

IntechOpen

# Biometric Systems, Design and Applications

*Edited by Zahid Riaz*





---

# **BIOMETRIC SYSTEMS, DESIGN AND APPLICATIONS**

---

Edited by **Zahid Riaz**

## Biometric Systems, Design and Applications

<http://dx.doi.org/10.5772/1014>

Edited by Zahid Riaz

### Contributors

Katsuko T. Nakahira, Miki Akahane, Yukiko Fukami, Carlos M. Manuel Travieso-Gonzalez, Marcos Del Pozo-Baños, Jesús B. Alonso, Milena Carneiro, Antônio Cláudio Paschoarelli Veiga, Edna Flores, Gilberto Carrijo, Jinfeng Yang, Yihua Shi, Renbiao Wu, Zahid Riaz, M. Saquib Sarfraz, Michael Beetz, Indra Devi, Wasil Elsadig Eltahir, Momoh Eyiomika Salami, Hasimah Ali, David Freire, Modesto Castrillón, Oscar Déniz-Suarez, Robert Brumnik, Iztok Podbregar, Teodora Ivanuša, Rahib Abiyev, Kemal Ihsan Kilic, Jesus F. Vargas, Miguel A. Ferrer, Der-Chin Chen, Ridza Azri Ramlee, Khairul Azha, Ranjit Singh Sarban Singh

### © The Editor(s) and the Author(s) 2011

The moral rights of the and the author(s) have been asserted.

All rights to the book as a whole are reserved by INTECH. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECH's written permission.

Enquiries concerning the use of the book should be directed to INTECH rights and permissions department ([permissions@intechopen.com](mailto:permissions@intechopen.com)).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

### 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.

First published in Croatia, 2011 by INTECH d.o.o.

eBook (PDF) Published by IN TECH d.o.o.

Place and year of publication of eBook (PDF): Rijeka, 2019.

IntechOpen is the global imprint of IN TECH d.o.o.

Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from [orders@intechopen.com](mailto:orders@intechopen.com)

Biometric Systems, Design and Applications

Edited by Zahid Riaz

p. cm.

ISBN 978-953-307-542-6

eBook (PDF) ISBN 978-953-51-5568-3

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

4,100+

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



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](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)





# Meet the editor



Zahid Riaz is currently working as a full time researcher in Intelligent Autonomous Systems (IAS) at Technical University of Munich (TUM), Germany. He has received his B.Sc. and M.Sc. degree in Physics and Systems Engineering in 2001 and 2004 respectively from Pakistan. He has completed his PhD from Technical University of Munich (TUM), Germany in 2011 in the area of computer vision and image processing. He was awarded German Academic Exchange Service (DAAD) and Higher Education Commission of Pakistan (HEC) fellowship in 2007 and European Research Consortium for Informatics and Mathematics (ERCIM) fellowship in 2011. Zahid Riaz is also serving as project leader for Context Aware Perception for Assistive Environments under DAAD program for German-Pakistani universities collaboration. In 2009-10, he has worked as a visiting researcher in computer vision lab at University of Central Florida (UCF).





---

# Contents

---

**Preface** XI

**Part 1 Fingerprints Verification and Identification** 1

Chapter 1 **Reliability of Fingerprint Biometry (Weibull Approach)** 3  
Robert Brumnik, Iztok Podbregar and Teodora Ivanuša

Chapter 2 **Finger-Vein Recognition Based on Gabor Features** 17  
Jinfeng Yang, Yihua Shi and Renbiao Wu

Chapter 3 **Efficient Fingerprint Recognition Through Improvement of Feature Level Clustering, Indexing and Matching Using Discrete Cosine Transform** 33  
D. Indradevi

**Part 2 Face Recognition** 55

Chapter 4 **Facial Identification Based on Transform Domains for Images and Videos** 57  
Carlos M. Travieso-González,  
Marcos del Pozo-Baños and Jesús B. Alonso

