 Update the speed and position of each particle using Equation (9).

    2.3 For each candidate particle $p_i$, train the SVM classifier and compute the fitness function $f(i)$.

    2.4 Update the best position of each particle $p_{bi}$ if its current position has a smaller fitness function.
**3 Convergence**
    3.1 If the maximum iteration times have reached, then exit, else return 2.1.
**4 SVM Training and Classification**
    4.1 Select the best global position $p_g^*$ of the particle swarm and train the SVM with the detected feature subset mapped by $p_g^*$ and modeled with the optimized parameters C and $\sigma$.
    4.2 Make SVM classification based on the trained classifier.

---

Table 2. Algorithm for SVM Training and Classification with PSO

Many researchers have proposed cascaded SVMs architecture to solve the computational complexity problem. Among the cascaded structure, there are mainly two kinds of SVMs architecture. One is parallel SVMs, and the other is serial SVMs. The parallel cascaded structure of SVM is more relied on hardware architecture of computer, which is developed based on decomposing the original complex problem into a number of independent simplified smaller problems, and in the end the partial results are combined into an ultimate

output [41]. Compared to parallel SVMs, the serial structure cascaded SVMs is more feasible and easily to be implemented. Y. Ma and X. Ding proposed a cost-sensitive SVM classifier to detect face [42]. In many classification cases, the cost of False Negative is far more than False Positive. In their cost-sensitive SVMs, different cost are assigned to two types of misclassification to train the cascaded SVMs in different stage of the face detector. The optimization goal is given by

$$\min\left\{\frac{1}{2}\|w\|^2 + C_p\sum_{y_i=1}\xi_i + C_n\sum_{y_i=-1}\xi_i\right\}$$

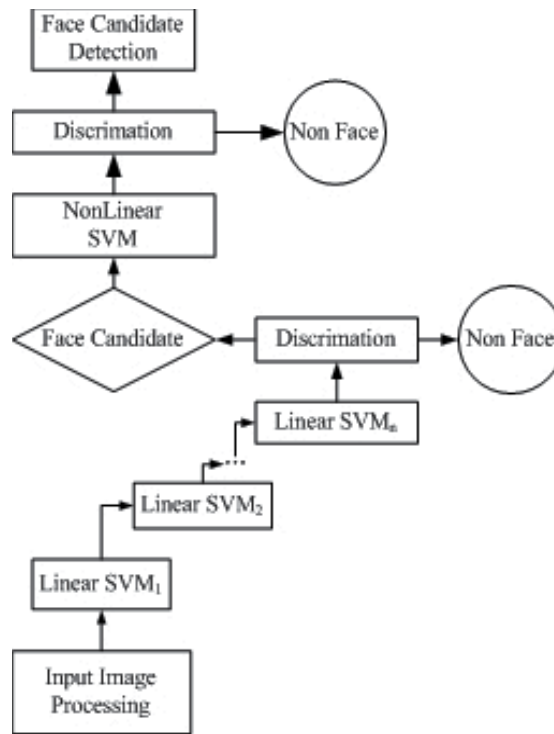where $C_p$ is the cost for face samples, and $C_n$ is for non-face samples, usually $C_p > C_n$.



Fig. 6. Cascaded Architecture of Serial SVMs

We propose here a novel approach named PSO-Cascaded SVMs classification method to combine Cascaded SVMs with particle swarm optimization algorithm. As shown in Figure 6, there are some linear SVMs cascaded to form a serial structure at the front end and a nonlinear SVM at the end of this system.

We think of two optional strategies to introduce particle swarm optimization into the cascaded architecture of serial SVMs: one is integrate PSO into the serial linear SVMs and the other is integrate PSO into the end nonlinear SVM only. Regarding that the serial SVMs structure at the front end is more easily to construct and implement than nonlinear SVM, we take the later strategy. The nonlinear SVM classification is more complicated because it has to deal with feature subset selection and kernel function parameters optimization.
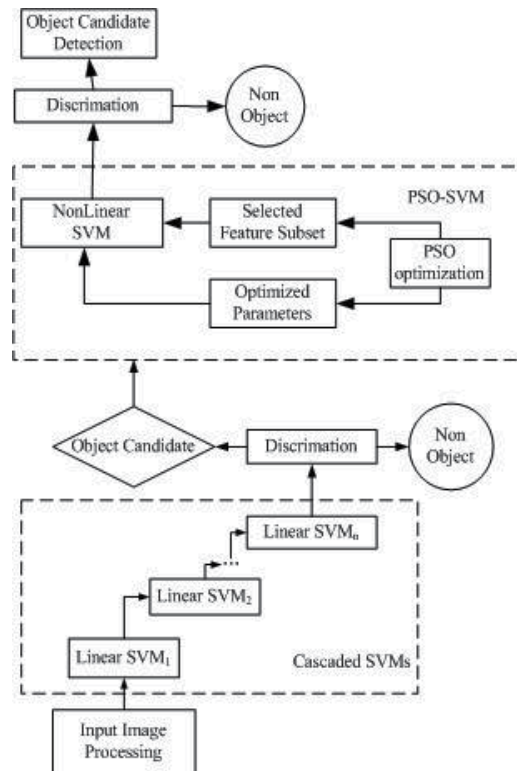
Fig. 7. PSO-Cascaded SVMs for Feature Selection and Parameters Optimization

In our experiment, only lip, eye and forehead changes are detected and analyzed. There are 4 states for lip, 3 states for eye and 3 states for forehead **individually**. We use 10-dimension feature vector to express 4*3*3 (36) kinds of states altogether, which means only 36 facial expressions are considered in our experiment, as shown in Table 3.

| Facial Regions | State Code | State Description | |
|---|---|---|---|
| | | Positive Class | Negative Class |
| lip | I.1 | Open | Close |
| | I.2 | Stretch | Depress |
| | I.3 | pout | pucker |
| | I.4 | Bend up at the corner of the mouth | Bend up at the corner of the mouth |
| eye | II.1 | Eyelid Open | Eyelid Mouth |
| | II.2 | Eyeball turn left | Eyeball turn left |
| | II.3 | Eyeball oversee | Eyeball look up |
| forehead | III.1 | Eyebrow raiser | Eyebrow lower |
| | III.2 | Eyebrow tighter | Eyebrow stretcher |
| | III.3 | Tipofbrow raiser, Eyeborws lower | Tipofbrow lower, Eyeborws raiser |

Table 3. Overview of methods used in facial expression recognition

The overall accuracy rate  for facial expression recognition with our proposed method can achieve 83.5% under uniform illumination. The performance of PSO-SVM classifier learning and testing for features selecting and parameters optimization outweighed other methods like PCA-SVM, PCA-RBF, whose overall accuracy rate achieved 91.46%, whereas PCA-SVM was 85.54% and PCA-RBF 83.27%,respectively.

## 4. Conclusion and future work

In this paper we have proposed a fusion method for facial expression and gesture recognition to build a surveillance system. Different from traditional six basic emotions on which many researchers have worked, we care only about the anxiety emotion. However, facial expression information is not enough for computer to recognize, we syncretize body gesture recognition to obtain a combined approach. Among many kinds of classifiers, SVM shows good performance for small samples classification. Meanwhile SVM is computational time expensive tool, we have to deal with some optimization such as feature subset selection before training and classifying to improve its performance. Especially for nonlinear SVM, the selection of kernel function and parameters optimization is more important. The cascaded SVMs structure show up superior performance in pattern classification.
Further more, as a simple stochastic global optimization technique inspired by social behavior of bird flocking, PSO can be used into the cascaded SVMs to select feature subset and optimize parameter for kernel function. The performance PSO-SVM strategies have been proposed in this paper for SVM to enhance its learning and classifying capability.

## 5. Acknowledgment

## 6. References

[1] H. Buxton. Learning and understanding dynamic scene activity: a review. Image and Vision Computing, 21(1):125-136, 2003.
[2] G. Lu, X. Li, H. Li. Research on Recognition for Facial Expression of Pain in Neonates. ACTA Optica Sinica, 28(11):2109-2114, 2008
[3] P.Ekman, W.V.Friesen. Facial Action Coding System (FACS). Consulting Psychologist Press, Inc.1978
[4] M. Grimm, D. Dastidar, K. Kroschel. Recognizing Emotion in Spontaneous Facial Expressions. International Conference on Intelligent Systems and Computing, 2006
[5] W. Liu, Z. Wang. Facial Expression Recognition Based on Fusion of Multiple Gabor Features. The 18th International Conference on Pattern Recognition, 2006
[6] G.Littlewort, M. Bartlett, I. Fasel, et.al. Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification. Neural Information Processing System, 2003

[7] B. Lee, J. Chun, P. Park. Classification of Facial Expression Using SVM for Emotion Care Service System. The 9th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing. 2008

[8] I. Kotsia, N. Nikolaidis, I. Pitas. Facial Expression Recognition in Video Using a Novel Multi-class Support Vector Machines Variant. ICASSP 2007

[9] P. Belhumeur, J.Hespanha,D.Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell, 19(7):711-720,1997

[10] H.Deng, J.Zhu,M.Lyu, I, King. Two-stage Multi-class AdaBoost for Facial Expression Recognition.Proceedings of International Joint Conference on Neural Network, 2007

[11] Y. Zhu, C. Silva, C.Ko. Using moment invariants and hmm in facial expression recognition. Pattern Recognition Letters, 23(1-3):83-91, 2002

[12] X.Mao,Y. Xue, Z. Li, K. Huang,S.Lv. Robust Facial Expression Recognition Based on RPCA and AdaBoost. WIAMIS 2009

[13] Z.Ying, X.Fang. Combining LBP and Adaboost for Facial Expression Recognition. ICSP 2008.

[14] S.Jung,D.Kim,K,An,M,Chung. Efficient Rectangle Feature Extraction for Real-time Facial Expression Recognition based on AdaBoost. International Conference on Intelligent Robots and Systems, 2005.

[15] Imed Bouchrika. Gait Analysis and Recognition for Automated Visual Surveillance. School of Electronics and Computer Science, University of Southampton, 2008

[16] S.Lajevardi, M.Lech. Facial Expression Recognition from Image Sequences Using Optimized Feature Selection.

[17] P. Rani, N.Sarkar, J. Adams. Anxiety-based affective communication for implicit human machine interaction. Advanced Engineering Informatics. 21(2007):323-334

[18] H.Kage,M.Seki,K.Sumi,K.Tanaka,K.Kyuma. Pattern Recognition for Video Surveillance and Physical Security. SICE Annual Conference 2007

[19] C.Huang, C.Wang. A GA-based feature selection and parameters optimization for support vector machines. Expert System with Applications, 31(2006):231:240

[20] L.Tang, Y.Zhou, J.Jiang, et.al. Radius Basis Function Network-Based Transform for a Nonlinear Support Vector Machine as Optimized by a Particle Swarm Optimization Algorithm with Application to QSAR Studies. J.Chen. Inf.Model, 47(2007):1438-1445

[21] F.Melgani,Y.Bazi. Classification of Electrocardiogram Signals with Support Vector Machines and Particle Swarm Optimization. IEEE Trans. On Information and Technology in Biomedicine, 12(5):667-677, 2008

[22] Z.Liu,C.Wang,S.Yi. A combination of modified particle swarm optimization algorithm and support vector machine for Pattern Recognition. The 3rd International Symposium on Intelligent Information Technology Application, 2009

[23] S.Romdhani, P.Torr, B.Acholkopf,A.Blake. Efficient face detection by a cascaded support vector machine expansion. Proceedings of the Royal Society, 2004

[24] Russel Eberhart, James Kennedy. A new optimizer using particle swarm theory. The sixth international symposium on micro machine and human science, 1995:39-43

[25] Yuhui Shi,R C Eberhart. Proceedings of IEEE International Conference on Evolutionary Computation, 1998,69-73

[26] R.T.Collins, A.J.Lipton, T.Kanade, et.al. A system for Video Surveillance and Monitoring. Technical Report CMU-RI-TR-00-12, Carnegie Mellon University, 2000

[27] Visual Surveillance System 1.0, An intelligent, turn-key, enterpriser-wide, visual surveillance system for any sized security installation, http://www.clarityvi.com

[28] Picard,R. Affective computing. Cambridge, MA: MIT Press, 1997

[29] C.Kim, S.Pang,M.Je. Constructing support vector machine ensemble. Pattern Recognition, 2005(36):2757-2767

[30] I.Buciu. Demonstrating the stability of support vector machines for classification. Signal Processing. 2006(86): 2364-2380.

[31] W.Jeevani. Performance Degradation in Boosting. In conf. MCS 2001:multiple classifier systems,11-21.

[32] J. Yang, V. Honvar. Feature subset selection using a genetic algorithm. IEEE Intelligent System and their Application, 13(2),44-49,1998

[33] Cheng Lung Huang, Chieh Jen Wang. A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications. 31(2006) 231-240

[34] Hsu, C.W., Chang,C.C., Lin,C.J.(2003). A pratical guide to support vector classification. Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[35] R.Wang., Z.Hu, L.Chen, J.Xiong. An Approach on Feature Selection and Parameters Optimization of Cascaded SVM with Particle Swarm Optimization Algorithm. The 3rd International Workshop on Computer Science and Engineering (WCSE), 2010

[36] James Kennedy and Russel Eberhart. Particle Swarm Optimization. Proceedings of IEEE international Conference on Neural Networks, IEEE Service Center, Piscataway, NJ, 1995

[37] Russel Eberhart, James Kennedy. A new optimizer using particle swarm theory. The sixth international symposium on micro machine and human science, 1995:39-43

[38] Yuhui Shi. Particel Swarm Optimization. IEEE Neural Network Society, February 2004

[39] Yuhui Shi, Russel Eberhart. A modified Particel Swarm Optimizer. Proceedings of the IEEE congress on Evolutionary Computation (CEC 1999:69-73, Piscataway NJ, 1999

[40] Li-Juan Tang, Yan-Ping Zhou et.al . Radius Basis Function Network-Based Transform for a Nonlinear Support Vector Machine as Optimized by a Particle Swarm Optimization Algorithm with Application to QSAR Studies[J]. J. Chem. Inf. Model. 2007,47,1438-1445

[41] J. Yang. An improved cascade SVM training algorithm with crossed feedbacks. Proceedings of the First International multi-symposiums on Computer and Computatioal Sciences, 2006

[42] Y. Ma, Xiaoqing Ding. Face Detection based on Cost-sensitive Support Vector Machines. Lecture Notes in Computer Science, Springer-Verlag Berlin, 2002

[43] D. Iakovidis, I. Flaounas, S. Karkanis, D. Maroulis. A cascading support vector machines system for gene expression data classification. Second IEEE international conference on intelligent system, june 2004

[44] G.C. Cawley, N. Talbot. Efficient leave-one-out cross validation of kernel Fisher discriminant classifiers. Pattern Recognition. Vol.36, no.11:2585-2592, 2003

# Design Approach to Improve *Kansei* Quality Based on *Kansei* Engineering

Nam-Gyu Kang

*Future University Hakodate,*
*Japan*

## 1. Introduction

In recent years, design has improved with development of manufacturing techniques. Only products that satisfy the consumer survive. *Karino* [1] has suggested the 3 types of quality based on relationship with physical fulfillment and individual satisfaction of designed objects; 1) *Must-be* quality, 2) *One-dimensional* quality, such as usability and operability and 3) *Attractive* quality, such as pleasantness, preference. *Attractive* quality is especially related to the user's potential needs because it is deeply related with the user's *Kansei* (emotional) satisfaction. Furthermore, Noman [2] has suggested 3 levels in design; visceral, behavioral, and reflective design engaging the appearance, efficiency and satisfaction (personal, memories etc) respectively.

Against this background, many approaches based on *Kansei* engineering have been conducted in Japan with the aim of offering more likeable designs. However, *Kansei* engineering does not have a long history. In a 1986 lecture at the University of Michigan the president of *Mazda* Motor Corporation introduced Professor *Nagamachi*'s car design process based on *Kansei* engineering. In the design process, perceptions of users' were analyzed statistically. *Kansei* engineering has since then been used worldwide leading to the development of the Japan Society of *Kansei* Engineering (JSKE). JSKE quantifies various characteristics of design to meet vast individual needs. In 2007 the Japanese Ministry of Economy accepted a 'Declaration for creating *Kansei* Value' as a national declaration further boosting public interest in *Kansei* quality [3]. These foregoing highlights the importance of *Kansei* quality in the future design. Clarifying the role and potential of *Kansei* engineering in design is crucial to the development of *Kansei* quality. Using *Kansei* engineering case studies this paper examines the role and potential of *Kansei* and *Kansei* quality.

## 2. *Kansei* and *Ksei* quality

The word *Kansei* has been used variously by researchers in relation not only to design but also to other research fields. It is therefore imperative to define *Kansei, Kansei quality* and subsequently, to address the relationship between the two in relation to the design process.

### 2.1 What is *Kansei* in design?

According to the Japanese dictionary, origin of *Kansei* is a German philosopher "Sinnlichkeit" of *Kant*. Elsewhere *Kansei* is interpreted as "the ability of sense", "the power of intuition" and "Sensibility".

According to many *Kansei* researches, *Kansei* is different from *Chisei* (This word is Japanese meaning intelligent), which works to increase knowledge or understanding by verbal description of logical facts also [4]. *Kansei* and *Chisei* are processed by the mind when information is received from the external world. *Tsuji* also refers to *Kansei* as the opposite of *Chisei* [5]. *Harada* [6] goes further to list three major characteristics of *Kansei* as; 1) human expression based on added knowledge and experience to inborn dispositions; 2) the ability to react to and evaluate external stimuli intuitively; and 3) the interaction of intuition and intelligent activity. P. Levy [7] defines *Kansei* as 'an internal process (or function) of the brain, involved in the construction of intuitive reaction to external stimuli'. For purposes of this paper therefore *Kansei* is defined as "the intuitive reaction to external stimuli based on one's past experiences".

In design, *Kansei* is understood as the important action of the heart to express imagery. *Kansei* works to evaluate an ambiguous feeling and impression of intuitive facts. The A design created through a designers' *Kansei* becomes a stimulus which the user psychologically evaluates. The designer must therefore understand his/her own *Kansei* as well as the user's *Kansei*.

### 2.2 What is *Kansei* quality in design?

As I have described previously, there are 3 types of qualities, *Must-be* quality, *One-dimentional* qulity and *Attractive* quality, in the designed object (Fig.1). Even if anyone evaluates the *Must-be* quality or *One-dimentional,* the evaluation results of those are almost same because those qualities were evaluated with the basic demand or the changing demand based on one's conscious standard. In contrast, the evaluation results of *Attractive* quality are different depend on the individual because the *Attractive* quality was evaluated with the potential demand based on the one's subconscious standard. The basic demand is the factor what is satisfied no matter what. The changing demand is depend on change of satisfaction level. However, the potential demand is factor to exceed the expectation of a user. In that case, almost of the users did not expectant or didn't even notice the potential demand.



Fig. 1. A reflection of physical fulfillment and user satisfaction from an object

According to *Shimaguchi*'s research, there are extensional meaning (=dictionary meaning) and intentional meaning (=implicative meaning) in a designed object. If anyone evaluates

the intentional meaning in the object, the results of evaluation are almost same. However, the evaluation results of intentional meaning are different depend on the individual. To synthesize the previous research, *Kansei* quality can be summarized as following; *Kansei quality is the intentional meaning of the object, it is evaluated intuitively with the tacit knowledge based on the human's past experiences.* This *Kansei* quality was related in human's potential demand. If we improve *Kansei* quality in design, we have to understand a users' tacit knowledge (they don't even notice.) based on their daily experiences.

## 3. Case study 1: *Kansei* library search system '*MegLook*'

There are several search systems in a library. However, almost of all of these systems use keywords such as author, title or genre for searching. These systems require previous knowledge of what is being searched for. In other words, these systems can be called *Chisei* (=intelligence) dependent systems. These library search (LS) systems are very useful for users such as adults who have previous specific knowledge about what they are searching for. However, these systems are not helpful for children who don't have previous specific knowledge. Sometimes a user lacks sufficient literal or specific information about a book they need. Moreover, in addition to keywords many library search systems re several require multiple step processes that can be time consuming and difficult.

The purpose of this study was to propose a new *Kansei* LS system for children based on behavior and to demonstrate the superiority of the *Kansei* LS system. The '*Kansei* system' in this case was defined as a system that can be operated intuitively with nothing but the tacit knowledge gained through one's past experiences. No special knowledge would be required to operate the system. Moreover, in the process of operating the system users would be able to experience positive changes of emotion such as 'comfort', 'pleasantness', 'a desire to use the system more' and so on [7].

In this chapter a *Kansei* quality LS system developed by our research team based on observation of child behavior is discussed.

### 3.1 Survey
### 3.1.1 Method of survey

Many researchers use human behavior as *Kansei* information that is rich and varied. There is even an 'Association for Behavior Analysis' in Japan [8]. It is therefore feasible to meet potential needs of users by observation of their behavior. In other words, observation of user behavior is very important for the development of *Kansei* quality in design.

### 3.1.2 Results of survey

There are two types of LS system in *Hakodate* City Library; one is for adults and the other is for children. However, if a child wants to search using the LS system for children, the child has to input the keywords such as author, title or genre etc. Many children cannot use the LS system without the previous specific knowledge about what they are searching for. Actually, there were few cases in which children successfully used the LS system in without assistance.

An analysis of behavioral patterns (Fig 2) revealed that the keyword search system was difficult for many users, especially children. Many children searched for and selected books based on ambiguous information such as graphical images from the front book cover, the book spine and perceptions from looking at 3 or 4, 5 pages.

Fig. 2. Analysis of behavioral patterns in selection of books at *Hakodate* Library

## 3.2 Proposal of Kansei LS system '*MegLook*'

Based on these results, we proposed a new *Kansei* Library search system called '*MegLook*' [9]. With the '*MegLook*' system, users can intuitively select by simply looking at the book spines. After making an initial selection, users can look at graphical images of the book cover and five pages (Fig. 3). Using this information, users can decide if they like the book. Using a touch screen, users can also save or print this information for future reference.
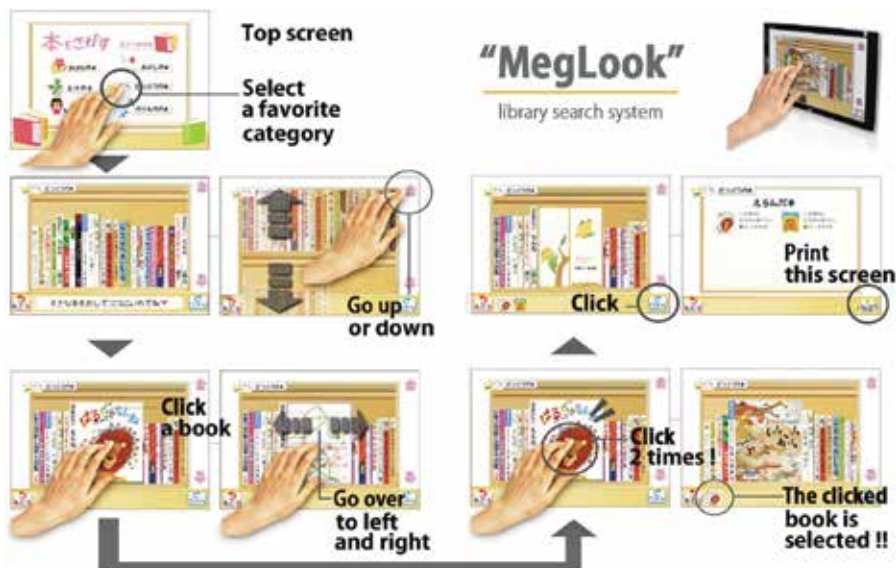


Fig. 3. The 'MegLook' Library search system based on *Kansei*

### 3.3 Evaluation experiment

The purpose of this experiment was to compare '*MegLook*' with the existing LS system. Special emphasis was given to *Kansei* quality of both systems.

### 3.3.1 Method of evaluation

Our field evaluations were conducted with two subject groups at the '*Hakodate* City Library'. There were 20 subjects in the child group (6 boys, 14 girls, average age 9.45 years old) and 10 subjects in the adult group (the children's parents). Subjects compared '*MegLook*' and the conventional LS system, using seven criteria: 1) Pleasantness (Pleasantness of experience), 2) Easy to understand (Ease of understanding system), 3) Easy to operate (Ease to operation), 4) Friendly (User-friendliness), 5) Preference (System preference), 6) Ease with which books could be found (Ease of search success) and 7) Wish to use the system more (User loyalty). They also wrote comments after their evaluations. We gave the subjects only three evaluation choices because most of them were children.

### 3.3.2 Results of evaluation

For each of the criteria *MegLook*' was preferable to the conventional LS system (Table 2). This was especially significant for criteria 1, 5 and 7. The results show that the subjects were more satisfied with '*MegLook*' than the existing LS system. The adults found Meg Look preferable with respect to criteria 1, 4, 5 and 7. However, the '*MegLook*' did not do well criteria 3. The adults even found the conventional system preferable in terms of criterion 6 (Fig 3). For the adult group, finding a book on the current LS system was rather easier. We feel this is because they already have the specific knowledge and keywords (author, title and genre) to know what they are looking for. Adults also do not *browse* the books like children do so '*MegLook*' does not greatly enhance their search. It must also be noted that '*MegLook*' system was designed mainly for children thus the lower performance on criterion 6 among adults was considered to be insignificant. The results were compared with with Karino's 3 types qualities; *Must-be* quality, *One-dimensional* quality and 3) *Attractive* quality.

'*MeguLook*' did not score highly among the adult group in *One-dimensional* quality such as usability and operability, but it scored very highly in terms of *Attractive* quality such as *pleasantness, friendliness and preference*. Results from the children's group were however different from the adults (Fig. 4). All seven of their evaluation scores for '*MegLook*' were higher than the conventional LS system. '*MegLook*' scored very highly in the *wish to use more, preference, friendliness* and *pleasant* categories. They also scored *easy to operate* and *easy to find a book* higher than the existing LS system. The results prove that children can use '*MegLook*' without specialized knowledge on how to operate the system. Overall '*MegLook*' was the preferred system for both groups.

|  | Child | | | Adult | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
|  | A | Neither A nor B | B | A | Neither A nor B | B | A | Neither A nor B | B |
| Pleasant | 1 | 2 | 17 | 0 | 0 | 10 | 1 | 2 | 27 |
| Easy to understand | 6 | 3 | 11 | 1 | 2 | 7 | 7 | 5 | 18 |
| Easy to operate | 2 | 4 | 14 | 2 | 4 | 4 | 4 | 8 | 18 |
| Friendly | 4 | 4 | 12 | 0 | 0 | 10 | 4 | 4 | 22 |
| Preference | 1 | 1 | 18 | 0 | 1 | 9 | 1 | 2 | 27 |
| Easy find of book | 3 | 5 | 12 | 5 | 1 | 4 | 8 | 6 | 16 |
| Wish to use more | 1 | 0 | 19 | 0 | 1 | 9 | 1 | 1 | 28 |

A: the existing LS system    B: MeguLook

Table 2. Results of evaluation
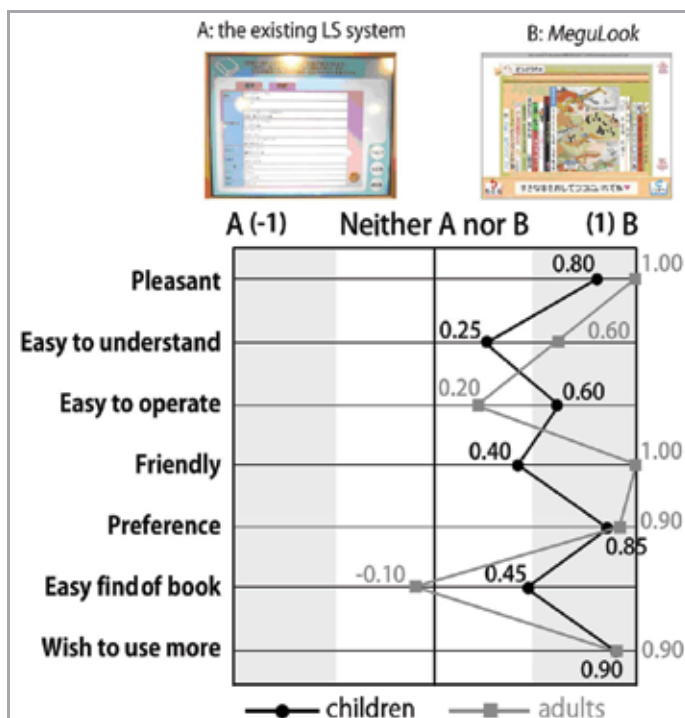
Fig. 4. Results of evaluation

### 3.4 Summary

We proposed '*MegLook*' based on behavioral analysis of how children search for and select books at a library. And through our field trials we proved '*MegLook*' is superior to existing LS systems. Since adults already have specific knowledge and are probably accustomed to the to the conventional LS system, their preferences for '*MegLook*' were not as high as those for the children's group. Even then, in terms of *Attractive* quality, '*MegLook*' was preferable for both groups. And in terms of one-dimension quality '*MegLook*' was preferable for the children's group. It can be concluded that '*MegLook*' was the preferable system that brought the most pleasure to the search experience.

## 4. Case study 2: Kansei medical information system '*mellonet*'

Recently, several proposals for the development of new medical environments for elderly patients through medical information systems (MI systems) have been made. However users often fail to adjust to the fast pace at which information technology products evolve. As a result many of proposed MI systems are of little use to elderly patients. Against this background, a new MI system, named '*mellonet*' was developed by our research team[10].

### 4.1 Proposal of '*mellonet*'

'*mellonet*' was proposed based on two observations; 1) usage of MI systems in a hospital by elderly patients and 2) behavior of elderly patients using special equipment during hospitalization. '*mellonet*' is operated using an intuitively operated simple touch screen or a

device similar to television remote control. Both of these tools were deemed easy for elderly users. *'mellonet'* helps to check patient's MI and to support communication between the elderly patient and the his/her caregiver or the medical staff. For easy accessibility, the system is placed at the bedside in a hospital or living room at home (Fig 5).



Fig. 5. *mellonet* system

## 4.2 Evaluation experiments
### 4.2.1 Method of evaluation

The purposes of this research were to visualize the characteristic of *Kansei* quality in our new MI system *'mellonet'* and to determine effectiveness of *'mellonet'*. The subject searched for medical information with '*mellonet*' and a conventional MI system. Two analyses were performed; 1) familiarity with information devices as a factor of age and 2) evaluation of *Kansei* quality in our new MI system *'mellonet'* by factor analysis using the SD method. The subjects were divided into 3 groups; 1) 20's group (19 subjects, average age: 20.8), 2) 30's to 40's group (8 subjects, average age: 42.5) and 3) 50's to 60's group (8 subjects, average age: 59.5). Each subject selected a familiar information device such as 'TV', 'internet device', 'cellular phone' and 'fixed-line phone'.

The subject then performed 15 tasks such as 'Checking for operation schedules', 'Watching the TV', 'looking up drugs' with '*mellonet*' and the existing MI system (Fig. 6). To improve reliability the order in which the systems were operated was random. Behavior of the subjects was recorded with a video camera for protocol analysis. Finally, the subject's perception of each system was evaluated using the SD method based 15 criteria selected from previous research [11] information system evaluation. A 5 level evaluation was used for each criterion.
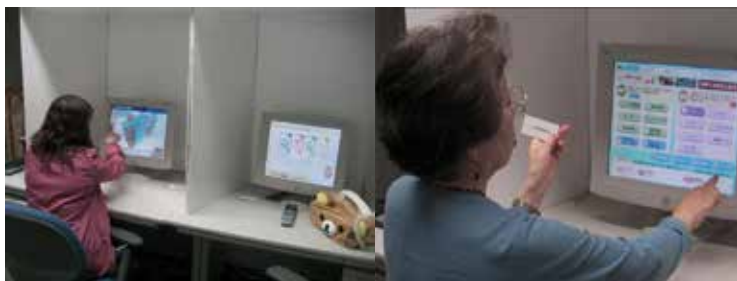
Fig. 6. Scenes from the experiment

## 4.2.2 Results of evaluation

From the survey of familiar devices, all subjects of the 50′ to 60′ group were more familiar with 'TV' and 'fixed-line phone', and fewer were familiar with 'internet device'. On the other hand, all subjects of the 20′ group were more familiar with 'internet device' and 'cellular phone'. Table 3 shows these results. From these results it was TV and TV remote control in were selected for use in '*mellonet*'.

|  | 20's | 30's-40's | 50's-60's | Average |
|---|---|---|---|---|
| Television | 84.2 | 100.0 | 100.0 | 94.7 |
| Internet device | 100.0 | 75.0 | 37.5 | 70.8 |
| Cellular phone | 100.0 | 75.0 | 62.5 | 79.2 |
| Fixed-line telephone | 15.8 | 100.0 | 100.0 | 71.9 |

Table 3. Adoption rate by age (Familiar device as a factor of age of subjects)

On behavior of subjects, protocol analysis revealed that average time for completion of tasks with the conventional MI increased with advancing age (F (2.32) =19.00, p< 0.01) (Table 4). However with '*mellonet*' average completion time per taks was shortened and there were no significant differences between age groups. Similarly the average number errors in operation increased with advancing age (F (2.32) =14.88, p< 0.01) when subjects used the conventional MI. However with '*mellonet*' there were fewer errors and no significant differences between age groups in the number of errors. The results indicate that even elderly patients can use '*mellonet*' intuitively without the special knowledge on how to operate the system.

| Subject | Average time for one operation (sec.) | | Average number of times of error in operation | |
|---|---|---|---|---|
|  | the exisitng MI system | *mellonet*' MI system | the exisitng MI system | *mellonet*' MI system |
| 20's | 20.73 | 14.48 | 0.67 | 0.03 |
| 30's to 40's | 28.53 | 16.22 | 1.75 | 0.25 |
| 50's to 60's | 34.99 | 19.63 | 2.5 | 0.00 |
| Variance analysis | F(2.32)=19.00, p<0.01 | n.s | F(2.32)=14.88, p<0.01 | n.s |

Table 4. Average taken time per task and average number of errors in operation

On the subject's perception of each system using SD method, '*mellonet*' was evaluated higher than the conventional MI system on most of the criteria (Figure 7). Significant differences

between subject perceptions of the two systems were observed for 8 criteria including convenience, unpleasantness, kindness, constraining, simplicity, gracefulness, harmony and complexity.
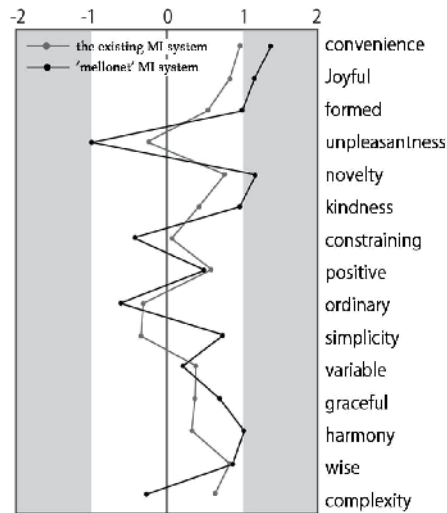


Fig. 7. Result of SD method

We conducted factor analysis using results from SD method. Table 5 shows the result of factor analysis. As a result of factor analysis, we could extract the 3 factors as comprehensive impression of both systems (table 5).

| Variable | Factor 1: Friendliness | Factor 2: Innovativeness | Factor 3: Sophisticatication |
|---|---|---|---|
| simplicity | -0.832 | 0.325 | 0.250 |
| difficult | 0.734 | -0.144 | -0.034 |
| free | -0.526 | -0.232 | 0.069 |
| unpleasantness | -0.542 | -0.287 | -0.206 |
| convenience | 0.510 | 0.198 | 0.233 |
| kindness | 0.491 | 0.004 | 0.321 |
| formed | 0.480 | -0.191 | 0.427 |
| original | 0.021 | -0.779 | -0.037 |
| variable | 0.063 | 0.777 | -0.486 |
| positive | -0.153 | 0.667 | 0.160 |
| novelty | 0.090 | 0.551 | 0.072 |
| joyful | 0.377 | 0.434 | 0.213 |
| awkward | 0.013 | -0.051 | 0.671 |
| wise | -0.004 | 0.297 | 0.581 |
| harmony | 0.425 | -0.118 | 0.444 |
| Eigenvalues | 3.987 | 2.976 | 3.698 |
| Variance explained (%) | 31.714 | 17.191 | 3.624 |
| Cummulative variance explained (%) | 31.714 | 48.905 | 52.529 |

Table 5. Results of factor analysis

The 3 factors analyzed were 'Friendliness', 'Innovativeness' and 'Sophistication' (table 5). The Figure 8 shows the scores of each factor for the two systems. These results show that many subjects, including the elderly, could operate the '*mellonet*' more quickly and easily than the conventional system. Moreover failure rate while using '*mellonet*' was lower than that while using the existing MI system. SD method results also showed that scores of "friendliness" and "refinement" were higher for the '*mellonet*' system. Moreover on analysis of the friendliness and sophistication factors shows that '*mellonet*' is significantly more friendly than the conventional system.
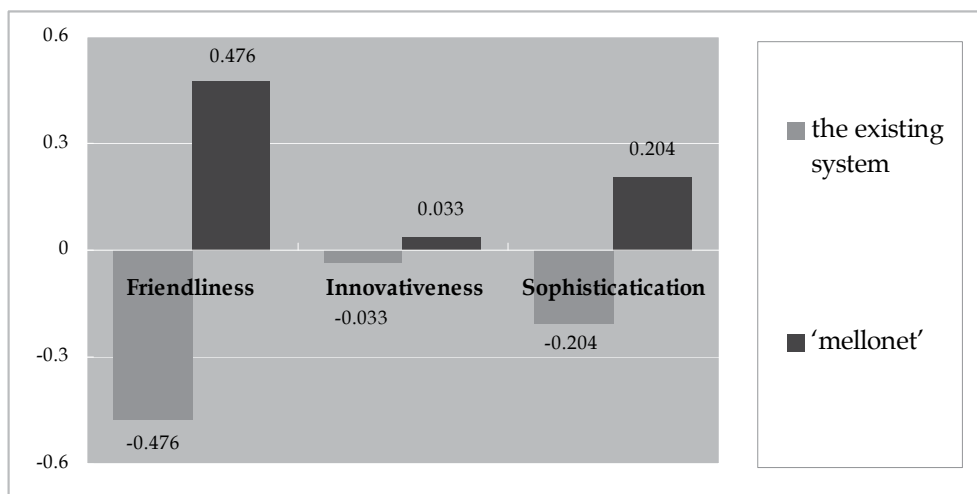


Fig. 8. Result of factor score of both systems.

### 4.3 Summary

From the results, we confirmed *mellonet*'s effectiveness and superiority as a *Kansei* MI System enabling even operation by elderly users through gained tacit knowledge based on daily experiences and without need for special knowledge. In other worlds, friendly characteristic of system is an indication of its *Kansei* quality.

## 5. Case study 3: *Kansei* photo browser system '*KanSya*'

Everyone has moving experiences resulting from stimuli such as movies, novels and so on. Such moving experiences can be described as *Kansei* (intuitive reaction to external stimuli based on one's past experiences). If we apply *Kansei* to photo browser system (PBS) design process, more users can have moving experiences. However, with the advent of digital photography and the resultant sheer volume of photographic data, browsing for a particular photograph can be difficult. One solution has been to add tags to the photos [4-6]. However, most tagging systems are based on the people in the photos, the places or dates. Unfortunately, such tagging methods do not allow *Kansei* based searches.

Therefore, this research focused on the construction of a PBS, with special attention paid to the emotive impressions from the photos. We propose creating *emotags (emotive + tags)* in addition to the tags already in use.

## 5.1 Survey
### 5.1.1 Method of survey
In order to design a PBS based on impression we must first define the characteristics of what one believes to be impressive. This survey sets out to clarify this by noting the details of the impression, the results of the impression (What made it impressive?) and emotions felt when recollecting these impressions, especially as to how they affected change. A questionnaire survey was conducted among 71 university students (40 men, 31 women). The survey method closely followed that of *Tokaji* [12]. With the help of three others, we categorized the results of the questionnaire survey using the KJ Method.

### 5.1.2 Results of survey
Most impact from photos (81.7%) was found to be from "Experience through others" (Table 6) rather than from "Self-Centered experiences." We concluded that people are impressed by experiences that are shared through and with other people.

| Experience Through Others | | Self-Centered Experience | |
|---|---|---|---|
| Friends · Family · Lovers | 18(25.4%) | Live Concert | 1(1.4%) |
| Club Activities | 31(43.7%) | Test | 4(5.6%) |
| Graduation | 4(5.6%) | Movie · TV | 3(4.2%) |
| People | 1(1.4%) | Nature | 2(2.8%) |
| Birth | 4(5.6%) | Music | 1(1.4%) |

Table 6. Results based on the KJ Method

The reasons for impressions can be seen in Table 7. Many of the subjects gave many different reasons for one impression, leading us to believe that it is possible to have various reasons for an impression. Having totaled the results, we found that 'Unexpectedness & Surprise', 'Achievement', 'Thoughtfulness & Love' and 'Strength of Memory' were the most common.

| Type | Subjects (%) | Details |
|---|---|---|
| Unexpected Surpised (26) | Friends / Family / Lovers 11 (42.3%) | Surprise birthday parties. Getting a farewell note from a friend when moving from Sapporo |
| | Club 7 (26.9%) | Teacher always being angry but praising me at the end... |
| Achievement (26) | Club 22 (84.6%) | Winning All-Hokkaido, Personal & Group, in Jr. High School Receiving big applause at concert. |
| Thoughtfulness Love (23) | Friend / Family / Lovers 10 (43.4%) | Receiving strength and courage from a friend. |
| | Club 6 (26.0%) | What people said after winning or losing a match or game. |
| Strength of Memory (22) | Club 9 (40.9%) | The Finals (sports) in high school |
| | Friend / Family / Lovers 8 (36.3%) | The feeling after a school play. |

Table 7. Reasons and Details of Impressions

The results of changes by impression were categorized to the three groups, 'motivation' (36.8%), 'outlook on others' (34.7%) and 'change in attitude' (28.5%). Within the group 'motivation', the criterion to try and/or try harder was the largest (20.8%). Within the group 'outlook on others', criteria of 'trust' (19.5%) and 'human love' (12.5%) were high. And finally, the criteria of 'broadened outlook' (11.1%) and 'change in thinking' (10.4%) were high in the group 'change in attitude'. We concluded that in terms of changes, these three groups are representative.

### 5.2 Proposal of PBS *'KanSya'*

The PBS '*KanSya*' [13] was proposed based on these results. The outline of the '*KanSya*' is shown figure 9.

1.  Tag adding photo ;
    The user browses photos and sets an attribute such as 'Surprise', 'Achievement', 'Love' and 'Strength of Thought' (1). By this operation, photos are added tags and compiled into a database.
2.  Photo browsing
    In photo browsing, the user enters the attribute of 'Surprise', 'Achievement', 'Love' and 'Strength of Thought' depending on their feelings (2). Then, the photos with the attribute similar to the entered one are displayed in sequence (3). By using this system, the user can browse photos using attributes such as 'moving experience by surprise', 'moving experience by achievement' for the many previous experiences.
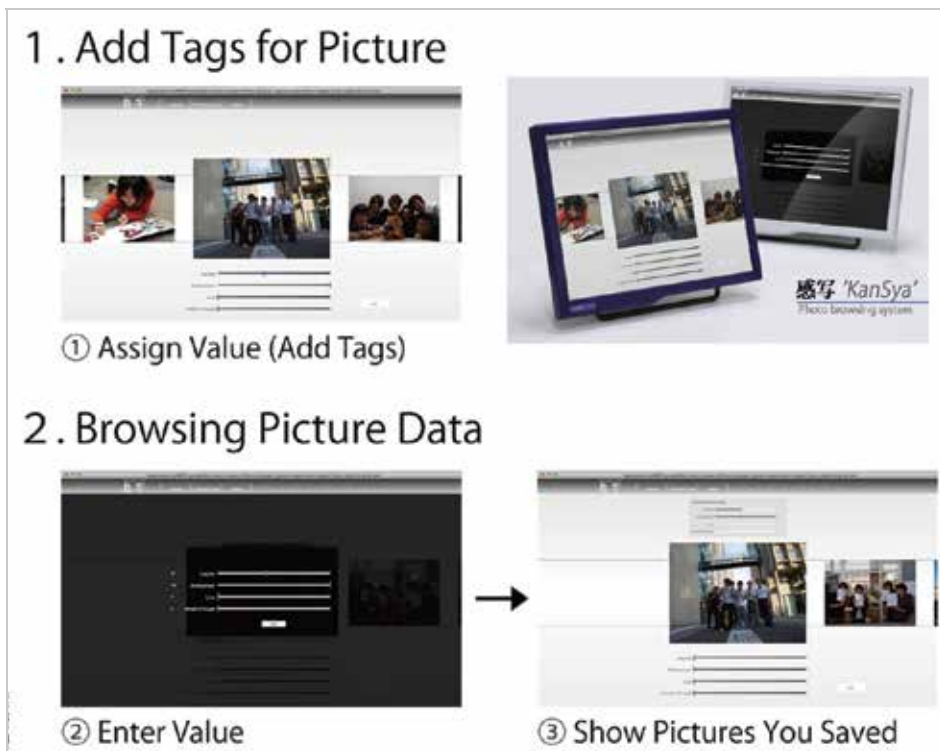


Fig. 9. PBS '*KanSya*'

### 5.3 Evaluation experiment
The purpose of this experiment is to determine whether the browsing with '*KanSya*' helped the impressed (moved) subjects more than browsing with a conventional PBS.

### 5.3.1 Method of evaluation
Each subject browsed the photos with '*KanSya*' and a conventional PBS (Mac OS X application 'Preview') to examine effectiveness of '*KanSya*'. The evaluation was divided it into two phases; 1) preparatory phase in which subjects set the attributes for selected personal photos and 2) main phase in which subjects browsed photos. A 3 week period was allowed between the two phases. The place was Future University Hakodate, and the subjects were 7male and 3 female university students.

During the first phase, each subject submitted 30 personal photos and set attributes like 'Surprise', 'Achievement', 'Love' and 'Strength of Thought'. The subject then answered a questionnaire about the difficulties in setting attributes ('Where you able to set the attributes easily?'). Alternatives answers included 'very easily', 'a little easily', 'neither easy nor difficult', 'a little difficult' and 'very difficult'. The subjects were allowed to freely give reasons for their answers.