Chapter 5 **Towards Unconstrained Face Recognition Using 3D Face Model** 77  
Zahid Riaz, M. Saquib Sarfraz and Michael Beetz

Chapter 6 **Digital Signature: A Novel Adaptive Image Segmentation Approach** 93  
David Freire-Obregón, Modesto Castrillón-Santana and Oscar Déniz-Suárez

**Part 3 Iris Segmentation and Identification** 109

- Chapter 7 **Solutions for Iris Segmentation 111**  
Milena Bueno Pereira Carneiro, Antônio Cláudio P. Veiga,  
Edna Lúcia Flôres and Gilberto A. Carrijo
- Chapter 8 **Detecting Cholesterol Presence  
with Iris Recognition Algorithm 129**  
Ridza Azri Ramlee, Khairul Azha and Ranjit Singh Sarban Singh
- Chapter 9 **Robust Feature Extraction and Iris  
Recognition for Biometric Personal Identification 149**  
Rahib Hidayat Abiyev and Kemal Ihsan Kilic
- Chapter 10 **Iris Recognition System Using Support Vector Machines 169**  
Hasimah Ali and Momoh J. E. Salami
- Part 4 Other Biometrics 183**
- Chapter 11 **Verification of the Effectiveness of  
Blended Learning in Teaching Performance  
Skills for Simultaneous Singing and Piano Playing 185**  
Katsuko T. Nakahira, Yukiko Fukami and Miki Akahane
- Chapter 12 **Portable Biometric System of  
High Sensitivity Absorption Detection 195**  
Der Chin Chen
- Chapter 13 **Texture Analysis for Off-Line Signature Verification 219**  
Jesus F. Vargas and Miguel E. Ferrer
- Chapter 14 **Design and Evaluation of a Pressure  
Based Typing Biometric Authentication System 235**  
Momoh J. E. Salami, Wasil Eltahir and Hashimah Ali

---

## Preface

---

Biometric authentication has been widely used for access control and security systems over the past few years. It is the study of the physiological (biometric) and behavioral (soft-biometric) traits of humans which are required to classify them. A general biometric system consists of different modules including single or multi-sensor data acquisition, enrollment, feature extraction and classification. A person can be identified on the basis of different physiological traits like fingerprints, live scans, faces, iris, hand geometry, gait, ear pattern and thermal signature etc. Behavioral or soft-biometric attributes could be helpful in classifying different persons however they have less discrimination power as compared to biometric attributes. For instance, facial expression recognition, height, gender etc. The choice of a biometric feature can be made on the basis of different factors like reliability, universality, uniqueness, non-intrusiveness and its discrimination power depending upon its application. Besides conventional applications of the biometrics in security systems, access and documentation control, different emerging applications of these systems have been discussed in this book. These applications include Human Robot Interaction (HRI), behavior in online learning and medical applications like finding cholesterol level in iris pattern. The purpose of this book is to provide the readers with life cycle of different biometric authentication systems from their design and development to qualification and final application. The major systems discussed in this book include fingerprint identification, face recognition, iris segmentation and classification, signature verification and other miscellaneous systems which describe management policies of biometrics, reliability measures, pressure based typing and signature verification, bio-chemical systems and behavioral characteristics.

Over the past few years, a major part of the revenue collected from the biometric industry is obtained from fingerprint identification systems and Automatic Fingerprint Identification Systems (AFIS) due to their reliability, collectability and application in document classification (e.g. biometric passports and identity cards). Section I provides details about the development of fingerprint identification and verification system and a new approach called finger-vein recognition which studies the vein patterns in the fingers. Finger-vein identification system has immunity to counterfeit, active liveliness, user friendliness and permanence over the conventional fingerprints identification systems. Fingerprints are easy to spoof however current approaches like liveliness detection and finger-vein pattern identification can easily

cope with such challenges. Moreover, reliability measure of fingerprint systems using Weibull approach is described in detail.