In the main phase, subjects 1) browsed the photos with the conventional PBS or '*KanSya*' in random order. 2) answered a questionnaire, 3) browsed the photos with a PBS which was not used in the first phase and 4) answered another questionnaire. Question included 'How much were you moved by browsing photos with the two PBSs?' Answers were selected from among 'was very moved', 'was moved a little', 'neither 'moved' or 'not moved'', 'wasn't moved a little' and 'wasn't moved at all'. Another question was "feelings from browsing photos 'to which subjects were allowed to answer freely.

### 5.3.2 Results of evaluation
With the conventional PBS only 20% of the subjects felt 'moved' or 'moved a little'. However with '*KanSya*', the figure rose to 80% (Figure 10). The results indicate that '*KanSya*' enhances moving experiences more than the conventional PBS.
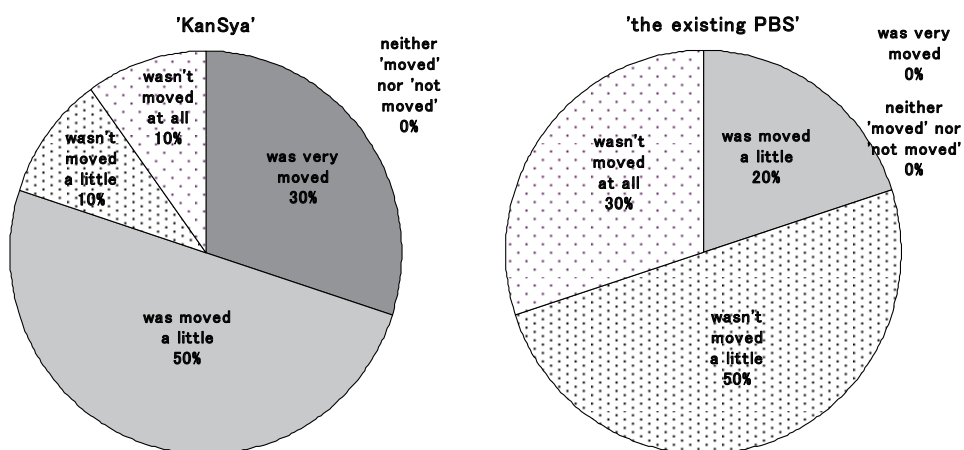


Fig. 10. Comparisons between '*KanSha*' and the existing PBS to evoke moving experiences

It is also notable that after the first phase, only 20% of the subjects found it difficult to set attributes for their photos (Figure 11). After browsing with the conventional PBS, there were many unspecific comments about 'nostalgia' such as "I felt nostalgic." And there were also feedback comments like "I could remember a past event, but I can't really remember what I thought or felt." However, browsing with *'KanSya'*, comments reconfirmed specific feelings such as "It was good for me to reconfirm what I love" and "It's fun for me to be able to get an opportunity to think of the feeling at that time" and so on. The results mean that photo browsing with *'KanSya'* gave a users an opportunity to experience some feelings again and notice the change.
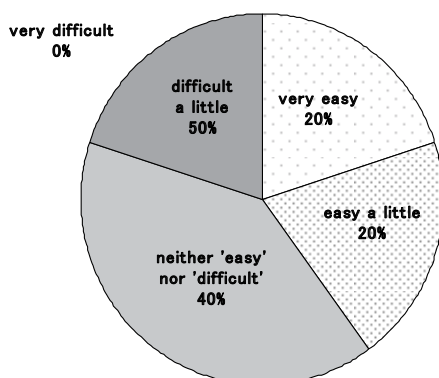


Fig. 11. How much smoothly the subjects set the parameter

### 5.4 Summary

This research identified the 4 key attributes that most people related to moving experiences. These are '**Surprise**', '**Achievement**', '**Love**' and '**Strength of Thought**'. Based on the results of our survey, we proposed *'KanSya'* which can enhance moving experiences through the use of *emotags* like 'Surprise', 'Achievement', 'Love' and 'Strength of Thought' depending on the user's feelings. Enhanced moving experiences from browsing with *'KanSya'* seemed to be related to the opportunity to reconfirm an emotion from a previous experience for oneself [14]. The results highlight the new *Kansei* qualities in the *'KanSya'* PBS.

## 6. Summary

The purpose of this paper was to examine the role and potential of Kansei and Kansei quality using Kansei engineering case studies, and introduced the 3 case studies which were approached to improve *Kansei* quality in system design. In the case studies, I conducted various evaluation methods such as factor analysis, SD method, behaviour protocol analysis, questionnaire investigation and observation of user's daily experiences to comprehend the users' various psychology evaluations to design. From those results, it is revealed that *Kansei* quality is related with evaluator or user' past experiences.

If we improve *Kansei* quality of a system, a user can use intuitively the system without the special knowledge about operating system. Moreover, the user will feel the positive emotion such as 'pleasant' and 'friendly', thereby the change of emotion affects up making a decision such as 'want to have it' or 'want to use more' etc. This *Kansei* quality was related in

human's potential demand. To improve *Kansei* quality in design, we have to understand a users' tacit knowledge based on daily experiences that they don't even notice.

The observing user's experiences is a good method to understand user's potential demand, these may be effective various the other methods. To visualize the characteristic of each method is in the future work.

## 7. References

Karino, N. (1984). Attractive quality and Must-be Quality, *Quality. The Japanes Sciety for Qulaity Control (JSQC)*, pp.39-48, ISSN 03868230

Noeman, A. D. (2004). *Emotional Design*, Basic Books, ISBN 0-465-05135-9, Cambridge, USA

Yamanaka, T & Pierre D. Levy (2009). Kansei Science and Kansei Value Creation through Kansei, Behavioral and Brain Sciences, COSMETIC STAGE, Vol.4, No.3, pp. 1-11, Tokyo, Japan

Tsuji, S. (1997). *Kansei Science*, SAIENSU, ISBN 4-7819-0828-4, Tokyo, Japan

Lee, S. (1999). A study of Design Approach by the Evaluation based on the Kansei Information, 'Images', doctoral dissertation of Tsukuba University, Tsukuba, Japan

Harada, A. (1998). The defibition of KANSEI, *Report of Modeling the Evaluation Structure of KANSEU 1998*, Evaluation of KANSEI 2, pp.49-56, ISBN 4-924843-33-4, Tsukuba, Japan

Pierre. D. L. & Ynmanaka, T. (2009). Design with event-related potentials: a Kansei information approach on CMC design, INDER SCIENCE, *International Journal of Product Development*, vol. 7 (1-2), pp.127-148

Yamaoka. T. (2008). HITTOSHOUHINWOUMU KSANSATSUKOUGAKU, Kyouritsu, ISBN 978-4320071698, Tokyo, Japan

Kang, N. & Nakaya, R. (2009). 2009 International Conference on Biometrics and Kansei Engineering (ICBAKE 2009), 978-0-7695-3692-7, 2009 IEEE, DOI 10.1109, pp.96-99

Kang, N. & Takamiya. K. (2009). A CONSTRUCTION AND *KANSEI* EVALUTION OF MEDICAL INFORMATION SYSTEM BASED ON EXPERIENCES, Journal of Japan Society of Kansei Engineering (JSKE), Vol.8, No.3, pp.489-498, Japan

Fukuda, T. (2006). *NINGENKOUGAKUGAIDO, Scientist*, ISBN 4-86079-014-6, Tokyo, Japan

Kana, S; Takayuki, I. & Satoshi, N. (2009), IPSJ SIG Technical Report, Information Processing Society of Japan, Available from
http://itolab.ito.is.ocha.ac.jp/~kana/image/SIGHCI2011.pdf

Tokaji A. (2001). Mechanisms for Evoking Emotional Responses of "Kandoh", *Cognitive Science*, Japanese Cognitive Science Society (JCSS), Vol.8, No.4, pp.360-368, ISSN 1341-7924

Niinomi, R. & Kang, N. (2010). RESEARCH ON EMOTAG PHOTO BROWSER SYSTEM BASED ON "IMPRESSION", INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2010 (KEER2010), pp.548-541, Paris, France

Niinomi, R. & Kang, N. (2010). Proposal and Evaluation of Photo Browser System to evoke Moving Experience, Proceedings of the 2nd international Service Innovation Design Conference (ISIDC 2010), pp. 217-222, Hakodate, Japan

# Part 5

# Biometrics Security

# Efficiency of Biometric Integration with Salt Value at an Enterprise Level and Data Centres

Bhargav Balakrishnan
*Sutherland Global Services*
*India*

## 1. Introduction

Biometric have been an effective tool in providing the authentication for the authorized user to access the resources of an organization. They have been widely used in data centres and at enterprise organizations as it require lot of security; those are termed as information security (confidentiality of data). Even then how the hackers are able to trace the network and break the passwords. Is there any weakness in the design of the operating system? Why the designer of the operating system have not come up with any tool that can provide better security for the servers. As we all know that securities are applied at different stages of an OS like at system boot, before login screen and the final password checkpoint at the logon screen. The network is designed in such a way that each stage from firewall till user web access is monitored, then how the hackers are able to trace the flaw. To avoid this happening especially loss of data can be prevented by including biometric at higher level of security. Biometric is one of the tools that provides authentication only for the registered/ authorized users of that respective server. Once if it joins with SALT value (randomly generated value of any length) which is nothing but the password of the authorized user and maps with the encrypted value to authorize the user access on to the server, the server level security goes high. Biometric have not been interfaced with SALT value yet and used for authentication of authorized user's at server level. Whenever the security is applied on the server level especially for Microsoft Servers, the complexity of the password alone is not sufficient as there has been lot of possible ways designed by the hackers to break that password. Here the biometric when included will not allow the hacker to penetrate as it (Biometric image) is unique for every user. There will be lot of FAQ's regarding this type of methodology for user authentication at server level.

When the authorized user gets hurt in his finger for example, how the server can be accessed?

Solution Here the application should be designed in such a way that it accepts maximum of two thumb impression. When it goes beyond this the user has to log on emergency mode in server by pressing f8, which can be accessed by means of a complex password with minimal access to applications. (Will be explained in depth in coming topics)

How does biometric image and password maps with Encrypted value store in the NTDS file of windows 2003 server?

Solution: -The authentication pattern is similar to any authentication methodology that is followed in NT authentication, mail servers etc... A slight modification will include a biometric image + SALT Value that automatically generate an encrypted value which maps with stored encrypted value. The encryption algorithm should be changed on regular basis accordingly the encrypted value corresponding to each user will change.

Here how the biometric is going to help?

Based on the image the value is generated even though SALT Value (here it is user's complex password) is known to stranger the respective system of a user can not logged with his (authorized user) thumb impression, that's where biometric provides security at enterprise level or data centre.

## 2. Biometric techniques

There are different biometric techniques and some of the commonly known techniques are as follows 1. Finger Print Technology is an impression of the friction ridges of all or any part of the finger. A friction ridge is a raised portion of the on the palmar (palm) or digits (fingers and toes) or plantar (sole) skin, consisting of one or more connected ridge units of friction ridge skin. These ridges are sometimes known as "dermal ridges" or "dermal" 2. Face Recognition Technology is an application of computer for automatically identifying or verifying a person from a digital image or a video frame from a video source. It is the most natural means of biometric identification. Facial recognition technologies have recently developed into two areas and they are Facial metric and Eigen faces 3.  IRIS Technology uses the iris of the eye which is colored area that surrounds the pupil. Iris patterns are unique and are obtained through video based image acquisition system. 4. Hand Geometry Technology include the estimation of length, width, thickness and surface area of the hand. Various method are used to measure the hands- Mechanical or optical principle 5. Retina Geometry Technology is based on the blood vessel pattern in the retina of the eye as the blood vessels at the back of the eye have a unique pattern, from eye to eye and person to person Retina is not directly visible and so a coherent infrared light source is necessary to illuminate the retina. The infrared energy is absorbed faster by blood vessels in the retina than by the surrounding tissue. The image of the retina blood vessel pattern is then analyzed 6. Speaker Recognition Technique focuses on the vocal characteristics that produce speech and not on the sound or the pronunciation of speech itself. The vocal characteristics depend on the dimensions of the vocal tract, mouth, nasal cavities and the other speech processing mechanism of the human body. It doesn't require any special and expensive hardware. The signature dynamics recognition is based on the dynamics of making the signature, rather than a direct comparison of the signature itself afterwards. The dynamics is measured as a means of the pressure, direction, acceleration and the length of the strokes, dynamics number of strokes and their duration. There are a lot of other biometric techniques like palm print, hand vein, DNA, thermal imaging, ear shape, body odour, keystrokes dynamics, fingernail bed. But these techniques are not been widely used in the authentication of the a person in attendance marking, server level authentication, authentication of a resident card holder as there are not feasible as the commonly used techniques which has been described above. As the authentication techniques should be feasible enough both in security and usability of the device. Based upon which only, the organization will accept for the implementation of Biometric authentication technique for their security purpose.

## 3. Evaluation on various biometric techniques

### 3.1 False Accept Rate (FAR) and False Match Rate (MAR)

The probability that the system incorrectly declares a successful match between the input pattern and a non matching pattern in the database is measured by the percent of invalid matches. These systems are critical since they are commonly used to forbid certain actions by disallowed people.

### 3.2 False Reject Rate (FRR) or False Non-Match Rate (FNMR)

The probability that the system incorrectly declares failure of match between the input pattern and the matching template in the database is measured by the percent of valid inputs being rejected. This happens in some of the biometric authentication technique as it will give a negative result when the log is generated as the image it has authenticated is different which will be considered as a negative parameter.

### 3.3 Relative Operating Characteristic (ROC)

In general, the matching algorithm performs a decision using some parameters (e.g. a threshold). In biometric systems the FAR and FRR can typically be traded off against each other by changing those parameters. The ROC plot is obtained by graphing the values of FAR and FRR, changing the variables implicitly. A common variation is the Detection Error Trade-off (DET), which is obtained using normal deviate scales on both axes. This more linear graph illuminates the differences for higher performances (rarer errors).

### 3.4 Equal Error Rate (EER)

The rates at which both accept and reject errors are equal.ROC or DET plotting is used because how FAR and FRR can be changed, is shown clearly. When quick comparison of two systems is required, the ERR is commonly used. Obtained from the ROC plot by taking the point where FAR and FRR have the same value. The lower the EER, the more accurate the system is considered to be.

### 3.5 Failure to Enrol Rate (FTE or FER)

The percentage of data input is considered invalid and fails to input into the system. Failure to enroll happens when the data obtained by the sensor are considered invalid or of poor quality.

### 3.6 Failure to Capture Rate (FTC)

Within automatic systems, the probability that the system fails to detect a biometric characteristic when presented correctly is generally treated as FTC.

### 3.7 Template capacity

It is defined as the maximum number of sets of data which can be input in to the system.

## 4. Basic setup of enterprise level security

As we see from the above diagram the security that is applied at each stage of a network. Even after applying these securities how the hackers are able to penetrate through the

network and able to steal the confidential data's of many user's. When the user is accessing his bank account through net banking or when he trying to do a transaction of money over a network all that is required is security for his password and his account. Even then the hackers are able to get the username but getting a password is what his challenge is with which he can manipulate anything on the customer's account. A lot of these things are happening in today's present scenario. But how to secure these kinds of flaws both at a server level as well as at a user level is what is going to be discussed in depth in this chapter and the methodology that is going to be used to prevent this using the biometric and salt value along with the encryption algorithm. The biometric can't be used at a wide level at a Net banking as every user will not have a laptop or can't get biometric devices separately. In order to apply even that at an enterprise server level, ho to do that is what is going to be discussed in this methodology of server and application authentication at an enterprise level.
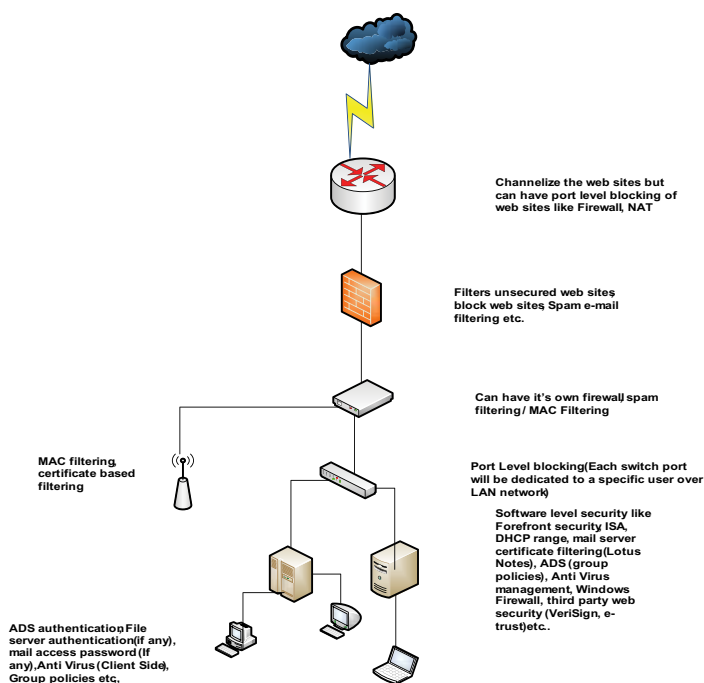


Fig. 1. Security applied at each stage of a network

Each stage has its own encryption algorithm but having something included unique within an Encryption algorithm is what to make the data centres to have their information's keep even more secured. Each application has an encryption algorithm right from Cisco routers but they are also broken in many ways them the hackers are able to get the IP address of the internal network by some means. Even some organizations allow users at higher level executive to use Pen Drives on their official computers. If the antivirus installed on the computer is not effective then the virus /spam that has affected the other computer can penetrate into the network and can affect many other computers over the network. So what happened to the security of the information's that are stored on the server? The main

drawback comes here is the server authentication are maintained with just the password and the encryption that comes with the server OS alone. But even though the password is set complex it is easy for hackers to reset the password. There are lot of encryption algorithm in today's world which are making the process of breaking up the security password. Even after using a lot of network monitoring many organizations are facing this issue. How to resolve these kinds of issues at Servers at enterprise level is the place the biometric and the salt value is going to play a vital role. As the biometric images are unique as we all know and can also provide a better level of security with both the SALT value and the encryption algorithm.

## 4.1 Parameters of biometric technique at security level
### 4.1.1 Permutation and combination
Why we have to choose permutation and combination while applying biometric at enterprise level? The main reason is to have a redundancy when there is a user gets hurt then what will be an alternate option. When we take the eye the possibility of generating a biometric image from a person will be two and it mainly depends on the characteristic of the light behind it that is the brightness. When there is some slight variation in the light that is generating this image can cause the authorized users from accessing and the probability that can be tried in this approach is also less. There are certain concepts like Voice, finger print where the probability becomes wider. The other techniques are also effective but each as specific criteria to bring that at complete enterprise will violate the security norm as well as it will be taken into account that is what we call as "Risk management". The live server are always are handled with a lot of risk and security measure taken for it will high. Let me explain you how this permutation combination concept is going to work.

For example, if I am going to be an authorized network engineer at an organization and have been given the permission to change certain things on the server regarding to the network and it's security monitoring. I have generated a biometric image with my fingers say 2 fingers from the left hand and 2 from the right hand. Now on that I have got fracture in my right hand. So the possibility of generating biometric images using the two fingers is there in the left hand. So the combinations that are accepted by the system is high and it becomes flexible for the authorized user to operate on the server and also secured as the images are unique to each person. Only the registered users along with their password (SALT value) and encryption algorithm that is getting generated internally once after accepting the biometric image and password of a user is going to map with the encryption table. The encryption algorithm can be varied on a weekly basis to ensure that the encrypted value are manipulated periodically to ensure high level of security at the server level as that is like the heart of an organization and the stage above it are like a wall or barriers for the hackers. Certain server can have an authentication from couple of biometric generated by the same person which will converted into the respective formats using any mathematical approaches and it is going to be discussed in topic where the generation of encryption is going to be done.

$$X1/X2 + SALT\ VALUE\ (Y1) = W1+Z1 = Final\ encrypted\ value\ (E1) \qquad (1)$$

Here X1/X2 are the biometric images in which either of them can be used. But in which biometric is this combination are more is in very less techniques. Then the ones that are having more probability will be the finger print and the voice. But the voice also has a

specific drawback.. When I generate a voice encryption the application should filter the unwanted sounds that come apart from the voice of the authorized user then the probability of using the voice in authentication techniques will be high. As the voice is having a lot of combination like the finger print and can be converted into different format before it comes with a different format of password (SALT value) thereby providing a highly security approach of security. The main thing that should be joined with voice is the filtering of the unwanted sound from the background every time right from the registration of authorized user on a server at enterprise level. There are a lot of combinations that needs to be taken into account as the server needs to be accessed regularly so the technique that can process easier will be voice and finger print. Let us look into the other techniques and according to the priotization, reliability, usability and feasibility the biometric techniques will be utilized but having a common will make the process of authentication easier. Let us see a brief description on the parameters that are going to play an important role in the implementation of sever authentication technique at an enterprise level using the biometric and SALT value as a source of generating the authentication code. Then the code is going to map with the encryption process for authenticating an authorized user.

### 4.1.1.1 Priotization

The servers at enterprise will be undergoing monitoring at regular intervals and accessing the servers for various purposes will be high. Certain servers will be accessed at specific intervals like data base server, web server, net banking, ATM servers etc... Like at the end of the day to generate the complete report on the transaction and they are accessed only by certain authorized users who are technical specialist on that application and also who can generate the end of the day report as the data's which are seem as highly confidential like user's account number, Pin numbers, account details which are normally kept highly secured for which this biometric approach of authentication will make it highly secured. In this sector the biometric authentication type should be highly secured and feasible. So in these sectors the highly recommended approach with finger prints and then comes voice recognition. Why these approaches are feasible in this section? The main reason the probability of generating the finger print image is more than the other biometric methodology. When the authorized user needs to access, there is no requirement of other criteria's like brightness of the room, the voice filtering, the position etc. The finger print is quite a simple approach of biometric and also gives high security for the authentication. The best example will be the Yahoo Mail where Yahoo has got finger print approach for accessing the e-mails. The other methodologies of biometric generation are also having the advantage over authentication but it is the division where we use them. The biometric image once generated should also be stored secured and then the Salt value generation should be random. Every day the SALT value should be different and it should get updated to the authorized user. The device that is used is generate the SALT Value every 60 seconds are been manufactured by EMC2. The current models of this SALT value device are RSA SecurID 900, RSA SecurID 700, RSA SecurID 800, RSA SecurID 200, and RSA SecurID 520. These are some of the device that are being widely used in today's enterprise where the security is give the most important priority when compared to other parameters of an organization policies.