Human faces are preferred over the other biometric systems due to their non-intrusive nature and applications at different public places for biometric and soft-biometric classification. Section II of the book describes detailed study on the segmentation, recognition and modeling of the human faces. A stand-alone system for 3D human face modeling from a single image has been developed in detail. This system is applied to HRI applications. The model parameters from a single face image contain identity, facial expressions, gender, age and ethnical information of the person and therefore can be applied to different public places for interactive applications. Moreover face identification in images and videos is studied using transform domains which include subspace learning methods like PCA, ICA and LDA and transforms like wavelet and cosine transforms. The features extracted from these methods are comparatively studied by using different standard classifiers. A novel approach towards face segmentation in cluttered backgrounds has also been described which provides an image descriptor based on self-similarities which captures the general structure of an image.

Current iris patterns recognition systems are reliable but collectability is the major challenge for them. A thorough study along with design and development of iris recognition systems has been provided in section III of this book. Image segmentation, normalization, feature extraction and classification stages are studied in detail. Besides conventional iris recognition systems, this section provides medical application to find presence of cholesterol level in iris pattern.

Finally, the last section of the book provides different biometric and soft-biometric systems. This provides management policies of the biometric systems, signature verification, pressure based system which uses signature and keyboard typing, behavior analysis of simultaneous singing and piano playing application for students of different categories and design of a portable biometric system that can measure the amount of absorption of the visible collimated beam that passes by the sample to know the absorbance of the sample.

In summary, this book provides the students and the researchers with different approaches to develop biometric authentication systems and at the same time provides state-of-the-art approaches in their design and development. The approaches have been thoroughly tested on standard databases and in real world applications.

**Zahid Riaz**  
Research Fellow  
Faculty of Informatics  
Technical University of Munich  
Garching, Germany

# **Part 1**

## **Fingerprints Verification and Identification**



# Reliability of Fingerprint Biometry (Weibull Approach)

Robert Brumnik<sup>1</sup>, Iztok Podbregar<sup>2</sup> and Teodora Ivanuša<sup>2</sup>

<sup>1</sup>*Metra inženiring Ltd.*

<sup>2</sup>*University of Maribor,  
Faculty of Criminal Justice and Security  
Slovenia*

## 1. Introduction

Biometrics refers to the identification of a person on the basis of their physical and behavioural characteristics. Today we know a lot of biometric systems which are based on the identification of these, for everyone's unique identity. Some biometric systems include the characteristics of: fingerprints, hand geometry, voice, iris, etc., and can be used for identification. Most biometric systems are based on the collection and comparison of biometric characteristics which can provide identification. This study begins with a historical review of biometric and radio frequency identification (RFID) methods and research areas. The study continues in the direction of biometric methods based on fingerprints. The survey parameters of reliability, which may affect the results of the biometric system in use, prove the hypothesis. A summary of the results obtained the measured parameters of reliability and the efficiency of the biometric system we discussed.

Each biometric system includes the following three processes: registration, preparation of a sample, and readings of the sample. Finally the system provides a comparison of the measured sample with digitized samples stored in the database. Also in this chapter we show the optimization of a biometric system with neural networks resulting in multi-biometric or multimodal biometric systems. This procedure combines two or more biometric methods in the form of a more efficient and more secure biometric system.

During our research we carried out a »Weibull« mathematical model for determining the effectiveness of the fingerprint identification system. By means of ongoing research and development projects in this area, this study is aimed at confirming its effectiveness empirically. Efficiency and reliability are important factors in the reading and operation of biometric systems. The research focuses on the measurement of activity in the process of the fingerprint biometric system, and explains what is meant by the result achieved.

The research we refer to reviews relevant standards, which are necessary to determine the policy of biometric measures and security mechanisms, and to successfully implement a quality identification system.

The hypothesis, we have assumed in the thesis to the survey has been fully confirmed. Biometric methods based on research parameters are both more reliable and effective than RFID identification systems while enabling a greater flow of people.