### 4.1.1.2 Reliability on biometric techniques

Biometric is highly reliable when it comes to information security. How it is going to be a feasible approach when it comes to authentication at enterprise level? What are the things

that needs to be customized in server OS especially Windows, Solaris? As customizing the server should be after getting the necessary approval from the OS developer and the license provider that is Microsoft/Sun whichever is going to be customized according to these authentication requirements. When the necessary approvals are processed and this customized OS needs to be approved by security norms designing bodies like ISO that this approach of server authentication can be practise. Once this is accepted then the methodology can be widely used in the enterprise level. So we are going to see the areas of justifying this methodology that is going to tell "HOW RELIABLE IS THIS METHODOLOGY?" 1. When this approach is implemented the possibility for the hacker to steal information becomes less as both the Biometric value and the SALT Value is going to be unique. Once these two numbers are going to joined as 0'1 and 1's using any calculation like XorY, Xand Y, XnorY etc...Then the number formatting is changed and then when encrypted using an encryption algorithm the output will be completely and doesn't gives even a clue on what number or image is used. Even if the number is able to decrypted getting the same biometric image is hardly possible 2.The risk that is involved in maintaining these biometric image are high but there is a modification that is done for avoiding this risk and in a secured approach which will discussed in depth in the topic that is going to give a complete explanation of this biometric technique 3. The reliability on this biometric approach of authenticating server access can be high as the Biometric technique that is chosen is based on the maximum combination not with the least combination where it can be of a risk and the management won't agree for that approach. All that management requires from its point of view is an application that can be feasible and at the same time keep the information's of the organization and the client safe and secured. This methodology will be highly as it is going to be an integration of know approach but encrypted and combined in a different approach which has the capability of getting compatible at server level more easily with high reliability 4. The finger print that has been know for many centuries as a source of authentication but it was ink based that is pressed on a paper by an individual when there is an election to avoid misusing the policies of making another vote by the same candidate. So this gave a unique approach which was slowly being used for the authentication purpose in the country's visa card to authenticate a resident expatriate. So the finger print has been widely used in various application and sectors. That's has been reliable and feasible in authenticating a user 5. The finger print methodology has requires simple enhancement over the existing keyboard. The keyboard needs to be interface with the finger print reader which should have a driver that should be getting installed automatically when it is interfaced with the port on the server. That should have an application that should have a transfer the image to the application that should to in turn go through the entire process of authentication which will be explained in the section where we are going to discuss on the complete process involved in this authentication 6. The finger print is a technique that can be used in this scenario. Let us analyze on other techniques also and bring a complete analysis work on that area of work. This will be a valid justification on the exact priority of the biometric techniques 7. The biometric techniques are more reliable when compared to third-party software requesting to remember an image as a source of authentication for the users over the Net Banking. This needs to be a part of risk when the management point of view especially when this kind of methodology is applied at Net Banking concepts 8. Now the bank ATM card systems are slowly bringing up this technology apart from the PIN system as it will make the process of authentication much easier and with high security.

### 4.1.1.3 Usability of the biometric techniques

When it comes to the usability of the biometric devices, it has been simple as the installation is done by the infrastructure team along with the maintenance. The biometric has been really user friendly in terms of registering their image like finger, eye, palm etc... But when you to login using it there has been a fault tolerant that sometime if the brightness of the eye was not equivalent to the brightness that was there during registering it might not accept and this sometime makes the user to use the password to login into the computer. But when it is going to be designed at a server level it should not be the considered as a negative parameter. Here the biometric that is going to be should be highly advanced in authenticating the user in a much quicker way. It should try to filter the brightness and auto adjust itself so that it should only take the exact picture of the authorized and not the brightness behind the picture. Apart from the finger print other biometric techniques needs filtering. This will increase the options of using the finger print in the biometric techniques. Transaction authorization: A subject can execute a transaction only if the transaction is authorized for the subject's active role. With rules 1 and 2, this rule ensures that users can execute only transactions for which they are authorized. Here the sensor is the device that is going to identify a authorized user's biometric image. When a user comes before it or swipes through the device it will take the image then it will it will go to pre process, the image is will be converted to the required format as designed in the parameters of the features followed by generate the template (Generate biometric template customized based on feature parameter). In the pre-processing, it has to remove artifacts from the sensor, to enhance the input (e.g. removing background noise), to use some kind of normalization, etc.Then it will be stored in the database of the Biometric device. Then again when the swipes, it will go through the process of customizing then goes to matcher and then check it matches with the one stored in the Biometric database. Here what type of conversion is being used?  The algorithm that is used is Matching algorithm. The matching program will analyze the template with the input. This will then be output for any specified use or purpose (e.g. entrance in a restricted area). So the probability of this algorithm securing the biometric image when compared to the biometric image into 0's and 1's will be discussed in the later topics. As the image when converted to 0's and 1's either by binary, octal, hexadecimal. It is then applied to digital conversion like 4B/5B, 6B/8B format then converted to the number makes a rearrangement of bits and it will be of high security when applied at Enterprise level. This makes the complete process of the Biometric authentication process. This diagram will be common for all the biometric but the encryption algorithm and approach of Biometric authentication varies a bit. In some device it will take only a biometric image for authentication like laptop, resident card authentication. But when you take for entrance security it has biometric image with a key to authenticate a user for his attendance. But how effectively they are used is comparatively less as the users finds it tedious with the work pressure they have and this process is mostly ignored in many places. Normally they have password authentication for access the organization or card system which is swiped and mark the attendance. But how far the card system has been effective is very less when compared to Biometric authentication. Even in major bank it has not been implemented. The Biometric is not implemented at entrance, locker section and the server room where all the confidential data's are stored.
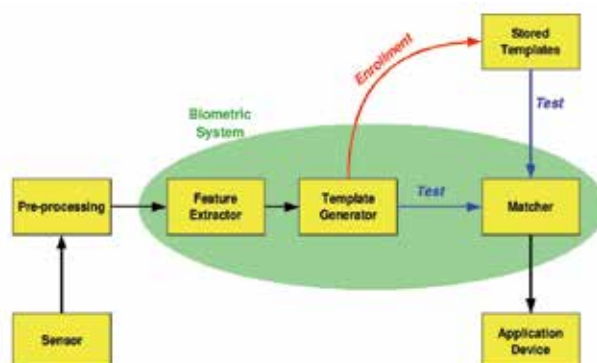
Fig. 2. The basic block diagram of a biometric system

### 4.1.1.4 Feasibility of the biometric authentication

The application designed for the biometric authentication has been highly feasible as it is just to store and authenticate the authorized user when he accesses a security location. The authentication should have a proper a backup and restore system as to a source of redundancy if the device gets damaged or the image template gets corrupted. As there will be fault tolerance in any software as it doesn't have a specific reason for it to get corrupted. So that is the only that need to be really careful. As the authentication devices are highly feasible bu the damage of the device depends on the life factor that is quoted for that device. So the backup of the authentication template has to be taken on a regular basis along with the report logs of authentication which forms a part of security auditing. When the auditing is done for giving an organization with the ISO certification authentication process is a part of it. So this biometric authentication should be proper and it should be approved approach so that the organization can be secured in keeping their data's and client information with high security. The feasibility depend on many factors like change management, updation of firmware, risk management. The feasibility as we all know is classified as economical feasibility, technical feasibility and operational feasibility. When all these conditions are satisfied then only the using a particular biometric technique will be approved in a enterprise organization. When the biometric is used at enterprise it should be reliable, quick on authentication and price for the installation should be reliable. But when it comes to server authentication all matters is Information security which is of high priority than all those feasibility of biometric authentication. Time is not major constraint when compared to Security. The security level is considered and the analysis of that which will seen in the coming topics. The Information security have been the major constraint and for which security at the network and server level is increased periodically to ensure that the data when it is transmitted over the network are not been easily decrypted by the hackers. It is a real challenge for the people at the security domain and auditing vertical of security. As the hackers are working very hard to track and try to create a lot of problems. But how this biometric is going to help in this process is going to being discussed and also later works on integrating biometric with confidential data's during transmission of data over the network. This will ensure that the data's are safe in both Inter ad Intra network locations. This type of strategic approach is much needed for this security level. Without having a proper approach towards the parameters which are mentioned selection of a biometric migh go wrong. So follow designing by means of the above parameters.

## 5. Proposed biometric techniques

This is the proposed biometric process flow for authentication at sever level mainly we call as "Enterprise level support" where huge data's of customers, client, companies confidential data's are stored. At this level generation of report log is mandatory which is going to be generated at the EOD. Along with the authentication at the network level should be monitored and should be generated that is going to form the consolidated report for the day. These complete consolidation of data's at the end of the year is going to presented for the auditing based upon which the security policy of the company can be seen, Many proposed model of security are available which is customized as per the companies and used for the generation of the security audit reports. Here at registration process the Biometric sensor is going to be getting the biometric image from the authorized user and then going to ask for the SALT value. Both of them are converted to the respective binary format, then going to perform the Logic gate operation which will rearrange the arrangement of bits.
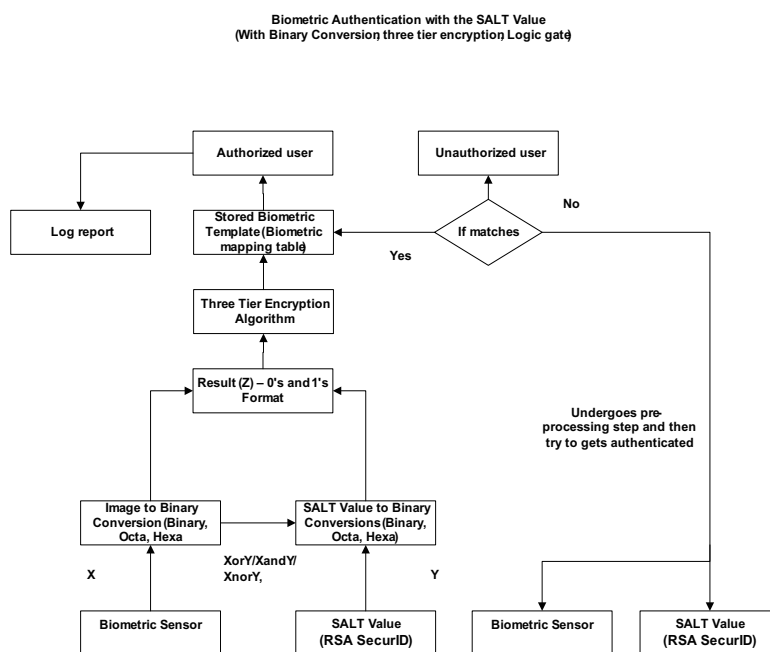


Fig. 3. Biometric authentication with SALT Value for server level authentication

Then at the Three Tier encryption algorithm the encryption is going to happen which is then stored as the generated template of the biometric image. Once the template is generated when the user logs in based on the biometric image and the SALT value the calculation automatically map with the generated value. Which will then going to authenticate the user? Here comes the question how the saved template is going to authenticate when the saved SALT value keeps changing. The application is also simultaneously within the server as the same SALT value which is there with the authorized. So the encryption table will have generated template value updated accordingly and it will easily authenticate the user. Then where is the security. Here the entire uniqueness of accessing is the biometric image

which cannot be easily generated. Apart from this SALT value will be in an encrypted format so tracing out the possible value will take many years for the hacker to break through. But they will never be able to manipulate any biometric image of an authorized user. That is the most important part of this methodology. Here the main intention was to prevent the unwanted access of servers by unauthorized users for checking the information's on the server without clear information. At the same the authorized user is not have to write the password and keep anywhere as it is his biometric image and SALT value that is randomly generated. The authorized will have to be careful with the SALT generating device. The administrator should be set with a lot of policies that for the security of the confidential data's. In the above you can see that the report log that needs to be generated it is something similar to the System log file which comes on the authorized who has accessed that particular server. The report log is generated for all the servers that are on the network of that organization. This report needs to be consolidated to get the final EOD report. It will give a clear picture of the user access control for the servers. If there is any loss of data's it is easy to trace with the authorized user log report. As there will be unmatched/ invalid user report that would be generated for the unauthorized user. This will make the process of tracking the unauthorized users simple. This is the main advantage behind this methodology. The IT infrastructure manager should be careful on the report log generation and should ensure that the LOG reports are generated at the EOD is properly consolidated for all the server and network infrastructure.

This methodology is going to be an enhancement in the current OS and it is going to be integrated with the necessary approvals from the OS developers like Microsoft, Linux, Solaris etc.. Once this approval is done then the customization of the OS can be done and the testing at the security level, feasibility level (mainly restoral level), reliability are going to studied in depth before executing at the Enterprise level. Here in this chapter this analysis is going to be seen in the coming topics. Why the format of 0's and 1's are being used and with a lot of rearrangement of bits. This is because when the hacker is having the ability for finding the number but the number format of 0's and 1's are quite difficult as it will be quite unique. So the hackers will not be able to identify that the exact that is being carried out with 0's and 1's. Let us see some of the analysis with the rearrangement of the bits and without it the security level will be comparatively less.

## 6. Generating Three Tier Encryption algorithm

The steps of using this encryption methodology are as follows step 1. In the first step the RSA algorithm will be carried with the following modifications a. Consider two prime numbers as 11 and 13 b. N= P *Q i.e.   143 c. M = (P-1) (Q-1) i.e. 120 d. D is the decryption key Example 3 which is a prime number e. E = D inverse (mod n)   i.e. 47 f. Let the password be "Hello "take the ASCII value of the password covert it as 7269767679 g. Concatenate this ASCII value with a SALT value (Randomly generated number) say 34 i.e. 247172101086 h. Finally multiply this with the Encryption value to get final encrypted word 9886884043440 There are certain constrain which are modified and the requirement in RSA Algorithm are as follows  a. The minimum requirement for P and Q values in RSA is 2048 bits which gives the utmost security to the file that is being transferred b. Modification is inclusion of ASCII value conversion and SALT Value. Here SALT is being left user defined c. The P and Q values are also user defined that is also a modification d. At this you can use any encryption algorithms which are being updated. Step 2.The above arrived result through RSA –

9886884043440 will be converted into 0's and 1's using number conversion. The above encrypted data (9886884043440) will be converted as 1001011011011100100100100101. This is for binary in the same way it can be done for octal /hexadecimal Step 3.This number conversion will be modified using Digital Encoding (Either Line or Block Encoding). Advantage: - Rearranges the bits of data i.e. 0's and 1's. Then use any of the line encoding schemes like NRZ, NRZ-I, RZ, biphase (Manchester, and differential Manchester), AMI and pseudo ternary, 2B /IQ, 8B/6T, and 4d –PAMS and MLTS that will convert the number which are being as binary in the above as follows Let us consider 4B/5B Block Encoding for the replacement of bits that were generated using the binary conversion. The output that would be generated by using the reverse conversion process with be different from the generated using the RSA algorithm. The output that is generated us mentioned below.

10011011101101111010101100101010110100111011111110100110111011011110101011001010100110111011011110101011001010101101001110111111101001101110110111101010110010101011010011101111111010011011101101111010101100101010110100111011111110



Fig. 4. 4B/5B Substitution Block Encoding

Step 4.In this step the conversions of data into 4B/5B will be converted back into numbers using number conversions. This is reverse process of Step 2 Conversion back to binary will be give different encrypted word because of the usage of 4B/5B line encoding. The solution will be 10205099. Here also the conversion can be any one of the following binary/octal/hexadecimal. Step 5.In this step the above obtained number in step 4 10205099 will be considered as the X Value. This will be substituted in the Mathematical Series. Here in the example the sine series is being used. Formulae: - $\sin(x) = X - X^3/3! + X^5/5!$ - Here X is the encrypted value 10205099. The series is used defined say N=3 then the series will be till $X^7/7!$ Then the final result will be 10205099 - 1771333826010.833333+ 246018586945945274.37 = 176887364014124447176.45856481478. Use the round off function to get the final encrypted word as 176887364014124447176.

## 6.1 Advantage of using digital encoding, number conversion and mathematical series
The main advantage of Step 2, 3, 4 is in Step 2 the encrypted data obtained by RSA is converted into 0's and 1's. Then by using Digital Encoding the rearrangement of Bit's are done. Finally in Step 4 the reverse process of number conversion. What it does? The hacker will never get a clue of this process that is being carried unless he gets an idea about this

algorithm. Then Step 5 also a vital role as here the number X i.e. the value obtained from Step 4 has to be determined by the hacker, for which he should what is used, if found what mathematical series used which will takes ages to refine.

But for an organisation to encrypt and decrypt will be a simple as the process involved in each data encryption will be stored in their database. So this twist in the algorithm will be playing the most important in preventing the hacking of data's. How this methodology gives utmost security to the file at the same time increases the complexity in identifying the content by the intruder. These are being described below If the Intruder gets this encrypted word the following things are to be determined. Determining those values is a long process and finding those will take many years in order to arrive at the conclusion 1. The value of N i.e. the length of the series has to be determined 2. After finding N values the value of X has to be determined that has been substituted in the series 3. In the line encoding process the split up of the bits has to be determined like 4 bits, 8 bits and so on 4. After determining this, the type of encoding has to be determined and the substitution used as in the B8SZ where 8 bit value is substituted in place of continuous 8 zero's 5. Based upon which the entire two stages can be revealed from this the first stage can be proceeded that is RSA instead of that AES, SHA, MD5 any encryption algorithm can be used 6. The speciality of RSA is in determining the prime numbers P and Q which itself will take many years to determine.

The end user can be a data center, search engine etc which will get utmost security because of the usage of Line Encoding and Mathematical series. The line coding will convert the original encrypted word into duplicate encrypted word by using the following) i) binary/octal/hexadecimal the encrypted word is converted as 0's and 1's ii) then line encodings is used. This will act as a protection. This will be even more protective by using the mathematical series. On the whole the methodology will be a secure path for the transfer of data's. Time for generating the Encrypted file using this method will be comparatively less in the high end PC's with dual core processor and above with 2GB RAM with processor speed of 2.2 GHz. The RSA encryption of about 2048 bits will take time other steps will take fraction of seconds for generating the desired output.
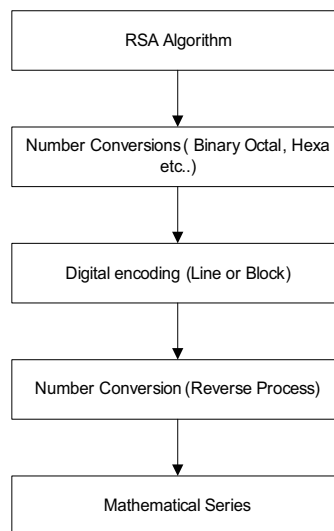


Fig. 5. Diagrammatic representation of entire encryption process

This will give the complete idea on this encryption algorithm flow. Here the important step is in the replacement of bits as that is making the complete change in final encrypted result.

## 6.2 Advantage of this encryption methodology

There are various advantages of this encryption methodology which are as follows 1. In the file transfer preferably in the low privilege servers which are an endangered place of hackers 2. In the WAN where the data transfer is not that secured, in order to give a firm security this methodology can be adopted 3. This methodology will be of high value in the defence sector where security is given high preference. Using this methodology the hacker will not be able to trace the ideas unless or until he is well versed in the mathematical and electrical technique of disclosing the data 4. This will also play a vital role in other sectors like Bank, IT, Aero Space and many more where the data transfer is given more security. These are some of the advantage of this encryption in secured file transfer over the low privileged and it will be to secure the server at the same level of security.

## 7. Proposed server authentications (complete analysis)

This is the proposed authentication model which is going to be integrated with the current server level authentication procedure. This manipulation will be done with a lot of software testing as to avoid to any flaw in the live operation. Here the fault tolerance should be replaced by a redundancy procedure which is also discussed in this topic.

The biometric integration with SALT value is explained below

**Step 1.** E.g. Biometric Image -> Binary/Oct/hexadecimal



--------→ 01000111110000011110111000001100001000000 1111000    (1)

**Step 2.** SALT Value (Randomly generated value used as user password -> Binary/Oct/hexadecimal

Each user password will be joined with a SALT Value and then converted to respective format. SALT Value is generated once and given to the user. User needs to remember his password and SALT value which he will get the RSA secure ID device.

   cristopher2101 + (concatenating) 2341 -> (01000111000100001100) (0010101010111)    (2)

**Step 3.** Converted value of Biometric image and result of SALT value + user password

   01000111110000011110 1110000 (OR) 01000111000100 = 011101010101011111 0010101    (3)

**Step 4.** Then apply this output to the encryption process (Three Tier Encryption Algorithm) which will do the replacement the replacement of the bits and then the output will be a number of the format as shown below.

   011101010101011111 0010101 -> 3242323131414113 -> 23456777888899989799 7123232354 (4)

Note: - The above value is just an example value not the true value.

The conversion is done as per the above example. The conversion format can be varied as per the requirement but the steps involved in the conversion will be as per the above mentioned example. Once this is converted in the above mentioned format, the hacker will just see it as a number but to decrypt this value will take many years and then to generate the image will not help the hacker in any ways to penetrate into the server thereby stealing the data's. This replacement of bits is done along with image conversion and concatenation of SALT value + password is only to bring about confusion for the hacker in tracing the original value. The value obtained after conversion will no way provide a trace on what is used in the conversion process. To make an analysis on this is a difficult task as the following things needs to be analyzed. In the authentication process even decrypting the encryption algorithm will be of a big challenge even though the steps used seems similar but input that are unique and especially biometric image is unique as well as SALT value changes for every server login and it is simultaneously matching with the template with the mapping output generated simultaneously. So penetrating and making a change is highly impossible. But that is how the authentication should work at the enterprise level and there should be a proper server authentication procedure 1. No of bit used in conversion 2. The value joined in the process concatenation (Password + SALT Value) 3. The value of image (which will generate only with the authorized user) Eventhough hacker derives the step 2, for step 3 he needs the authorized user to access, which is no way possible. That is where biometric provides an effective security feature with encryption. This methodology of Encryption has been designed in such a way that the authentication process is secured as the time to authenticate is also less. In the step 4 the output that is shown is how the value appears after the rearrangement of bits and after applying the Mathematical series. So the complexity of the output will be very high and also make a trace of exact authentication flow will be quite difficult. That is going to be final template and end of day reports are going to be generated based upon this authentication flow. When the encryption is done all that matters it the time to take the input, generate the output and authenticate. So how this going to be calculated will be show with a breakage with time duration in each stage of authentication process. We will see the complete analysis for other authentication techniques and also see which is going to be effective in authentication, probability of generation, easy to generate a biometric image with being less affected with the environmental effect like Sound, brightness etc... Then we are also going to see how the biometric is going to be used in message authentication too. That is going tell the positives of Biometric usage in authentication procedure at the server level.  Let us know the exact manipulation that I have proposed for the redundancy in server level authentication when we use Biometric authentication. When the authorized has got hurt but has to make change in the biometric image to authenticate the server to Login when needed. How can we do that? Is that any procedure that can be done with high level of security and without breaking up the security norms of the organization and the client? This will be done with a proper approval from the management team of both the organization and the client. How is it going to be done is going to be seen in the next section of this topic. Here there are going to be two options that will be there in this application Update and reset but that can be seen only in the "emergency access mode". Here the access for the application will be very minimal as this mode is dedicated for the only the update or reset the biometric image by authorized with a specific password that is again generated using the RSASecure Id device. This process is going to allow the authorized to go and change the biometric image in emergency or a periodic updation in the biometric image to make sure that the combination

provided should periodically been changed and also make the authentication process go without any flaw. This is also used when the authorized is hurt. This can be done with a proper approval from the managers of IT, change management. IT security managers, risk managers etc...Let us now see on that process in depth and the procedure that needs to be followed before making those changes in the live servers.

## 7.1 Authentication at server level

Let us see how the redundancy in biometric image can be generated in emergency level that is when the administrator has met with an accident or due to some unavoidable circumstances. It is pretty much simple procedure but this is also highly secured methodology of accessing the server. 1. When we press F8, the OS opens in **Safe Mode** 2. In this another option needs to be included for server OS alone is **"Emergency Access Mode."** 3. When we access this, there will an option to insert **biometric image, generate new SALT Value and press update + reset button**. Only that window alone opens. This will not allow access to any other resource on the server. Let us see the advantage of this methodology.