## 2. Theoretical overview

Personal identification is a means of associating a particular individual with an identity. The term “biometrics” derives from Bio,(meaning “life” and metric being a “measurement”. Variations of biometrics have long been in use in past history. Cave paintings were one of the earliest samples of a biometric form. A signature could presumably be deciphered from the outline of a human hand in some of the paintings. In ancient China, thumb prints were found on clay seals. In the 14th century in China, biometrics was used to identify children to merchants (Daniel, 2006). The merchants would take ink and make an impression of the child’s hand and footprint in order to distinguish between them. French police developed the first anthropometric system in 1883 to identify criminals by measuring the head and body widths and lengths. Fingerprints were used for business transactions in ancient Babylon, on clay tablets (Barnes, 2011).

Throughout history many other forms of biometrics, which include the fingerprint technique, were utilized to identify criminals and these are still in use today. The fingerprint method has been successfully used for many years in law enforcement and is now a very accurate and reliable method to determine an individual’s identity in many security access systems.

The production logistics must ensure an effective flow of material, tools and services during the whole production process and between companies. Solutions for the traceability of products and people (identification and authentication) are very important parts of the production process. The entire production efficacy and final product quality depends on the organization and efficiency of the logistics process. The capability of a company to develop, exploit and retain its competitive position is the key to increasing company value (Polajnar, 2005). Globalization dictates to industrial management the need for an effective and lean manufacturing process, downsizing and outsourcing where appropriate. The requirements of modern times are the development and use of wireless technologies such as the mobile phone. The intent is to develop remote maintenance, remote servicing and remote diagnostics (Polajnar, 2003). With the increasing use of new identification technologies, it is necessary to explore their reliability and efficacy in the logistics process. With the evolution of microelectronics, new identification systems have been achieving rapid development during the last ten years thus enabling practical application in the branch of automation of logistics and production. It is necessary to research and justify every economic investment in these applications.

Biometrics is not really a new technology. With the evolution of computer science the consecutive manner in which we can now use these unique features with the aid of computers contemporaneousness. In the future, modern computers will aid biometric technology playing a critical role in our society to assist questions related to the identity of individuals in a global world.

“Who is this person?”, “Is this the person he/she claims to be?”, “Should this individual be given access to our system or building?”, etc. These are examples of the every day questions asked by many organizations in the fields of telecommunication, financial services, health care, electronic commerce, governments and others all over the world.

The requirements and needs of quantity data and information processing are growing by the day. Also, people’s global mobility is becoming an everyday matter as is the necessity to ensure modern and discreet identification systems from different real and virtual access points on a global basis.



### 3. Quality parameters of biometrics technologies (ER, FRR, FAR, SL, EC)

In order to adopt biometric technologies such as fingerprint, iris, face, hand geometry and voice etc., we will evaluate some factors including the ease of use, error rate and cost. When we evaluate the score for each of the biometric technologies, we find that there is a range between the upper and lower scores for each item evaluated. Therefore we have to recognize that there is no perfect biometric technology.

For example, if a biometric system uses fingerprint technology, we will determine several factors as follows:

- a. What is the error rate (ER), as we use the False Acceptance Rate (FAR) or False Rejection Rate (FRR) that the system will allow?
- b. False Acceptance Rate (FAR) is the probability that a biometrics verification device will fail to reject an impostor.
- c. False Rejection Rate (FRR) is the probability that a biometrics verification device will fail to recognize the identity, or verify the claimed identity, of an enrollee.
- d. What is the security level (SL) to protect privacy and fraud that the system will require?
- e. Which environmental conditions (EC) for sensing fingerprints will be considered as dry or wet and dusty on the glass of a fingerprint scanner?

In the last ten years, new identification systems have been achieving extremely rapid development. The evolution of microelectronics has enabled practical application in the branch of automation of logistics and production. It is necessary to research and justify every economic investment in these applications. In this work the most important quantitative characteristics of reliability are explained. The authors also show the methodology for defining the reliability and efficacy of biometric identification systems in the process of identification and provide experimental research of personal identification systems<sup>1</sup> based upon reliability and efficacy parameters. Furthermore, a real identification system was upgraded based on automation and informatization.