In this portion of the OS this option needs to be brought about and then the same needs be linked with the application too which needs to go through some of the process of approvals in risk management, change management. Normally bringing that change is not an issue but this option is linking with the access control, application access control and its database where this biometric images get stored. The complete analysis procedure will be seen in the coming topics. How this procedure is going to be implemented is what is going to be seen and the time that is roughly required for the resting of this modification. So there is going to two modification (1) inclusion of Emergency access Mode (2) Integration of Biometric with user password at Server login authentication. How this option is integrated with the application that is installed inside which will be authenticating the user in place of server authentication which includes only password. So that is where it is going to be a real challenge for the developers who are going to make this change with testing, approvals etc. Let us see how this entire process flow for this modification is going to be made. Here there will be a doubt that why the reset of the new biometric image can't be like change of new password at the login page as the biometric image can be changed periodically the following the main security reasons behind not keeping that option there are as follows (1) It will become an option that would not be known to the unauthorized user to misuse it in the absence of the authorized user. This option which is integrated in the safe mode should not to known to anyone else other the authorized users of that server and the management executives. The integration is complex as the updation should happen properly when the biometric images are changed it should generated the final template and then when the user login back again it should be able to properly authenticate the user without any issues. Those are some of the places where the testing needs to be done and then deploy this OS in the live environment. Let us how the management going to take a decision on this change. The management which will be the main body for the approval of such important options like this which is going to be a part of the redundancy in the live operation. When a biometric image is going to changed or going to add a new biometric image. Here we can see how the management view a modification when it is brought about in an OS. Here we are seeing the parameters like % of validity, % of redundancy,% of probability, % of feasibility  that are normally used to authorize a modification.
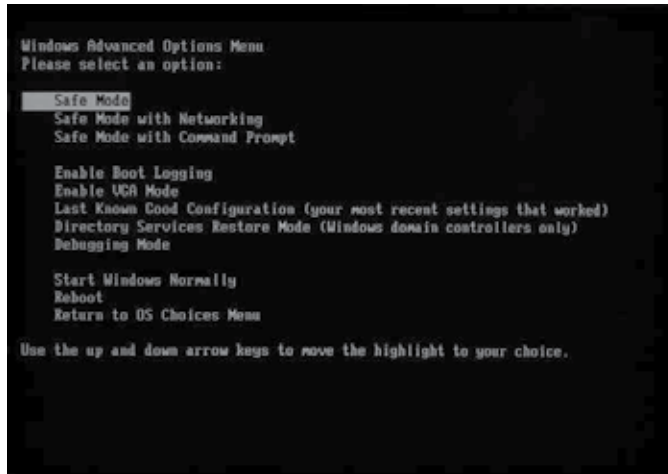
Fig. 6. Current Safe mode options

When a change is brought about the modifications needs to be discussed with the above mentioned and justifies the reason why this change is brought about. How this change is going to help in the authentication level. Here it is all about the redundancy step followed in authenticate. The entire process flow should be explained and also tell them if a new biometric is inserted by an authorized how the updation happens and how that changes and the new authentication results can be seen in the report log. What are the key data's that needs to be seen a report, all these things needs to be explained to the management as they are the people who are also responsible if there is any loss of data by chance. How to trace an intruder's access from the log and also how to track his network path is what are the queries that a management will have. The justification should be from the development team, infrastructure team and IT security as those are bodies who are designing this application.  What are the justifications for this modification? In this "Emergency Access Mode" the password is generated by a RSASecure ID which keeps changing the code every 60 seconds along with the password for the emergency access which will allow the authorized user to access the application with his server logon password and RSASecure ID code. Then the modification of the biometric image is done and saved in the encryption table. How this process is going to be secure approach for the modification. Here two things are unique (1) Code generated by RSASecure ID device (2) The user biometric image can't be caught and misused. Here when the biometric image is changed in the emergency situation or when there is a need to add a new profile how it is being done. When it is adding a new profile even the RSASecure ID device needs to be registered with this as the new user should be able to access the server without any issues. That is the reason once changed it needs to logged in and checked if that works without any issues. This needs to be carried out in the testing phase and not in the "live environment" as this will be a very costly issue if there is some unavoidable circumstances where the authorized is not able to produce his biometric authentication. Have 3-4 authentication techniques for a server authentication is always not advisable as it is like giving an option to the hacker to know the process that we are trying to manipulate. It should be a unique approach and there should be no trace of this modification to anyone even within the organization. In today's world the approaches are being leaked out in Media and the hackers are consolidating those

techniques to hack some valuable information's from many data centers. How to avoid this is by maintaining privacy within the organization on certain information's that are related to the confidential data's, security techniques and policies behind it. Once if these things are known to any of the user within the organization they can try to misuse it using any third party tool. If the user is not given access to a information's. He will try to threaten the System admin and can try to manipulate the things within the organization. To avoid all these things the security policy and methodology should not shared to employee even though he is friend or relative of the System admin. That is the reason why the agreement needs to be signed by the system admin and organization as a agreement normally called as OLA in management term. This will allow the system admin to take a risk on this as it will in turn going to question him and not the management by the Client. So this will bring about a strict policy in the security approaches. This is how the process is secured even in the modification or adding of profile in Biometric authentication. Some of the advantages of this approach can be seen before seeing the complete advantage of Biometric authentication over current authentication methodology which are as follows (1) the NT authentication will be highly secured protecting the data's on the server (2) The possibility of breaking up the password will be highly impossible, as the encryption algorithm is changed on regular intervals along with biometric image on a quarterly basis (that is image of another finger of the user). The possibilities/ probability of changing the encrypted value is high with biometric and encryption (3) the approach for encryption is simple and decryption process for hacker is highly impossible. This methodology has lot of other advantages which will be seen after on how to frame this methodology on the basis of ITIL framework. Then we see on the analysis of each biometric technique on the basis of the parameters like error rate in authentication, error rate in initial registration, error rate in accepting new user, error rate in other factors like Light, sound etc. Finally in this chapter we are going to see some information on the security policies that needs to adopted by the organized for this methodology as it involves a lot of confidential information's before getting an authentication to a server and this methodology is  specifically designed for the Enterprise level date centers. Finally but not the least we will see the future modification and other techniques going to brought about using the similar kind of methodologies. These are some of the things that are going to be discussed in the coming topics. As we have in the Fig 23 there is "NO" condition which tells there requires some modification so what can be the possible reason behind it will be (1) % of redundancy (2) % of error in wrong authentication acceptance (3) % of error is not accepting a authorized biometric image (4) % of flaw in application (5) % of time taken in authenticating at critical situations (6) %of feasibility in using that application are of the some of the common factors that comes in the minds of a managements. As a management employee he won't see on how this application is going to function but on how it going to keep the information secure as well as how it helps a n organization to drive a business easily with it secured approach in maintaining the data's of it clients. They see that whenever a organization needs to be retrieved from a specific server the authorized user should be able to get it without any issues in getting the authentication from the server. If the server is not accepting it then going to the Emergency access mode is quite a risky approach as it is the live server which he is turning down for a minute which is going to have a negative impact on a organization. The approach over here should be different and this is what going to come under the account of % of authentication acceptance.

$$\% \text{ of acceptance } = \text{No of acceptance } - \frac{\text{No of rejection}}{\text{total no of logins}} \tag{1}$$

$$\% \text{ of EAM usage } = \text{Total no of valid access } - \frac{\text{total no of invalid access}}{\text{Total of access in EAM}} \tag{2}$$

$$\% \text{ of redundancy } = \frac{\text{total number of possible biometric image generated}}{\text{Total possibility by that biometric technique}} * 100 \tag{3}$$

These are some of the parameters which are normally calculated at the end of the day report which will show the complete statistics of the biometric authentication results at the end of the day. Now this information's are normally consolidated at the testing phase of the application itself. Let us now see how the testing of this integration of the application is going to be done. The testing's commonly done at each module as well as complete application. Let us see how this module level testing is done and then on complete application.

### 7.2 Change management - biometric authentication methodology

In the change management of this biometric authentication methodology there are few possible that can be made either in biometric technique or the encryption flow in the application which will be carried out phase by phase after discussing the test result with the change advisory board and other management team before deploying it in the Live operation. There can be an emergency that can be brought about if it required if the hackers if finds a possible approach to reach the confidential information. In that there is going to be a decision going to be made by the emergency change advisory board to deploy the approach immediately here it involves a lot of risk and the downtime requirement if any. So all the information needs to be tested initially itself and should be submitted at the time of need to the management to understand about that approach .So the change requires a proper documentation on the location of the application where the  modifications are made. Once this documentation is done along with the test results, then it provides complete test result with justification for bringing about the change when required for the application. So let see how the process flow is designed for this methodology. This is the process by which the change is going to take place for this biometric authentication methodology. When it is accepted then the It team has analyze on the reason why this change was not accepted it is going to affect the live operations of other application or bringing about this change is not going to bring any effect on the hacker penetrating the network.

As far as the biometric authentication is concerned the entry into the server is highly impossible by making these changes the security on the server infrastructure will be high so that the data's that are stored on the server are highly secured without any flaw that allows a third party person to access and view any data's that are stored on the server. Making periodical changes with proper testing on Test server by giving sample inputs will never lead to any issues in the change of an application on a live environment. That is the effective approach for the change management in common. Once this changes are made this needs to document is secured location as these things need not be shared to anyone other the users of this methodology. Bringing out this methodology in the live is not good as it will create an alarm to the hacker that the biometric is going to be used for the authentication purpose.

This biometric authentication is going to have a change only to make sure that the process flow is updated periodically with a new one so that there is security that is maintained on the data's that are stored on the server. Once the server security is properly updated then the output can be seen in the security at the enterprise level. Why the client is involved in this change which is going to be a minor change. It is very important to convey the changes that are being made on the server where they are storing their data's or hosting their websites. As this allow them to give their point of view on this change. Then based upon the approvals from both of them will allow the IT team to go ahead with the modification or else suggest them another approach or reason for not making the change. Based upon which the IT team can provide their justification why this change is made and what is the benefit of it behind it. Then accordingly the changes can be made on the authentication methodology.
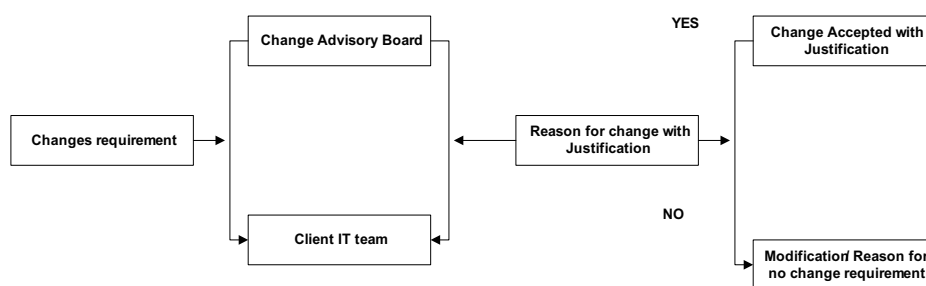


Fig. 7. Change management process flow (Biometric authentication methodology)

### 7.3 Risk management - biometric authentication methodology

The risk that is involved in this methodology is very less and that is only the report generation when the change or modification that is done on the process flow. Even this can be avoided when it is tested during the testing phase using the sample inputs. So the risk involved in the biometric image is also an important one that needs to be taken into consideration but that can be justified as a server is accessible by the entire authorized administrator as it is not that a user when registered on a server can access only that server alone. Let me provide you a screenshot on how it looks when a authorized user tries to access any other server in the need of emergency. Here there is no risk involved as this option server name is asked when the biometric image is generated on other server. Let me provide this with an example If the user X has registered his authentication on the server CNHDLADS01 now due to some emergency in restoring the e-mail he access CHNDLML01 then it will ask for this server name along with the biometric image in order to check for the biometric image, the password and the RSASecure ID certification code which is already stored on other server then it will generate the encrypted value to authenticate the user to access the mail server. This process doesn't involve any risk if the testing and the integration are done properly by the development team. Once this is done the user can access the server and restore the e-mails to the user. There are some of the risks that can be handled without any issues in this methodology. This is the main speciality of this methodology as it involves a risk free approach. Even the small risk also can be handled within no time. So the management will get a justifiable reason for this methodology. This is how the risk management is carried out for this methodology. When the justification reason for the risk is

not agreed then it needs to be analyzed and produced with a sample input value that is going to convey the management that why this risk is there and how this can be overcome.



Fig. 8. Server authentication (Accessed by other authorized admin)

When a change is made the reports generated on the entire server and the consolidated report generated on the repository server all should get properly aligned with the change that is made in the process flow of the authentication. That can be tested in the testing phase itself so there is only a risk about 0.0000000001% in this approach. So then this it will tell how easy it is for executing and maintaining this approach of server level authentication. The risk management is important as without understanding the risk involved in a application the redundancy can't be developed for an application which is going to be integrated with a Live server.

## 7.4 Redundancy management - biometric authentication methodology

In the redundancy management we are going to see how the biometric images from all the servers are stored on common repository. When there is any fault in a server and it is being reinstalled or replaced then the same biometric image will be loaded back with the same user password. The only change will be the SALT value as it is being generated every 60 seconds RSASecure ID device. How this is going to be carried out and the reason for which it is carried out is to avoid more downtime. Once the server is up and running, the authorized user has to just start login into the server and also to make the process easy as in the live operation this is what they expect from the vendor organization which is maintaining their information. Less downtime with high security is what should be the goal of a data centers at enterprise level. The redundancy of this application is only the above mentioned things as rest all is just deployed if there is any crash in the server or server is being replaced. So what is the situation regarding the log reports, they are taken backup regularly from all the servers on a daily basis and they are sent to all the management team. So there will be no loss in any of the information that is being generated by this authentication methodology. So this is going to tell how redundant the application at the time of emergency is. This will help an organization to keep itself secured with a much easier approach and maintain the same high level of security. This is how all the three parameters is going to keep the organization secured, and also provides a proper approach maintain the changes and handle the situation which are mentioned as risk. In the

methodology all parameters are simple and can be restored easily. The important factors that need to be seen in order to achieve this perfection in the implementation of this methodology are training the security policies that need to be set for this methodology. Let us some information's on them after the analysis of this methodology.

## 8. Advantage of this biometric authentication technique

There are lot of advantages of this biometric technique which are as follows 1. These kinds of biometric authentication technique on the server side have not been implemented as the operating systems like Linux and Solaris are considered to be highly security. Even then the hackers are able to hack the data's from the server by breaking up the password. Here this biometric authentication technique will be effective as generating the biometric image is not possible other the authorized IT administrator of that server 2. This biometric authentication has been designed in such a way that it includes a highly secured authentication login technique with encryption algorithm which uses a different approach for generating the final encrypted template using the rearrangement of bits methodology. This makes a very highly secured approach in accessing a server. This kind of authentication can't be seen in recent server authentication. The server authentication application is designed in such a way that it is used in multiple platform which just a small package of deployment to integrate this with the existing authentication technique 3. This authentication technique has a unique feature in authentication when a user is authorized on the server CHNMCADS01 and he access the information on server CHNMCMXL01 then it will ask for the server name where he is registered as it will map with that server for the biometric image and password in order to generate the final encryption template to authenticate the IT administrator. This is the greatest advantage of this approach as the IT administrator need not have to create a new profile on this server in the Emergency Access Mode in order to access this server or need not have to call the respective authorized administrator to access this server. This is the main advantage of this technique which is not in of the current server infrastructure 4. This kind of authentication technique is unique in both the report generation and Emergency access mode as the report is generated and sent automatically sent to the management with the information of the unauthorized user access with his information of the IP address. This authentication technique acts more similar to the network monitoring tool.  The Emergency access mode is designed with a lot of limitation with just opening the application which can be used to reset the biometric image/password or can add a new user profile. Nothing else can be seen or can be accessed with this Emergency access. There will be a separate password which will be there with the authorized user of that server. These password should be used anywhere as per the security policy of this application. These Emergency access mode passwords are generated by the provider and provided to the authorized user at the time of delivery of the application. This is accessed only when it is needed with the approvals from the management and after the working hours/non peak hours 5. This methodology has it special encryption technique which has the process of rearrangement of bits which will unique as the output is a number so the final encrypted template will be a number so the hacker even gets this number he will be not be able to get the password, SALT value which keeps changing every login and the biometric image which is unique with all the users. So the authentication is based on biometric image and the user password but the SALT value is to manipulate the final template periodically after every login. This will never give to the hacker on the process flow of this encryption algorithm.

## 9. Future enhancement of this biometric authentication technique

The future enhancement of this methodology is the file transfer authentication using the biometric and the SALT value. Here these two concepts will be integrated with the file that is being transmitted over the network. There will be authorized user only they can decrypt all the confidential files with their biometric image so the hacker will not be able to read any information without this SALT vale and biometric as the process flow will be something which is used for the server authentication technique. Here the File sent over the network will be encrypted using this application which will be electrically signed using this Biometric image of the authorized user with his password and SALT value and it will be decrypted by the authorized user at the other end. This concept is not related to stenography where the information's are embedded in a common images which was used for the 9/11 world trade center attack. Let us see the process flow of this methodology which will does not includes the encryption process as it is yet to designed for this. As you can see from the diagram below how the authentication for the file transfer has been done. The hackers can decrypt any format even if it 0's and 1's. There are tools that can try to give them the clue on those information's. Even after that how this biometric is going to play a vital role in this authentication process of file transfer. Biometric authentication is something unique as it can't by any other person other than the authorized user. When the biometric authentication is considered for this information security it is sometimes considered not a feasible approach as everything file that is sent has to be encrypted and sent manually by an authorized user. But on the other hand confidential data's when transferred with a proper security authentication technique then it is going to provide high level of security not only for their data's but preventing the hackers from stealing the information's of an organization. This is the authentication technique on which I am going to work on with a new encryption algorithm that will make this encryption process much feasible and much suitable for the enterprise organizations. The complete analysis is going to do with the security attacks that are going on currently. Along with this I am going to work on wider scope of this technique even in ATM transactions and net banking where quite a higher level of security is required. The biometric usage has been bought in Yahoo mail but don't how far this technique is followed by the user as her also the feasibility and the awareness of this biometric usage has to be explained more over the users with laptop will have this biometric option that too on higher configuration models alone.
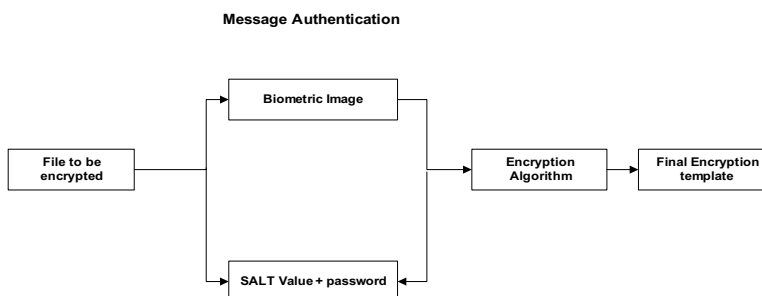


Fig. 9. Message authentication using the Biometric and SALT Value

Users who have desktop have to get biometric device as separate component Right now the biometric should be integrated with the existing keyboard so that both can be used for the

authentication purpose. Net banking have an image option which needs to be selected while using the net banking for the first time. Then it will be displayed for the user when he accesses the account for the consecutive time. But it will not considered as a proper authentication technique as we hear a lot of hackers who are stealing customer's information even though the web site is secured by a third party security provider. That is where the biometric integration with the information is going to play a vital role as when the information is encrypted with the image then dividing them is not an easy task when compared to the information just encrypted with the encryption algorithm. That is where we can see the real usage of biometric integration with confidential information. The biometric approach of authenticating the user is considered to be the most positive sign regarding the biometric techniques in today world where requirement of security is high. When the biometric is used with any form of security section say authenticating a resident person, employee of an organization, authorized user of a server (In infrastructure support) it has been feasible, reliable and above all the security that it provide is very high as each biometric image that is generated is unique. But here comes the technique that can be used and it varies with the sector like Retina and face recognition can be for employee authentication, retina and finger print for laptop authentication. Based upon these criteria's, going to design the next authentication technique for messaging system at enterprise organizations. This application is going to maintain a high level of security for the messages that are been shared between the clients and the branches of the organization. This application is going to maintain the logs in similar fashion but the server utilized with a centralized with a backup server as a source of redundancy. This application is going to be operating system based which will be implemented at the enterprise and not at a consumer level.

## 10. Conclusion

In this chapter we have seen how a biometric integration is going to used in Server authentication at enterprise level with the SALT value and encryption algorithms. When the biometric technique is used in authentication what are the parameters that needs to be followed and how it needs to taken into consideration with the management views are been discussed in this chapter. Then finally we have seen how to provide training for this application as training is considered as an important part of the IT transition. Based upon the assessment only the enhancement of this application also can be carried. But this will discussed in the initial stage itself before providing this application as when the security norms are signed and followed it should be followed like a holy book. As the IT admin should be simultaneously updated with the recent security threat and what is the solution that is enhanced from the application side. Then we have seen about the testing phase of the application during the implementation of the application at enterprise level and the information's that need to be checked while implementation of this authentication technique. Then we have seen recommended technique in biometric, which are completely based on the above parameter based on both management and the biometric technique parameters. Finally it is all about the IT security policy which is set for the application based on the current policy norms that are set by ISO standards, information policy as mentioned in CISSP, CISA and CISM. All these deal with the information security policy. These are some of topic we have discussed with the some real time examples of biometric authentication in other technologies. Then we have a topic that is "Need to know principle"

which is a wonderful topic that tell about the limitations that needs to be set in security policies. These are the information's that has been discussed in this chapter. It is like providing a complete overview of this Biometric integration with the SALT value at the enterprise level server authentications. Here it not only show the technique but followed by analysis, view of management with the important parameter and how this technique is going to be better that just a normal password authentication.

## 11. Acknowledgement

## 12. References

Amar Merrad, Noureddine Goléa. (2010). Multi-Layer Perceptrons Approach to Human Face Recognition. *Journal of Automation& Systems Engineering,* PP. (165-172)

Emmanuel Opara, Mohammad Rob, Vance Etnyre. (2006). Biometric and Systems Security: An Overview of End-To-End Security System. *Communications of the IIMA,* Volume 6, Issue 2, PP. (53-58)

Seifedine Kadry, Hussam Kassem, A new secure design for mobile communication. *Journal of Theoretical and Applied Information Technology*, PP. (652-657)

Debnath Bhattacharyya, Rahul Ranjan, Farkhod Alisherov A., and Minkyu Choi, (2009), Biometric Authentication: A Review. *International Journal of u- and e- Service, Science and Technology,* Volume 2, Issue 3, (September, 2009), PP. (13-28)

Anil Kapil, Atul Garg. (2010).Secure Web Access Model for sensitive data, *International Journal of Computer Science & Communication,* Volume 1, Issue no 1, (January-June 2010), PP. (13-16)

Pijush Kanti Bhattacharjee, Chandan Koner, Chandan Tilak Bhunia, Ujjwal Maulik. (2010), Biometric Entity Based Mutual Authentication Technique for 3-G Mobile Communications. *International Journal of Computer Theory and Engineering,* Volume 2, Issue 1, (February, 2010), PP. (26-30)

K. Saraswathi, Dr. R. Balasubramaniam. (2010). Bio Cryptosystems for Authentication and Network Security-A Survey *Global Journal of Computer Science and Technology,* Volume 10, Issue 3, (April 2010), PP. (12-16)

S. Akrouf, *Member,* A. Bouziane, A. Hacine. Gharbi, M. Mostefai, Y. Chahir. (2010). Towards an Intelligent Multimodal Biometric Identification System, *International Journal of Computer and Electrical Engineering,* Volume 2, Issue 6, (December 2010), PP. (1001-1004)

K.Sasidhar, Vijaya L Kakulapati, Kolikipogu, Ramakrishna K KailasaRao. (2010). Multimodal biometric systems – Study to improve accuracy and performance, *International Journal of Computer Science & Engineering Survey* (IJCSES), Volume 1, Issue 3, (November 2010), PP. (54-61)

Dr.R.Seshadri, T.Raghu Trivedi. (2010). Generate a key for MAC Algorithm using Biometric Fingerprint. *International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC)*, Volume 1, Issue 4, (December 2010), PP. (38-45)

Arian Rahimi, Sharhriar Mohammadi, Rozita Rahimi. (2010). A New Web-based Architecture Based on Iris Biometrics Technique to Decrease Credit Cards Frauds over Internet., *International Journal of Digital Society (IJDS),*Volume 1, Issue no 2, (June 2010)

Jin-Woo Jung, Dae-Jin Kim, Z. Zenn Bien. Realization of Personalized Services for Intelligent Residential Space based on User Identification Method using Sequential Walking Footprints, *Systemics, Cybernetics AND Informatics*, Volume 3, Issue 2

ITIL V3 Improves Information Security Management Ginger TaylorEast Carolina University http://www.infosecwriters.com/text_resources/pdf/GTaylor_ITIL.pdf

Guidance on Aligning COBIT, ITIL and ISO 17799 http://www.isaca.org/Journal/Past-Issues/2006/Volume1/Documents/jpdf0601-Guidance-on-Aligning.pdf

Lakxman Kumar C, Arunachalam P, Sandhya S. (2009). Biometric Anti-theft and Tracking System for mobiles – BATS. *International Journal of Recent Trends in Engineering,* Volume 1, Issue 1, (May 2009), PP. (237-242)

# Chaos-Based Biometrics Template Protection and Secure Authentication

Xiaomin Wang, Taihua Xu and Wenfang Zhang
*School of Information Science and Technology, Southwest Jiaotong University*
*China*

## 1. Introduction

With the increasing development of global economy and information technology, more and more fields require reliable identity authentication. And with information age characterized by digitalization and recessiveness of identity, a key problem to be solved is how to identify a person's identity accurately and ensure information security. In this regard, a variety of inherent human biometrics were gradually understood and studied, thus the development of biometric identification technology is considerable. The gradual yet profound application of biometric identification system today has improved security and creates much convenience to identity authentication. However, there are still some inherent problems that need to be solved. For instance, masquerade attack, difficulties to republish when the template is lost and a series of other potential threats. The existences of these threats have created a bottleneck, constraining further development of the biometric identification technology.

In this chapter, we will firstly give a review mainly on the theories and techniques of biometrics template protection, and then present a novel chaos-based biometrics template protection with secure authentication scheme. The proposed scheme is lightened by fuzzy extractor, yet includes two-layer error-correcting (one is BCH error-correcting code, the other is chaotic spread spectrum encryption) to achieve a good authentication performance of GAR=99.5% and FAR=0%. In addition, the functional features of proposed authentication scheme are: (1) do not need user to remember secret information such as password, or store them into physical media such as token or smart card; (2) no biometric template and any other secret information stored in server end; (3) the user's biometric template is cancellable; (4) user's registering information can be updated freely and easily. (5) with the help of user's inaccurate biometric template, secret information (user maybe knows or unknowns) can be accurately recovered. These interesting features push forward the proposed scheme having potential application in biometric-based authentication/identification systems.

### 1.1 Biometric and biometric identification systems

Traditional identity authentication methods are based on what is physically possessed such as ID cards and what can be mentally stored in the memory such as passwords and keys. The shortfalls of both are for instance ID cards can easily be lost or forged while passwords and keys can either be easily guessed or forgotten respectively. Short passwords are often easy for memory but easily guessed by others. On the other hand, long passwords (commonly known as keys) although cannot be easily guessed are prone to memory

problem. Key storage is therefore an issue and it is recommended that general long keys are stored in key cards and at the same time use short passwords to protect the Key Cards (Wang et al., 2006, Wang et al., 2007). Eventually, short passwords are still essential to identity authentication security.

Biometric (Tian, 2005) features inherited in person include two major categories which are person's physical characteristics and behavioural characteristics. Physiological characteristics are fingerprints, face, iris, palm prints, and voice to name but a few. Behavioural characteristics include gait, signature, keystrokes etc. These characteristics have attracted a large number of scholars who conducted extensive and thorough research on them. In order to perform the identification, an automatic technology is adopted to measure these features, and have them compared with data from a database template. This infers that identification and biometric identification technology is the solution to the certification.

Before the popularization and application of computers, biometrics was carried out manually mainly by artificial experts (e.g. American FBI for instance have large fingerprint experts). The development of productivity and popularity of information technology today have made biometrics to be automated using computers. The Automatic Fingerprint Identification System (AFIS) for example is one of the automated systems ever established. A typical AFIS includes an off-line register and an on-line identification process, as shown in Fig.1 (Li et al, 2009). The off-line register includes signal acquisition, feature extraction, template storage and other necessary steps. The on-line identification includes a signal acquisition, feature extraction, registration, template matching etc. Biometric identification system has two modes for identity authentication: authentication (1:1) and identification (1: N). Authentication mode test are "you the person you claimed", and identification mode test verifies "your identity information in the database and who you are". The two methods have large gap in aspects of their algorithm processing time complexity.
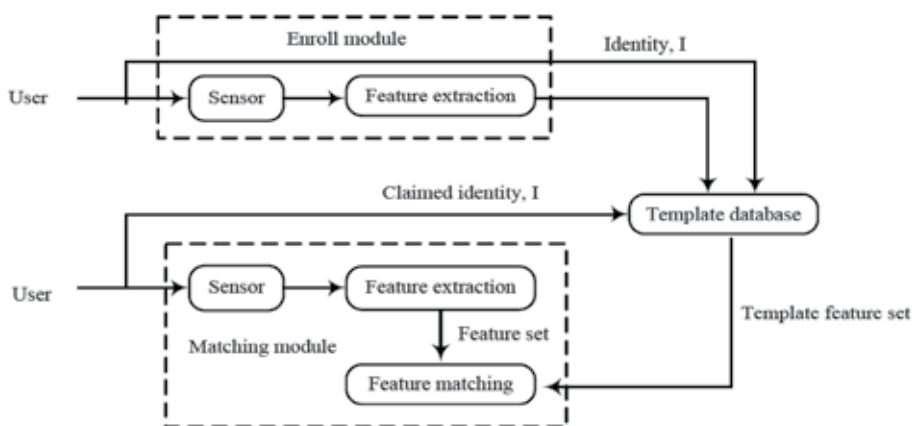


Fig. 1. Enroll model and matching module of biometric system (Li et al, 2009).

## 1.2 The defects of traditional biometric identification system

Traditional biometric identification system has increased in terms of recognition accuracy and speed. Yet, most traditional fingerprint identification systems adopt minutiae as their recognition features and the information of location where the direction of minutiae are stored for comparison in the form of pure data. The traditional system stores original

coordinates of minutiae and their value of direction, unfortunately, without any encryption. With the development of hardware attack and crack technology the whole biometrics identification system will be completely exposed to the scope of hacker attacks, threatening the security and privacy of user identity. Unlike passwords and keys that can be reset after their loss, the loss of biometric is permanent.

Cappelli et al. (2007) shows in a novel approach that the original fingerprint can be reconstructed automatically from standard minutiae-based templates. This may unlikely fool a human expert but is definitely possible to successfully attack even state-of-the-art automatic recognition systems, provided that one is able to present reconstructed images to the system. Thus there is the higher need for template security of biometric identification systems. Besides outside threats to template security, biometrics identification system is also facing a variety of other types of attacks.

In particular, Ratha et al. (2001a) did specific analysis on the sources of vulnerable attacks on the biometric identification system, and put them into 8 categories, as shown in Fig.2.
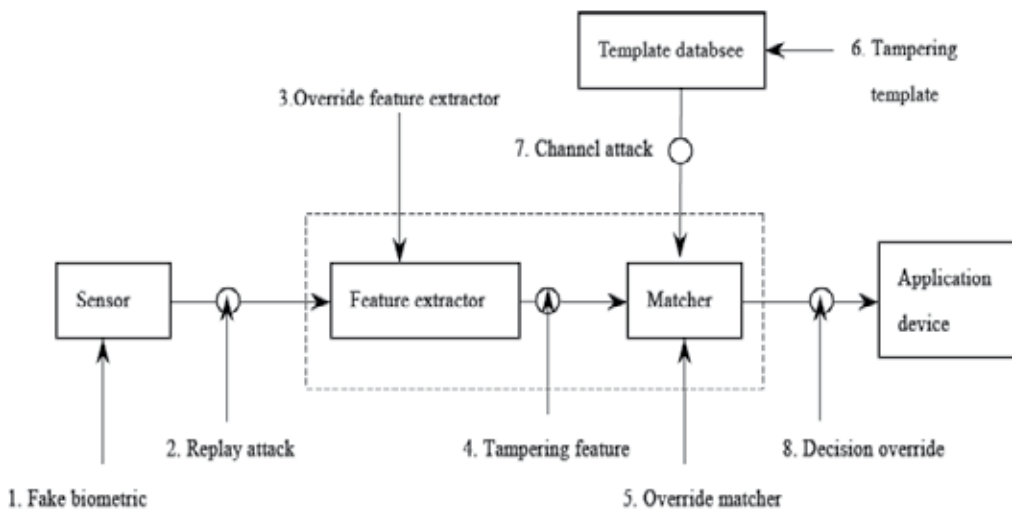


Fig. 2. Possible attack points in a generic biometrics-based system (Ratha et al, 2001a).

The eight basic sources of attack are described as below:
1. Fake biometric at the sensor: In this mode of attack, a possible reproduction of the biometric being used will be presented to the system. Examples include a fake finger, a copy of a signature, a face mask.
2. Resubmission of old digitally stored biometrics signal: In this mode of attack, an old recorded signal is replayed into the system bypassing the sensor.
3. Override feature extract: The feature extractor could be attacked with a Trojan horse so that it would produce feature sets chosen by the hacker.
4. Tampering with the feature representation: After the features have been extracted from the input signal they are replaced with a different synthesized feature set (assuming the representation is known).
5. Override matcher: The matcher is attacked to always directly produce an artificial high or low match score.

6. Tampering with stored templates: The stored template attacker tries to modify one or more templates in the database which could result in authorization for a fraudulent individual, or at least denial of service for the person associated with the corrupted template.
7. Channel attack between stored templates and the matcher: The templates from the stored database are sent to the matcher through a channel which could be attacked to change the contents of the templates before they reach the matcher.
8. Overriding Yes/No response: If the final result can be overridden with the choice of result from the hacker, the final outcome is very dangerous. Even if the actual pattern recognition system had excellent performance characteristics, it has been rendered useless by the simple exercise of overriding the result.

Due to the existence of the above threats to biometric system, it can be said that biometrics have degenerated gradually from "inherent features of you" to "features of what you have" to a certain extent. On the contrary passwords and keys can overcome this danger through encryption. Biometric cannot be protected directly through encryption, for instance, the hash function, as the great Hash intra-variance of it. However, it provides a feasible way for protecting the safety of biometric templates that combined biometric science and cryptography. There is the biggest obstacle to above combination that the contradiction between accuracy required by cryptography and inherent ambiguity of biometrics even if more and more researchers realized the advancement of the combination. How to overcome that contradiction in the condition of guarantying authentication performance of the system is the content of study on various biometric templates protection algorithm.

## 2. Review of biometric template protection technologies

This section focuses on classical biometric template protection theory and algorithms in the academic field. In a general viewpoint, we divided the biometric template protection into four groups: (1) **Biohashing** (Jin et al, 2004a, 2004b, 2004c, 2005, 2006, 2007, 2008; Lumini & Nanni, 2006, 2007; Jain et al, 1999; Nanni & Lumini, 2006, 2008a, 2008b; Connie et al, 2004; Ling et al, 2004, 2006; Maio & Nanni, 2005); (2) **Template encryption** (Soutar et al,1999; Davida et al, 1998; Juels & Sudan, 2002); (3) **Geometric transform of template technology** (Ratha et al, 2006, 2007; Ang et al, 2005; Clancy et al, 2003; Lee C et al, 2007; Lee Y et al, 2007; Tulyakov et al, 2005, 2007; Hao et al, 2006; Jain et al, 2006; Juels & Wattenberg, 1999; Juels & Sudan, 2002; Davida et al, 1998; Wang & Plataniotis, 2008; Uludag et al, 2005; Nandakumar et al, 2007; Kholmatov & Yanikoglu, 2008; Chang, 2006; Dodis et al, 2004, 2006; Mihailescu, 2007; Scheirer & Boult, 2007; Nyang & Lee, 2007; Jin et al, 2007;  Buhan et al, 2007; Boyen, 2004; Boyen et al, 2005; Li, Q et al, 2006; Sutcu, 2007; Tong et al, 2007; Arakala et al, 2007; Zhang et al, 2008); and (4) **Template hiding transmission** ( Khan et al, 2007, 2010).

### 2.1 Biohashing
The cancellable biometrics issue was addressed by Connie et al. (2004) which adopted a technique known as BioHashing. Jin et al. (2004c) proposed a novel approach of two-factor authenticator, based on iterated inner products between tokenised pseudo-random number and the user specific fingerprint feature, which generated from the integrated wavelet and Fourier–Mellin transform (WFMT), and hence produced a set of user specific compact code that named as BioHashing. WFMT features were chosen in this algorithms because in WFMT framework, wavelet transform preserves the local edges and noise reduction in the

low-frequency domain (high energy compacted) after the image decomposition, and hence makes the fingerprint images less sensitive to shape distortion. In addition to that, the reduced dimension of the images also helps to improve the computation efficiency.

The fingerprint feature vector is acquired after fingerprint image passed through wavelet transform, FFT transform, log-polar transform and high-pass filtering. As log-polar transform, the vector is invariable to translation, rotation and scale. Pseudo-random number can be calculated based on a seed that stores in USB token or smart card microprocessor through a random number generator. And a data T can be produced by iterating inner product between the pseudo-random number and the wavelet FMT fingerprint feature. Then the biohashing code is obtained by quantizing T with T=0 if T$\leq \tau$, otherwise T=1, where $\tau$ is a preset threshold. The BioHashing progression can be illustrated as in Fig. 3.
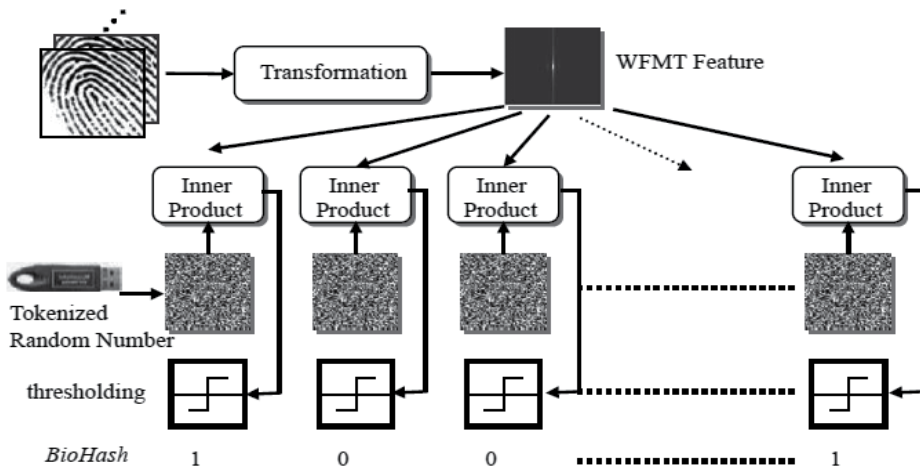


Fig. 3. Demonstration of BioHashing process (Jin et al, 2004c).

However, if the user token was stolen, the performance of BioHashing would be lower than that using only the biometric data (Lumini & Nanni, 2007; Nanni & Lumini, 2006, 2008). It can be concluded that the main factor is pseudo-random number, instead of fingerprint itself.

Lumini & Nanni (2007) proposed an improved BioHashing approach which is more robust than the original method. They consider that the case of loss of random number can be solved by extending the length of hashing key. Then they put forward four improvement measures to extend the length of key, i.e.

- **NORMALIZATION:** Processing with orthogonalization of generated vector.
- **$\tau$ VARIATION:** Instead of using a fixed value for $\tau$, use several values for $\tau$ and obtain varying $\tau$ between $\tau_{\max}$ and $\tau_{\min}$, with $p$ steps of

$$\tau_{step} = \left( \tau_{\max} - \tau_{\min} \right) / p \tag{1}$$

- **SPACES AUGMENTATION:** Augment the length of key to $k$ times of origin by space augmentation to be $K$ spaces.
- **FEATURES PERMUTATION:** Using $q$ permutations of biometric vector and obtained by round-shifting the coefficients of a fixed amount thus obtaining $q$ bit vectors.

The result of improved BioHashing procedure, if all the above solutions are exploited, is a set of $k \cdot p \cdot q$ BioHash codes, which are compared by the Hamming distance. The verification task is performed by training a classifier for each BioHash code and finally by combining these classifiers by a fusion rule (we suggest the SUM rule).Thus it enormously increased length of hashing key, the problem of original algorithm is solved.

Biohashing algorithm was originally proposed for the fingerprint, but the algorithm requires highly differentiated fixed-length features which are very difficult to extract in the fingerprint. FingerCode (Jain et al, 1999) has a fixed length, but a low discriminabiltity, can not assure the certificated performance under the circumstance of loss of random number (Lumini & Nanni, 2007). The Biohashing algorithms of other biometrics, such as face, palmprint, have been proposed and carried out relevant research (Jin et al, 2004a, 2004b, 2006; Nanni & Lumini, 2006, 2008a; Connie et al, 2004; Jin & Ling, 2005; Ling et al, 2004, 2006). Some of the new technology applied also to Biohashing algorithms, such as probabilistic neural network (PNN) (Lumini & Nanni, 2006), Gray coding (Jin et al, 2007, 2008).It also applied to Biohashing algorithms that the technology of multimodal fusion and multi-feature fusion, to settle the problem of high EER in the term of loss of random number (Maio & Nanni, 2005; Lumini & Nanni, 2006; Nanni & Lumini, 2008).

## 2.2 Biometric template encryption

Bioscrypt algorithm was proposed by Soutar et al. (1999), which is one of the earliest algorithms about biometric encryption. The basic idea is based on image processing and Fourier transform. The algorithm has two steps: enrollment (as shown in Fig. 4(a)) and verification (as shown in Fig. 4(b)).

**Enrollment phase:** In the stage E-1 called Image Processing, combine a series of input fingerprint images with a random (phase) array to create two output arrays that are $H_{stored}(u)$ and $c_0(x)$; In the stage E-2 called Key linking, link a cryptographic key $k_0$, to the pattern, $c_0(x)$, via the link algorithm; In the stage E-3 called Identification code creation, create an identification code $id_0$, derived from the key $k_0$.

**Verification phase:** In the stage V-1 called Image Processing, combine $H_{stored}(u)$ from the bioscrypt, with a new series of input fingerprint images to create an output pattern, $c_1(x)$; In the stage V-2 called Key Retrieval, extract a key $k_1$ from $c_1(x)$ using the retrieval algorithm; In the stage V-3 called Key Validation, validate $k_1$ by creating a new identification code $id_1$, and comparing it with $id_0$.

Also, there are criticisms to the algorithm from literature (Davida et al, 1998; Juels & Sudan, 2002) that the algorithm carried no rigorous security guarantees. It does not count the entropy loss of algorithm in enrollment phase and not present definitely the rejection rate and false acceptance rate. In addition, the authors assume that the corresponding fingerprint image is pre-registration in the course of the experiment, in fact, it is difficult to achieve.

## 2.3 Geometric transform of template technology
### 2.3.1 Geometric features transform

Ang et al. (2005) consider a key-dependent geometric transform that is applied to the features extracted from a fingerprint, to generate a key-dependent cancellable template for the fingerprint. The method reduce the EER according to the experiment with FVC2002 database, while the drawback of the method is that it has to detect singularity, and singularity itself is difficult to detect precisely, so the associated error will be introduced,

what's more, some types of fingerprints does not have singularity(such as arch). In addition, there is some inaptitude when folded templates are treated with common matching, such as there may be a coincidence that the minutiae to be overwritten while folded.
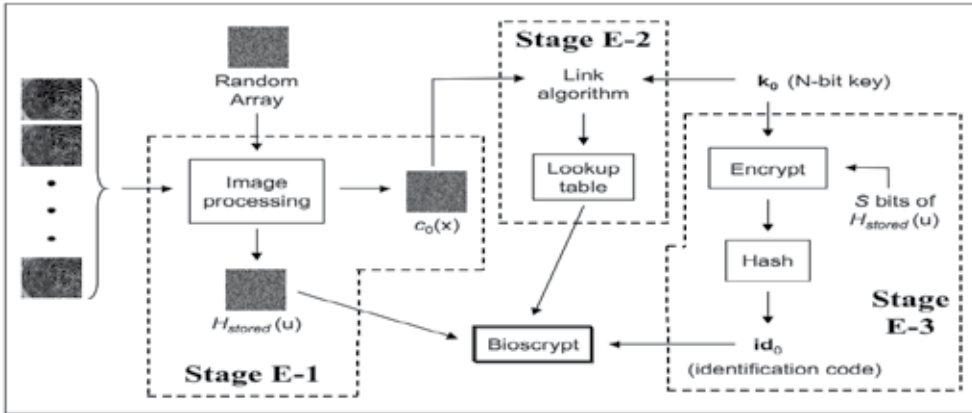


(a)
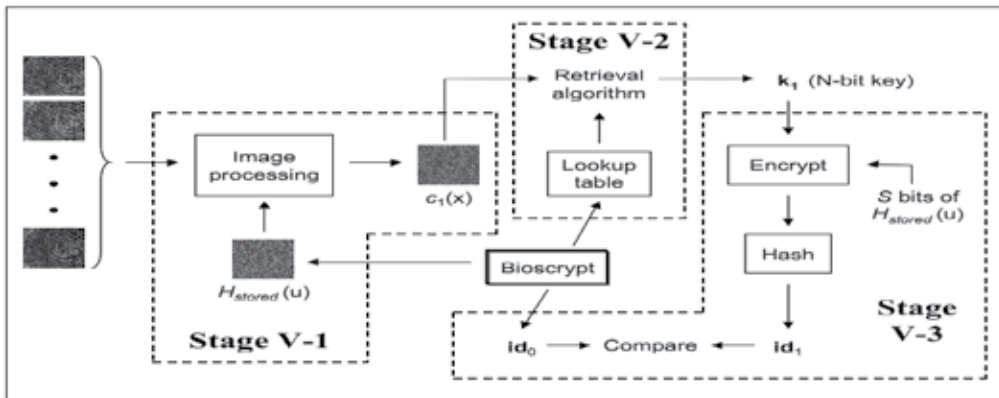


(b)

Fig. 4. (a). Enrollment phase of Bioscrypt algorithm (Soutar et al, 1999). (b). Verification phase of Bioscrypt algorithm (Soutar et al, 1999).