In this article based on Biometric Identification Systems, we:

- show the availability and efficacy of analyses in the identification processes,
- extend reliability estimations of biometric identification systems based on significant reliability characteristics,
- provide a contribution to science by researching the biometric automated identification process to ensure optimal procedures.

A review of scientific databases shows that the area of assessing the reliability of identification systems in the process of production and logistics is not well explored. In modern production and logistics processes (automobile industry, aerospace industry, pharmacy, forensics, etc.) it is necessary to have fast and reliable control over the flow of people.

### 4. Defining the problem and research parameters

The availability of a production-logistic process is the probability that the system is functioning well at a given moment or is capable of functioning when used during certain

---

<sup>1</sup> Personal Identification Systems; Recent events have heightened interest in implementing more secure personal identification (ID) systems to improve confidence in verifying the identity of individuals seeking access to physical or virtual locations in the logistic process. A secure personal ID system must be designed to address government and business policy issues and individual privacy concerns. The ID system must be secure, provide fast and effective verification of an individual's identity, and protect the privacy of the individual's identity information.

circumstances. Reliability, by definition, is probability (capability) of the system to perform under the stated conditions defined by function and time (Hudoklin & Rozman, 2004). It is one of the most important characteristics of efficacy of identification systems and has an impact on safety and efficiency of the system. Military standard MIL HDBK 217 is also used to estimate the inherent reliability of electronic equipment and systems based on component failure data. It consists of two basic prediction methods: Parts-Count Analysis and Part-Stress Prediction. Increasing the system's reliability means less improper use, greater safety, fewer repair procedures and shorter identification times, consequently causing higher system availability. Implementing higher reliability in early development phases and its assurance during the use of the identification system, requires the knowledge of methods and techniques of reliability theory and their interactions.

Many different characteristics are used to measure the reliability of identification systems and their components. Some of them are connected to time functions others represent average time functions. Which of these characteristic are relevant in specified cases depends on the set goals, selected method of analysis, and the availability of data.

Characteristics of reliability are based on mean time intervals to the occurrence of failure. Time to failure is a random magnitude and we will mark it with the symbol " $t_i$ ". In this article we give definitions and statistical estimations of basic reliability characteristics. Reliability characteristics used in this research are:

- MTTF - mean time to failure
- MTBF - mean time between failures
- MTTR - mean time to repair
- $F(t)$  - unreliability function
- $\lambda(t)$ - failure rate
- $\beta$  - shape parameter
  - a.  $\beta < 1$  temporary failure frequency  $\lambda(t)$  decreases (early period, system implementation)
  - b.  $\beta = 1$  temporary failure frequency  $\lambda(t)$  is constant (normal system operation)
  - c.  $\beta > 1$  temporary failure frequency  $\lambda(t)$  increases (exploitation, ageing)

The shape parameter ( $\beta$ ) changes the configuration of the temporal distribution of operational failures.

## 5. Quantitative reliability characteristics

The theory of reliability was obtained by the authors Hudoklin and Rozman (2004):

Unreliability function  $F(t)$  is defined by the equation:

$$F(t) = P(X \leq t) \quad (1)$$

$F(t)$  is therefore the probability of a system to become non-functional in the interval between 0 and  $t$ .

If we observe a number of systems, or system components, we can calculate the statistical estimation for the unreliability function by the equation:

$$\hat{F}(t) = \frac{N_0 - N(t)}{N_0} \quad (2)$$

$N(t)$  - number of working/functional samples in the interval (0,t)