Ratha et al. (2006, 2007) presented a method of template transform. The method transforms the set of fingerprint minutiae from original space to another space using a one-way function. However, the performance of transformed template is lower than original template using the method. The reason is that there is deviation of transformed minutiae position from expectation, and additional registration to transformation function can avoid the descend mentioned above, but the registration is difficult to control. Lee C et al. (2007) presented a method without additional registration to transformation function, whereas, the method still does not reduce the risk of that system is attacked as loss of key.

Actually, Tulyakov et al. proposed a method named Symmetric Hash Functions for Fingerprint Minutiae (Tulyakov et al, 2007; Jain et al, 2006). They presented a method of hashing fingerprint minutia information and performing fingerprint identification in

hashing space. Due to the disorder of templates minutiae, input of hash function was not dependent on sequence (i.e. symmetric). Specifically, given n minutia points $\{c_1, c_2, \ldots, c_n\}$, they constructed following *m* symmetric hash functions and employed one or some of them:

$$
\begin{aligned}
\mathrm{h}_1(c_1, c_2, \cdots, c_n) &= c_1 + c_2 + \cdots + c_n \\
\mathrm{h}_2(c_1, c_2, \cdots, c_n) &= c_1^2 + c_2^2 + \cdots + c_n^2 \\
&\cdots \\
\mathrm{h}_m(c_1, c_2, \cdots, c_n) &= c_1^m + c_2^m + \cdots + c_n^m
\end{aligned}
\tag{2}
$$

where $c_i$ ( $i = 1, 2, \ldots, n$) are complex numbers, represent the information of minutiae structure.

They spread the concept of two factor authentication using key binding method. In order to enhance the security, they establish random relationship between a class of hash function and pair of minutiae structure by a particular user's key, so different user has different relationship between hash function and pair of minutiae structure.

### 2.3.2 Fuzzy commitment scheme

Juels & Wattenberg (1999) proposed a fuzzy commitment scheme. The early theoretical research combined well-known techniques from the areas of error-correcting codes and cryptography to achieve a typical key binding scheme. Actually, this scheme derived from bit commitment scheme of cryptography, and follows the concept of commitment and witness and uses them for the inherently fuzzy biometric data. Fuzzy commitment scheme *F* has two sections: commitment and decommitment. In terms of commitment, *F* shall be constructed so as to commit an error-correcting codeword *c* using a witness *x*, where both *c* and *x* are *n*-bit strings. In biometric scenarios, *x* typically represents a biometric template, such as a fingerprint. The codeword *c* represents a secret key protected under this template. Deviation $\delta = x\text{-}c$, so commit: $\{\mathbf{hash}(c), \delta\}$, where $\mathrm{hash}(\bullet)$ is hash function. While consider the decommitment, user input a biometric vector $x'$, a secret $c'$ can unlocked from commitment according the formula: $c' = x' - \delta = x' - x + c$. If *x* is very closed to $x'$ in a certain distance (i.e. Hamming distance), $c'$ can be considered to be identical to *c*, as well as verification of $\mathbf{hash}(c')$ and hash(c) , and thus achieve the authentication.

Based on the fuzzy commitment scheme, Hao et al. (2006) designed and implemented an iris encryption scheme. Compared to the fingerprint, iris is more suitable for the search of encryption because IrisCode is more canonical in coding. IrisCode has a fixed length of 2048-bit, together with some encryption algorithm to generate immediately, and the encryption and decryption is very easy to operate.

### 2.3.3 Fuzzy vault scheme

Juels & Sudan (2002) presented the fuzzy vault scheme on the foundation of fuzzy commitment scheme. The most valued characteristic of the algorithm is linking the fuzziness of biometric with accuracy of cryptography perfectly.

The detailed implementation of the algorithm can be described as follows:

a.  "Lock" vault: Alice aims to lock a secret *K* under an unordered set A. She selects a polynomial p in a single variable *x* such that p encodes *K* in some way and computes the p (A), projection of A lying on the polynomial p, thus form a finite point set (A, p

(A)). She then creates a number of random chaff points, with point set (A, p (A)) constitute the Vault

b. "Unlock" vault: Suppose now that Bob wishes to unlock $K$ by means of an unordered set B. If B overlaps substantially with A, then B identifies many points in R that lie on polynomial p. Using error correction, he is able to reconstruct p exactly and thereby $K$. If B does not overlap substantially with A, then it is infeasible for Bob to learn $K$, because of the presence of many chaff points.

Based on the work of Juels et al, Clancy et al. (2003) advanced the conception of fingerprint vault. Firstly, use user's five fingerprints to register, extract position of minutiae as input, manage correspondence problem between fingerprint features by nearest neighbor algorithm. In considering the size of fingerprint pressing region, author add $N$ chaff points to the minutiae set, where the distance of chaff points to the minutiae and the distance between chaff points themselves aren't smaller than $d$, thus form the encrypted fingerprint vault. Being different from Juels et al, Clancy et al. describes the order of fingerprint polynomial in detail. Considering the decryption, using the nearest neighbor algorithm for extracted minutiae feature from matching fingerprint, search out the corresponding points in fingerprint vault, then take the points as input of RS correction code algorithm to compute the correct form of encrypted polynomials. The work contributes to describe the implementation method of fuzzy vault in the field of fingerprint in detail, achieve 69-bit security on the basis of 20% to 30% of the rejection. While like reference (Davida et al, 1998), the drawback is the corresponding pre-registration fingerprint image which the authors assume.

Uludag et al. (2005) presented a more practical scheme named Fuzzy Vault for Fingerprint on the basis of Fuzzy Vault and Fingerprint Vault. Nandakumar et al. (2007) notice that since the fuzzy vault stores only a transformed version of the template, aligning the query fingerprint with the template is a challenging task. So they propose the idea that add a password to the periphery of fuzzy vault system, and it is deformed minutiae parameter that are stored in new template but original data, where the deformed parameter is correlated to the user set-up password. Encryption mechanism is independent on the security of fuzzy vault, so system is under double protection and attacker can take the legality user data only by breaching two systems in the one time. Compared to ordinary fuzzy vault system, enhanced system has a higher rejection rate, but the cost is enhanced algorithm time complexity.

Gradually fuzzy vault is extended to other biometric (Nyang & Lee, 2007; Wang & Plataniotis, 2008; Lee, Y, 2007). Nyang & Lee (2007) show how can fuzzy vault be introduced to the weighted principal component analysis (PCA) of face, and introduce a so-called intermediate layer so that more points heavy weighted feature construct, at the same time, hash the feature and corresponding construction data using the SHA-1 function, whereas there is no concrete experimental validation. The PCA features of face are mapped into binary data with two random orthonormal matrixes ($R_1$, $R_2$), the result is some binary features in the 16-bit length and used for the encoding and decoding of fuzzy vault (Wang & Plataniotis, 2008). Lee, Y (2007) proposes a new method of applying iris data to the fuzzy vault. The author obtains 16 27-bit length iris features by the methods of independent component analysis (ICA)-based feature extraction and K-means cluster pattern. Experiment on the database BERC iris, which have 99×10=990 iris images, constituted by author. Zero FAR and about 0.775% FRR are obtained.

Fuzzy Vault has become one of the most potential methods on biometric template protection technology. With the gradually abroad research and application of it, some researchers attend the corresponding attacks strategy (Scheirer & Boult, 2007; Kholmatov & Yanikoglu, 2008; Mihailescu, 2007; Chang, 2006). Scheirer & Boult (2007) review briefly some of the known attacks against biometric fuzzy vault (BFV) and biometric encryption (BE) techniques, including attack via record multiplicity, surreptitious key-inversion attack, and novel blended substitution attacks. And apply each of these attacks on the Fuzzy Vault and biometric encryption system. Kholmatov & Yanikoglu (2008) implemented attack via record multiplicity using 200×2+400 fingerprints and can correlate 59% of vaults approving the claim of fuzzy vault's vulnerability against attack by comparison between two vaults from same finger, which show that the fuzzy vault is threatened by attack via record multiplicity on the ratio more than 50%, the ratio will increase when there are three or more correlated vaults. Mihailescu (2007) proved that the system is vulnerable to the brute force attack and also gave several suggestions which can improve the fingerprint vault to a cryptographically secure algorithm by mathematic analysis. Chang (2006) thought that genuine minutiae can be distinguished from chaff points by statistical characteristics of all points, actually chaff points tend to concentrate, they proved that the genuine minutiae can be found in much less searching time than force attack in the means of mathematic analysis and experimental validation. All of these attack are based on the fact that the vault contain genuine minutiae data, in other words, there is definitely entropy loss. So, these attacks will have no entry point if those genuine minutiae are not stored in vault by some certain transformation.

### 2.3.4 Fuzzy extractor

Dodis et al. (2004) proposed a concept of secure sketch and fuzzy extractor, aimed to achieve reliable and secure authentication to user, they attempt to convert random biometric signal into stable key which can be used in encryption. Some certain information of secure sketch can be extracted from biometric signal by the operation that can tolerate error in a certain degree. The published information can reconstruct original template perfectly while signal similar with original template is input. Meanwhile, the linchpin of the method is that the original template cannot be reconstructed by the republished information. Fuzzy extractor extracts approximate uniformly distributed random signal **R** from the input biometric signal, so **R** can be applied as a Key to all of the encryption.

In order to construct concrete algorithm for various biometric signal, Dodis et al. make use of three measure spaces, such as hamming distance, set distance, and edit distance. In the space of hamming distance, Dodis et al. view fuzzy commitment (Jin et al, 2007) as optimal secure sketch, and reform it into approximate optimal fuzzy extractor using general construction method. In the space of set distance, they view fuzzy vault as approximate optimal secure sketch, and reform it into approximate optimal fuzzy extractor using same construction method. In the space of edit distance, they define the transformation from edit space to set space in order to transform optimal fuzzy extractor of set space into edit space. Also, authors prove that the optimal secure sketch and fuzzy extractor can be constructed if entropy loss satisfies some certain condition.

Literatures (Dodis et al, 2006; Buhan et al, 2007; Boyen, 2004; Boyen et al, 2005; Li, Q et al, 2006; Sutcu, 2007) contribute to the study of key generation method. Literatures (Tong et al, 2007; Arakala et al, 2007) extract robust key respectively from feature of fingercode and

feature of minutiae structure, and progress attempt of practical algorithm. Although the result isn't ideal, they contribute exploratively to the research of the issue. Literature (Zhang et al, 2008) actualizes iris-based fuzzy extractor, analyzes the influence on the performance of identification of difference between iris feature codes, and designs two layer cascade error-correcting scheme in which iterative codes and Reed-Solomon codes are applied.

### 2.4 Hidden transmission of biometric template

Khan et al. (2007) presented a chaotic secure content-based hidden transmission scheme of biometric data. Encryption and data hiding techniques are used to improve the security and secrecy of the transmitted templates. Secret keys are generated by the biometric image and used as the parameter value and initial condition of chaotic map, and each transaction session has different secret keys to protect from the attacks. Two chaotic maps are incorporated for the encryption to improve the system's resistance against attacks. Encryption is applied on the biometric templates before hiding into the cover/host images to make them secure, and then templates are hidden into the cover image. Experimental results show that the security, performance, and accuracy of the presented scheme are encouraging comparable with other methods found in the current literature. In 2010, Khan et al. proposed another means of hidden biometric template transmission named chaos and NDFT-based spread spectrum technique to conceal fingerprint-biometrics templates into audio signals. Fingerprint templates are encrypted by chaotic encryption, encoded by the BCH codes, modulated by chaotic parameter modulation (CPM), and then hid into the chaotically selected random sampling points of the host speech signal by non-uniform discrete Fourier transform (NDFT). The template extraction process is completely blind and does not require original speech signal, thus the extraction depends on the secret key. Experimental and simulation results show that the scheme is robust against common signal processing attacks, and accomplishes perceptual transparency by exploiting the masking effects of human auditory system (HAS).

## 3. The biometric template protection with secure authentication scheme based on fuzzy extractor and chaotic spread spectrum encryption

In this section, a biometric template protection scheme based on fuzzy extractor for biometric authentication is proposed. Instead of only using one layer error-correcting code (ECC) or two cascaded ECCs in published literatures, a ECC followed by chaotic spread spectrum encryption is utilized in our scheme. The scheme is evaluated using 160 4095-bit fingerprint codes from 20 different fingers, with 8 samples for each finger. Simulation experiments show that both security and privacy of biometric template can be effectively protected.

### 3.1 Chaotic spread spectrum encryption using coupled n-NDFs

Since the intra-class variance among the samples from same finger may achieve to 25%-30%, the chaotic spread spectrum encryption technique, instead of ECC, is used here to improve the error-correcting ability, with attendant encryption function. In the following subsection, $n$-dimensional nonlinear digital filter ($n$-NDF) is preferred to serve as the underlying chaotic system to produce secure spread spectrum code.

### 3.1.1 Chaotic spread spectrum code base on *n*-dimensional NDF

Nonlinear digital filters (NDFs) have received attention in chaotic secure communication, hash function and pseudorandom bit generator. The reason is that the $n$-NDF outputs $n$-dimensional uniform distributed chaotic signal when it satisfies Kelber conditions (Wang & Zhang, 2007). Fig.5. depicts the block diagram of an $n$th-order NDF, whose state equation is given by

$$\begin{cases} z_1(t+1) = h \circ \mathrm{mod}(\sum_{i=1}^{n} c_i\, z_i(t) + \phi) \\ \quad z_q(t+1) = z_{q-1}(t) \qquad\qquad , q = 2,3,\cdots,n \\ \qquad y(t) = z_1(t+1) \end{cases} \qquad (3)$$

where $\phi \in (-1,1)$ denotes input signal, y(t) the output signal, $z = \{z_1, z_2, \cdots, z_n\}^T \in (-1,1)^n$ the initial states of filter, $\mathbf{c} = \{c_1, c_2, \cdots c_n\}$ the filter coefficients, $h(.)$ the piecewise linear map defined by

$$h(x,p) = \begin{cases} (2x+1-p)/(1+p) & x \in (-1,p] \\ (-2x+1+p)/(1-p) & x \in (p,1) \end{cases}, \text{ and } \quad mod(v) = v - 2 \cdot \left\lfloor \frac{v+1}{2} \right\rfloor.$$

For describing convenience, the discretization form of $n$-NDF above is denoted as $y(i+1) = \mathrm{F}(\phi, \mathbf{z}, \mathbf{c}, i)$. It has been proven that $n$-NDF is an ergodic chaotic system with $n$-D uniform distribution provided that the system is not decomposable and the coefficients $c_n \in Z, |c_n| > 1, c_i \neq 0, i \in \{1, 2, ..., n-1\}$.
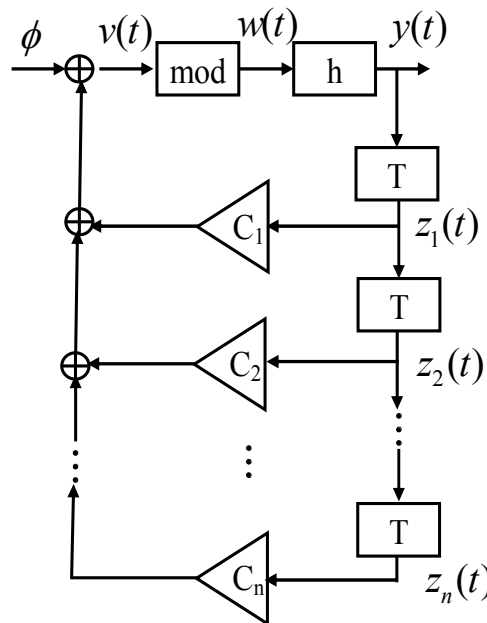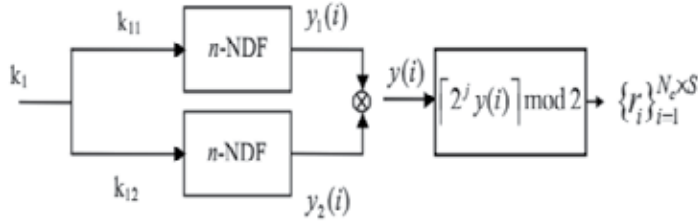


Fig. 5. Block diagram of the $n$th-order NDF.

Fig. 6. Generating chaotic spread spectrum sequence by coupling two $n$-NDFs.

In the following, we couple two independent $n$-NDFs, as depicted in Fig. 6, to generate chaotic spread spectrum sequence. The two independent $n$-NDFs are expressed as

$$\begin{cases} y_1(i+1) = F_1(\phi_1, \mathbf{z}_1, \mathbf{c}_1, i) \\ y_2(i+1) = F_2(\phi_2, \mathbf{z}_2, \mathbf{c}_2, i) \end{cases}, \ y_1, y_2 \in (-1, 1) \tag{4}$$

Then couple two outputs of Eq.(4) as $y(i) = mod(y_1(i) + y_2(i))$ ( The symbol " $\otimes$ " in Fig.6.), and quantize y(i) uniformly to get the binary spread spectrum sequence $r_i = \lceil 2y(i) \rceil \mathrm{mod}\, 2$.

### 3.1.2 Chaotic spread spectrum encryption

Figure 7 shows that the process of chaotic spread spectrum encryption is with the encryped operation XOR, at the same time with code spectrum spread. Specifically, under the control of key $k_1$, chaotic spread spectrum sequence $\{r_i\}_{i=1}^{N_c \times S}$ can be obtained, then XOR it with each error correction encoded binary code $c_j (j = 1, \cdots N_c)$, the result $w_i = r_i \oplus c_{\lceil i/s \rceil}$ is the spreading encryption information corresponding to $C = \{c_j\}_{j=1}^{N_c}$.



Fig. 7. The process of chaotic spread spectrum operation.
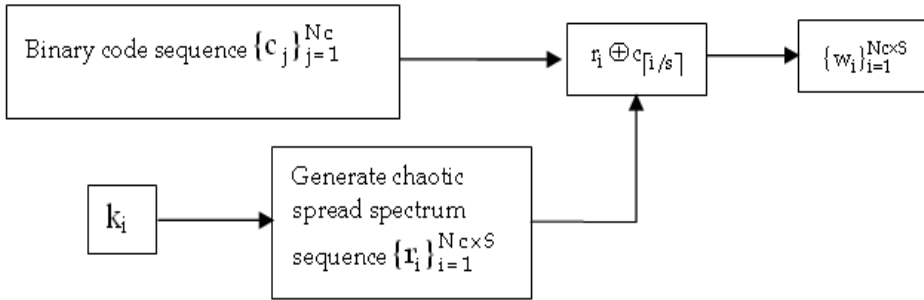
Based on the chaotic spread spectrum sequence $r_i$, the process of chaotic spread spectrum encryption is defined as

$$\begin{aligned} w &= \{w_i\}_{i=1}^{N_w} \\ &= \{r_i \oplus c_{\lceil i/s \rceil}\}_{i=1}^{N_C \times S} \\ &= \{c_j \oplus r_{(j-1) \times S+1}, c_j \oplus r_{(j-1) \times S+2}, \cdots, c_j \oplus r_{(j-1) \times S+S}\}_{j=1}^{N_c} \end{aligned} \tag{5}$$

where symbol "$\oplus$" denotes bit-XOR operation, **S** is spread factor, **Nc** the bit-length of original message, $r_i$ the spread spectrum sequence, and $w$ the spreaded sequence with bit-length $N_w = S \times N_c$. With the increasing **S**, the error correction capability can also be improved. The critical work is to decide a suitable **S** by experiments to discriminate the intra-class samples and inter-class samples.

Regarding the de-spread spectrum, it is the inverse process of Fig.7. Assume the spread information is $w^* = \{w_j^*\}_{j=1}^{N_w}$, corresponding to the original message $w$, the de-spread process is composed of correlation and decision phases defined by Eq.(6) and Eq.(7), respectively. The $c_j^*$ in Eq.(7) is the recovered binary code sequence corresponding to $c_j$.

$$d = \{d_j\}_{j=1}^{N_c}$$
$$= \{\sum_{i=1}^{s} w_{(j-1) \times S+i}^* \oplus r_{(j-1) \times S+i}\}_{j=1}^{N_c} \tag{6}$$

$$c_j^* = \begin{cases} 0 & d_j < S/2 \\ 1 & d_j \geq S/2 \end{cases} \tag{7}$$

## 3.2 The proposed biometric template protection scheme

The way of centralized storage of biometric data in the database have security guarantees by using the chaotic $n$-NDF, where the hash value $H(R)$ of random secret information **R** instead of biometric $w_0$ itself stored in the database, can play the same protection effect as a password on authentication system. Given that the one-way hash function $H(.)$ is safe and collision free, the proposed scheme is a safe fingerprint identification system.

The proposed scheme includes two stages: registration and authentication. In the stage of registration, $l$-bit random number $R$ was selected first, and then carry out BCH encoding operation on it and R′ is obtained. Next, perfrom chaotic spread spectrum on R′ to get sequence R″. At the same time $w_0$′ is reached from the user's fingerprint code $w_0$ after BCH decoding operation on the $w_0$, then publish $pub$=R″ $\oplus w_0$′. The stage of authentication is the recovery process of R. Suppose $w_1$ is the fingerprint code what is to be authenticated, similarly the $w_1$′ is the data obtained from the BCH decoding on the $w_1$, as $pub \oplus w_1$′= (R″ $\oplus w_0$′) $\oplus w_1$′=R″ $\oplus (w_0$′ $\oplus w_1$′), while $w_0$′ $\oplus w_1$′can be viewd as noise which disturbes R″. The registration and identification process can be seen as that R passes an additive noise channel of digital communication system. Similar fingerprint feature code have less different bits equivalently less noise, while the different fingerprint feature code have more different bits, resulting in greater noise. When the R″ is disturbed by noise P, through the appropriate error-correcting code that R can be recovered when similar fingerprint feature is authenticated while different fingerprint feature can not. Assume R be recovered as $R_1$, the authentication is valid or not depending on whether the hash value of $R_1$ equals the pre-stored hash value H ($R$) or not.

Utilizing the ECC, chaotic spread spectrum and fuzzy extractor, the proposed scheme consists of registration process and authentication process, which is illustrated in Fig.8. and described as follows.

### Registration process

User's fingerprint data is collected firstly in the registration phase, and carry out features extraction and coding, calculate R and *pub* from the BCH decoding of fingerprint template $w_0'$, where R is the secret random number and *pub* is public data. H(R) is calculated by the one-way hash function. R, H(R) and *pub* are stored in server database, thus complete the registration.

1. Randomly select a secret R and perform BCH encoding: $R' \leftarrow$ BCH(R);
2. Perform chaotic spread spectrum operation on R': $R'' \leftarrow$ Chaotic_SS(R');
3. To decrease the distance of intra-class samples, perform BCH decoding on the user's biometric template $w_0$: $w_0' \leftarrow$ De_BCH($w_0$);
4. Perform bit-XOR operation on R'' and $w_0'$ to get public information: $pub \leftarrow R'' \oplus w_0'$;
5. Store *pub* and Hash value of R on server for user authentication: $server \leftarrow \{pub, H(R)\}$, where H(.) is a cryptographic hash function.

### Authentication process

In the authentication phase the user's fingerprint information is collected and the fingerprint feature is denoted as $w_1$. The authentication process is as follows.

1. extract the user's fingerprint template $w_1$ and execute BCH decoding on it: $w_1' \leftarrow$ De_BCH($w_1$);
2. retrieve the *pub* information from server, and bit–XOR it with $w_1'$: $R_1'' \leftarrow w_1' \oplus pub$;
3. perform chaotic de-spread spectrum operation on $R_1''$: $R_1' \leftarrow$ Chaotic_DS($R_1''$);
4. perform BCH decoding on $R_1'$: $R_1 \leftarrow$ De_BCH($R_1'$);
5. match the hash value of $R_1$ and the hash value of R stored in server, if $H(R_1)==H(R)$, the user is authenticated, otherwise, the user is rejected.
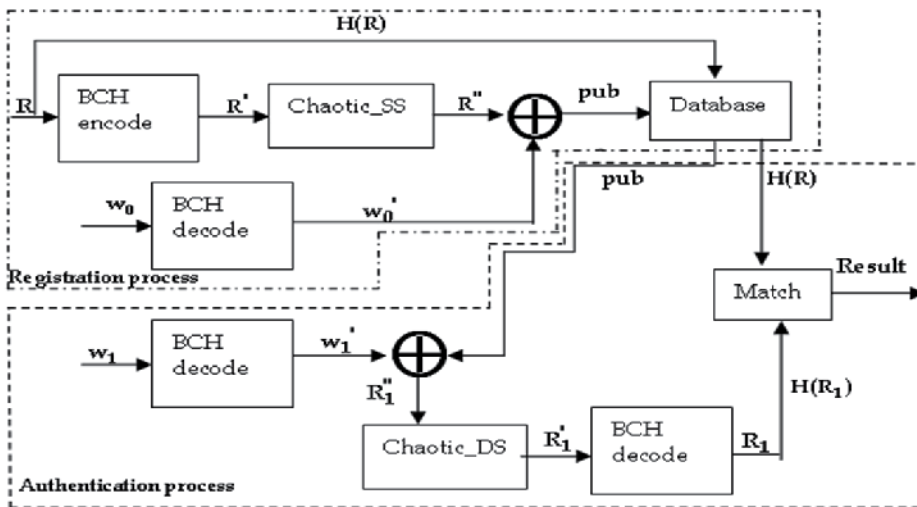


Fig. 8. Block diagram of the proposed scheme.

Note that two minor things in registration and authentication phases have to be processed. One is how to initialize the initial states and coefficients of coupled *n*-NDFs in chaotic spread/de-spread spectrum process. This can optionally split H(R) into 32-bit strings for

each state and coefficient. If the length of H(R) is not enough long, we can hashing R one more times until the total hash length meets requirement. The other is bit-XOR operations in $pub \leftarrow R'' \oplus w_0'$ and $R_1'' \leftarrow w_1' \oplus pub$, where two operands are required to be identical bit length, otherwise bit-expansion is necessary. That is, if the bit length of $w_0'$ is smaller than that of $R''$, repeatedly concatenate $w_0'$ so that its length is enough long. Otherwise, trim $w_0'$ so that its length equals to that of $R''$. As for $R_1'' \leftarrow w_1' \oplus pub$, the way of processing is similar.

### 3.3 Security analysis

In this subsection, we will briefly illustrate the privacy protection and cancellable ability of proposed scheme.

**Privacy protection:** Early biometric-based authentication systems directly store user's biometric templates in server, this way may cause template disclosing by database manager or hacker, even the templates are stored in smart card. In the proposed scheme, only H(R) and *pub* are stored in server. Since H($\cdot$) is one-way cryptographic hash function, it's computationally infeasible to recover R. Moreover, R is randomly selected by authenticated user, the attacker can not derive R and biometric template from H(R) and *pub*. Therefore, the proposed scheme has strong privacy protection.

**Template cancellation:** The template cancellation of proposed scheme is different from traditional template cancellations, but in fact it can achieve to the purpose of "template cancellation". In this scheme, on the one hand, users select different random secret R for different application systems, and thus different systems stored different information H(R) and *pub*. This way adversary can not obtain any secret information R or biometric template of a user, though they collect all the stored information of the same user from multiple authentication systems. On the other hand, when user's register information requires update, user only need reselect random secret information $R_{new}$ and calculate H($R_{new}$) and $pub_{new}$. After re-registering, the old information H($R_{old}$) and $pub_{old}$ are not valid any more. Moreover, it is not conductive to derive the user's biometric template from the newly registered information H($R_{new}$) and $pub_{new}$, even when attacker got the H($R_{old}$) and $pub_{old}$. Therefore, the multiple re-registering information from the same user does not decrease the security. From system function point of view, the proposed scheme inherently owns revocable-biometric ability.

### 3.4 Experimental results

The proposed method is evaluated using the fingerprint database of FVC 2004 [FVC 2004], where there are 8 impressions for each of the 100 distinct fingers with image size of 328x364 at a resolution of 500dpi.

We select 8 impressions for each of the 20 distinct fingers. Among these fingerprint images, 60 images for 20 fingers (each finger has 3 images) are used to parameter tuning before testing, while the rest fingerprint images are used to evaluate the scheme. Fig.9 shows 3 images of one finger of 20 fingers. The evaluation criteria used here are fault accept rate (FAR) and fault reject rate (FRR).

Firstly, we use 60 images for parameter optimizing. There are two parameters (i.e. *n*, *k*) in BCH error-correcting code, and one parameter (i.e. spread factor S) in chaotic spread spectrum. The optimization target is balancing the FRR for intra-class samples, the FAR for

inter-class samples and computational load. Based on such optimization principle, one of the tuning parameter set are valued as $n$=63, $k$=10 and spread factor S=40.



Fig. 9. Three images of one finger of 20 fingers for parameter tuning.

In the rest 100 samples, we select 2 samples from the rest 5 samples of each finger, that one sample is used to registration while the other is used to authentication. We perform such intra-class experiments for $20 \times C_5^2 = 200$ times. The experiment result is listed in table 1. The data of table 1 shows that the FRR=0.5% and GAR=99.5% in the scheme. When we improve the error-correcting capability by increasing the spread factor or BCH parameters, the FRR will decrease as expected at the cost of time complexity and storing volume.

| parameters | Right accept number | False refuse number | FRR |
|---|---|---|---|
| N=63, k=10, spread factor=40 | 199 | 1 | 0.5% |

Table 1. FRR experiment result for intra-class samples

In addition, we randomly select 2 inter-class samples from the rest 100 samples to evaluate the FAR. Fig.10 shows one experimented group of that. Such experiments are performed for $C_{100}^2 - 20 \times C_5^2 = 4750$ times with the same parameters as table 1, and the statistical result is summarized in table 2.



Fig. 10. One experimented group for the FAR evaluation.

| parameters | Right refuse number | False accept number | FAR |
|---|---|---|---|
| N=63, k=10, spread factor=40 | 4750 | 0 | 0 |

Table 2. FAR experiment result for inter-class samples

The inter-class experiments show that no fingerprint sample has been accepted by fault, i.e. the FAR=0. It should not be surprise for such result, because the difference of two inter-class samples is so large that exceeds the error-correcting capability of BCH and spread spectrum under the selected parameters.

From the experimental FRR and FAR index of the proposed scheme, it can be seen that the scheme has high right accept rate for the intra-class fingerprints while keep ideal fault accept rate for the inter-class fingerprints. Of course, the above experiments are not enough to test the scheme and come to final conclusion. More samples, more kinds of biometrics and great number of experiments are necessary to evaluate the biometric system.

## 4. Conclusion

In this chapter, we have presented a biometric template protection scheme based on fuzzy extractor for biometric authentication. Instead of only using one layer error-correcting code (ECC) or two cascaded ECCs in published literatures, an ECC followed by chaotic spread spectrum encryption is utilized in this scheme. We performed a series of experiments to evaluate the performance of the system and the experimental results show that the proposed system is robust against noises and attacks. Moreover, the proposed system can be easily realized in the real biometric applications.

## 5. References

Ang, R. Rei, S. & Luke, M. (2005). Cancellable key-based fingerprint templates, In: *Information Security and Privacy,* Boyd, C. & Nieto, J, pp. 242−252, Springer Berlin, ISBN 978-3-540-26547-4, Heidelberg, Germany

Arakala, A. Jeffers, J. & Horadam, K. (2008). Fuzzy extractors for minutiae-based fingerprint authentication. In: *Proceedings of the ICB 2007,* Lee SW, Li SZ, pp.760−769, Springer Berlin, ISBN 978-3-540-74548-8, Heidelberg, Germany

Boyen, X. (2004). Reusable cryptographic fuzzy extractors, *Proceedings of The Conference on Computer and Communications Security,* ISBN 1-58113-961-6, Washington DC, USA, Octorber 2004

Boyen, X. Dodis, Y. Katz, J. Ostrovsky, R. & Smith, A. (2005). Secure remote authentication using biometric data. In: *Advances in Cryptology â€" EUROCRYPT 2005,* Cramer, R, pp. 147−163, Springer Berlin, ISBN 978-3-540-25910-7 , Heidelberg, Germany

Buhan, I. Doumen, J. Hartel, P. & Veldhuis, R. (2007). Fuzzy extractors for continuous distributions, *Proceedings of The Conference on Computer and Communications Security,* ISBN 1-59593-574-6, Singapore , March 2007

Cappelli, R. Lumini, A. Daio, D. & Maltoni D. (2007). Fingerprint image reconstruction from standard templates. *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol.29, No.9, (September 2007), pp.1489−1503, ISSN 0162-8828

Chang, E. Shen, R. & Teo, F. (2006). Finding the original point set hidden among chaff, *Proceedings of Conference on Computer and Communications Security,* ISBN 1-59593-272-0, Taipei, China, March 2006

Clancy, T. Kiyavash, N. & Lin, D. (2003). Secure smartcard-based fingerprint authentication, *Proceedings of the ACMSIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop. Association for Computing Machinery*, (November 2003), pp. 45−52, ISSN 1-58113-779-6

Connie, T. Jin, A. Goh, A. & Ling, D. (2004). PalmHashing: A novel approach for dual-factor *authentication. Pattern Analysis & Applications,* Vol.7, No.3, (August 2004), pp. 255−268, ISSN 1433-7541

Davida, G. Frankel, Y. & Matt, B. (1998). On enabling secure applications through off-line biometric identification, *Proceedings of the IEEE Symposium on Security and Privacy*, ISBN 0-8186-8386-4, Oakland, May 1998

Dodis, Y. Reyzin, L. & Smith, A. (2004). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In: *Advances in Cryptology - EUROCRYPT 2004,* Cachin, C. & Camenisch, J, pp.523−540, Springer Berlin, ISBN 978-3-540-21935-4, Heidelberg, Germany

Dodis, Y. Katz, J. Reyzin, L, Smith A. (2006). Robust fuzzy extractors and authenticated key agreement from close secrets. *Advances in Cryptology-Crypto*, Vol.4117, (2006), pp.232−250, ISSN 0302-9743

FVC2004 http://bias.csr.unibo.it/fvc2004

Hao, F. Anderson, R. & Daugman, J. (2006). Combining crypto with biometrics effectively. I*EEE Trans on Computers*, Vol.55, No.9, (September 2006), pp. 1081−1088, ISSN 0018-9340

Jain, A. Prabhakar, S. Hong, L. & Pankanti, S. (1999). FingerCode: A filterbank for fingerprint representation and matching. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, ISBN 0-7695-0149-4, Fort Collins, CO, June 1999

Jain, A. Ross, A. & Pankanti, S. (2006). Biometrics: A tool for information security. *IEEE Trans on Information Forensics and Security*, Vol.1, No.2, (June 2006), pp. 125−143, ISSN 1556-6013

Jin, A. Ling, D. & Goh, A. (2004). An integrated dual factor authenticator based on the face data and tokenised random number, In: *Biometric Authentication*,  Zhang, D. & Jain, A. pp. 117−123, Springer Berlin, ISBN 978-3-540-22146-3, Heidelberg, Germany

Jin, A. Ling, D. & Goh, A. (2004). Personalised cryptographic key generation based on FaceHashing. *Computers & Security*, Vol.23, No.7, (October 2004), pp. 606−614, ISSN 01674048

Jin, A. Ling, D. & Goh, A. (2004). Biohashing: Two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition*, Vol.37, No.11, (November 2004), pp.2245−2255, ISSN 0031-3203

Jin, A. & Ling, D. (2005). Cancellable biometerics featuring with tokenised random number. *Pattern Recognition Letters*, Vol.26, No.10, (July 2005), pp.1454−1460, ISSN 01678655

Jin, A. Goh, A. & Ling, D. (2006). Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol.28, No.12, (December 2006), pp.1892−1901, ISSN 0162-8828

Jin, A. Toh, K. & Yip, W. (2007). 2^N Discretisation of biophasor in cancellable biometrics. In: *Advances in Biomet*rics, Lee, S. & Li, S, pp. 435−444, Springer Berlin, ISBN 978-3-540-74548-8, Heidelberg, Germany

Jin, A. Yip, W. & Lee, S. (2008). Cancellable biometrics and annotations on BioHash. P*attern Recognition*, Vol.41, No.6, (June 2008), pp. 2034−2044, ISSN 00313203

Juels, A. & Wattenberg, M. (1999). A fuzzy commitment scheme, *Proceedings of the 6th ACM conference on Computer and communications security,* ISBN 1-58113-148-8, Singapore, November 1999

Juels, A. & Sudan, M. (2002). A fuzzy vault scheme, *Proceedings of the 2002 IEEE International Symposium on Information Theory,* (2002), pp.408

Khan, MK. Zhang, JS. & Tian, L. (2007). Chaotic secure content-based hidden transmission of biometrics templates. *Chaos, Solitons, and Fractals,* Vol.32, No.5, (June 2007), pp. 1749–1759, ISSN 09600779

Khan, MK. Xie, L. & Zhang, JS. (2010). Chaos and NDFT-based concealing of fingerprint biometric data into audio signals for trustworthy person authentication. *Digital Signal Processing: A Review Journal,* Vol.20, No.1, (January 2010), pp. 179–190, ISSN 10512004

Khan, MK. Zhang, JS. Wang, XM. (2008). Chaotic Hash-based Fingerprint Biometric Remote User Authentication Scheme on Mobile Devices, *Chaos, Solitons and Fractals,* vol.35, No.3, (2008), pp.519-524, ISSN 09600779

Kholmatov, A. & Yanikoglu, B. (2008). Realization of correlation attack against the fuzzy vault scheme, *Proceedings of SPIE - The International Society for Optical Engineering,* ISBN 9780819469915, San Jose, CA, United states, January 2008

Lee, C. Choi, J. Toh, K. Lee, S. & Kim, J. (2007). Alignment-Free cancelable fingerprint templates based on local minutiae information. *IEEE Trans on Systems, Man, and Cybernetics, Part B: Cybernetics,* Vol.37, No.4, (August 2007), pp. 980−992, ISSN 1083-4419

Lee, Y. Bae, K. Lee, S. Park, K. & Kim, J. (2007). Biometric key binding: Fuzzy vault based on iris images. In: *International Conference on Advances in Biometrics, Proceedings of the ICB 2007,* Lee, S. & Li, S, pp. 800−808, Springer Berlin, ISBN 978-3-540-74548-8, Heidelberg, Germany

Li, Q. Sutcu, Y. & Memon, N. (2006). Secure sketch for biometric templates. In: *Advances in Cryptology - ASIACRYPT 2006,* Lai, XJ. & Chen, KF, pp. 99−113, Springer Berlin, ISBN 978-3-540-49475-1, Heidelberg, Germany

Li, P. Tian, J. Yang, X. Shi, P. & Zhang, YY. (2009).Biometric Template Protection. *Journal of Software,* Vol.20, No.6,2009,(June 2009), pp.1553−1573

Ling, D. Jin, A. & Goh, A. (2004). Eigenspace-Based face hashing. In: *Biometric Authentication,* Zhang, D. & Jain, A. pp. 195−199, Springer Berlin, ISBN 978-3-540-22146-3, Heidelberg, Germany

Ling, D. Jin, A. & Goh, A. (2006). Biometric Hash: High-Confidence face recognition. *IEEE Trans on Circuits And Systems for Video Technology,* Vol.16, No.6, (June 2006), pp. 771−775, ISSN 1051-8215

Lumini, A. & Nanni, L. (2006). An advanced multi-modal method for human authentication featuring biometrics data and tokenised random numbers. *Neurocomputing,* Vol.69, No.13-15, (August 2006), pp. 1706−1710, ISSN 09252312

Lumini, A. & Nanni, L. (2007). An improved BioHashing for human authentication. *Pattern Recognition,* Vol.40, No.3, (March 2007), pp.1057−1065, ISSN 0031-3203

Maio, D. & Nanni, L. (2005). Multihashing, human authentication featuring biometrics data and tokenized random number: A case study FVC2004. *Neurocomputing,* Vol.69, No.1-3, (December 2005), pp. 242−249, ISSN 09252312

Mihailescu, P. (2007). The fuzzy vault for fingerprints is vulnerable to brute force attack,In: *Computer Vision and Pattern Recognition,* 22.08.2007, Available from:

http://arxiv.org/abs/0708.2974v1

Nandakumar, K. Jain, A. & Pankanti, S. (2007). Fingerprint-Based fuzzy vault: Implementation and performance. *IEEE Trans on Information Forensics and Security*, Vol.2, No.4, (November 2007), pp. 744−757, ISSN 1556-6013

Nanni, L. & Lumini, A. (2006). Empirical tests on BioHashing. *Neurocomputing*, Vol.69, No.16-18, (Octorber 2006), pp.2390−2395, ISSN 09252312

Nanni, L. & Lumini, A. (2008). Random subspace for an improved BioHashing for face authentication. *Pattern Recognition Letters*, Vol.29, No.3, (February 2008), pp. 295−300, ISSN 01678655

Nyang, D. & Lee, K. (2007). Fuzzy Face Vault. How to implement fuzzy vault with weighted features. In: *Proceedings of the Universal Access in HCI,(HCII 2007)*, Stephanidis, C, pp.491-496, Springer Berlin, ISBN 978-3-540-73278-5, Heidelberg, Germany

Ratha, N. Connell, J. & Bolle RM. (2001).An analysis of minutiae matching strength, In: *Audio and Video-Based Biometric Person Authentication,* Bigun, J. & Smeraldi, F, pp. 223−228, Springer Berlin, ISBN 978-3-540-42216-7, Heidelberg, Germany

Ratha, N. Connell, J. Bolle, R. & Chikkerur, S. (2006). Cancelable biometrics: A case study in fingerprints, *Proceedinds of the 18th Int'l Conf. on Pattern Recognition (ICPR 2006)*, ISBN 1051-4651, HongKong, September 2006

Ratha, N. Chikkerur, S. Connell, J. & Bolle, R. (2007). Generating cancelable fingerprint templates. *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol.29, No.4, (April 2007), pp.561−572, ISSN 0162-8828

Scheirer, W. & Boult, T. (2007). Cracking fuzzy vaults and biometric encryption, *Proceedings of Biometrics Symposium*, ISBN 978-1-4244-1549-6, Colorado, USA, September 2007

Soutar, C. Roberge, D. Stoianov, A. Gilroy, R. & Vijaya, K. (1999). Biometric encryption,In: *ICSA Guide to Cryptography*, McGraw-Hill, Available from

http://www.bioscrypt.com/assets/Biometric_Encryption.pdf

Sutcu, Y. Li, Q. & Memon, N. (2007). Protecting biometric templates with sketch: Theory and practice. *IEEE Trans on Information Forensics and Security*, Vol.2, No.3, (August 2007), pp.503−512, ISSN 1556-6013

Tian, J. & Yang X. (2005).*Biometric Recognition Theory and Application*, Publishing House of Electronics Industry, ISBN 9787302184195 ,Beijing,China

Tong, V. Sibert, H. Lecoeur, J. & Girault, M. (2007). Biometric fuzzy extractors made practical: A proposal based on FingerCodes. In: *Proceedings of the ICB 2007,* Lee SW, Li SZ, pp. 604−613, Springer Berlin, ISBN 978-3-540-74548-8, Heidelberg, Germany

Tulyakov, S. Farooq, F. Govindaraju, V. (2005). Symmetric hash functions for fingerprint minutiae. In: *Pattern Recognition and Image Analysis*, Singh, S. Singh, M. Apte, C. & Perner, P, pp.30-38, Springer Berlin, ISBN 978-3-540-28833-6, Heidelberg, Germany

Tulyakov, S. Farooq, F. Mansukhani, P. & Govindaraju, V. (2007). Symmetric hash functions for secure fingerprint biometric systems. *Pattern Recognition Letters*, Vol.28, No.16, (December 2007), pp. 2427−2436, ISSN 01678655

Uludag, U. Pankanti, S. & Jain, A. (2005). Fuzzy Vault for Fingerprints. In: *Audio- and Video-Based Biometric Person Authentication*, Kanade T, Jai AK, Ratha NK, pp. 310−319, Springer Berlin, ISBN 978-3-540-27887-0, Heidelberg, Germany

Wang, XM. & Zhang, JS. (2007). Secure and Efficient Pseudorandom Bit Generator for Chaotic Stream Ciphers. *Chinese Physics Letters*, Vol.24, No.5, (feburary 2007), pp.1166−1169, ISSN 0256-307X

Wang, XM. Zhang, JS. Zhang, WF. & Khan, MK. (2006). Security Improvement on the Timestamp-based Password Authentication Scheme Using Smart Cards, *Proceedings of IEEE International Conference on Engineering of Intelligent Systems*, Islamabad, April 2006.

Wang, XM. Zhang, WF. Zhang, JS. Khan, MK. (2007).  Cryptanalysis and Improvement on Two Efficient Remote User Authentication Schemes Using Smart Cards, *Computer Standards & Interfaces*, vol.29, No.5, (July 2007), pp.507-512, ISSN 0920-5489.

Wang, XM. Zhang, WF. (2008). An efficient and secure biometric remote user authentication scheme using smart cards, *IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, Wuhan China, December 2008.

Wang, Y. & Plataniotis, K. (2008). Fuzzy vault for face based cryptographic key generation, *Proceedings of the Biometrics Symposium*, ISBN 978-1-4244-1549-6, Baltimore, January 2008

Zhang, F. Feng, D. & Sun, Z. (2008). An iris authentication scheme based on fuzzy extractor. *Journal of Computer Research and Development*, Vol.45, No.6, (December 2007), pp.1036−1042, ISSN 100021239

*Edited by Jucheng Yang and Norman Poh*

In the recent years, a number of recognition and authentication systems based on biometric measurements have been proposed. Algorithms and sensors have been developed to acquire and process many different biometric traits. Moreover, the biometric technology is being used in novel ways, with potential commercial and practical implications to our daily activities. The key objective of the book is to provide a collection of comprehensive references on some recent theoretical development as well as novel applications in biometrics. The topics covered in this book reflect well both aspects of development. They include biometric sample quality, privacy preserving and cancellable biometrics, contactless biometrics, novel and unconventional biometrics, and the technical challenges in implementing the technology in portable devices. The book consists of 15 chapters. It is divided into four sections, namely, biometric applications on mobile platforms, cancelable biometrics, biometric encryption, and other applications. The book was reviewed by editors Dr. Jucheng Yang and Dr. Norman Poh. We deeply appreciate the efforts of our guest editors: Dr. Girija Chetty, Dr. Loris Nanni, Dr. Jianjiang Feng, Dr. Dongsun Park and Dr. Sook Yoon, as well as a number of anonymous reviewers.

IntechOpen