$N_0$  - number of samples at the start of observation at  $t=0$

Reliability function  $R(t)$  is complementary to the unreliability function. We can define it using the equation:

$$R(t) = 1 - F(t) = P(X > t) \quad (3)$$

$R(t)$  is the probability that a system or component will become non-functional after a time period  $t$ . We can define a statistical estimation of the reliability function using the equation:

$$\hat{R}(t) = \frac{N(t)}{N_0} \quad (4)$$

The product of the time to failure function and  $dt$  is the probability of the system or its component to become non-functional in the interval  $(t, t+\Delta t)$ . We can calculate the function  $F(t)$  by differentiation of the unreliability function by time:

$$F(t) = \frac{dF(t)}{dt} \quad (5)$$

The statistical estimation for  $f(t)$  can be calculated with the equation:

$$\hat{f}(t) = \frac{N(t) - N(t + \Delta t)}{N_0 \cdot \Delta t} \quad (6)$$

Where  $\Delta t$  is interval  $(t, t+\Delta t)$ .

Product of Failure rate  $\lambda(t)$  and  $dt$  is the conditional probability of a system/part of a system to become non-functional in the interval  $(t, t+\Delta t)$ . Momentary frequency of failure rate can be written as:

$$\lambda(t) = \frac{f(t)}{R(t)} \quad (7)$$

The statistical estimation for  $\lambda(t)$  is defined with the equation:

$$\hat{\lambda}(t) = \frac{N(t) - N(t + \Delta t)}{N(t) \cdot \Delta t} \quad (8)$$

The mean time to failure (MTTF) of the system reliability is a characteristic and not a function of time, but the average value of the probability density function for the times to failure:

$$MTTF = \int_0^{\infty} tf(t)dt = \int_0^{\infty} R(t)dt \quad (9)$$

An estimate point for the mean time to failure (MTTF) is calculated for  $n$  times to failure with the estimator:

$$MT\hat{T}F = \frac{1}{n} \sum_{i=1}^n t_i \quad (10)$$

During normal operation, the MTTF is equal to:

$$MTTF = \frac{1}{\lambda} \quad (11)$$

For many systems, or system parts, the function  $\lambda(t)$  has a characteristic “bathtub” configuration (Figure 1.). The life cycle of systems can be divided into three periods: an early damaging period, a normal working period and an ageing or exploitation period. In the first period  $\lambda(t)$  decreases, in the second period  $\lambda(t)$  is constant, and in the third period  $\lambda(t)$  rises.

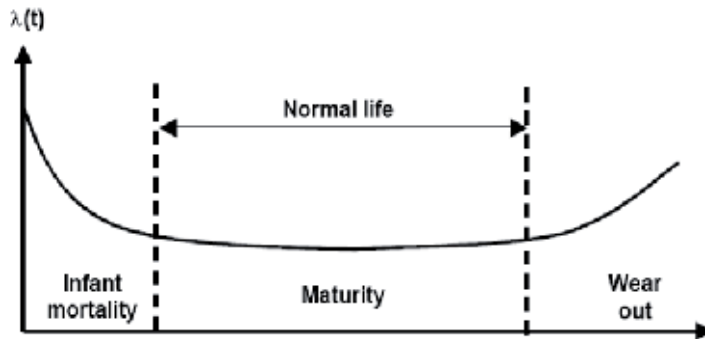


Fig. 1. “Bathtub” curve (FIDES, 2006).

### 5.1 Reliability of biometric identification systems

Definitions used in reliability calculation of biometric identification systems and terminology:

- FAR<sup>2</sup> is defined as the percentage of identification instances in which false acceptance occurs,
- FRR<sup>3</sup> is defined as the percentage of identification instances in which false rejection occurs,
- Mean time to failure (MTTF), mean time between failures (MTBF) and mean time to repair (MTTR),
- classification of failures,
- failures data bases.

In biometric methods, in contrast to the classic methods of identification, probability needs to be considered. All sensors are subject to noise and errors. The largest problem is the development and implementation of a safe crypto-algorithm. All limitations are summarized in the two terms: FRR and FAR. If a system is highly sensitive, the FAR value is low, but FRR is higher. In a system of low sensitivity the situation is reversed. Such a system is accepted by almost everyone (FAR>FRR). It is therefore necessary to make a compromise in the sensitivity of a system. It can also be regulated so that the FAR and FRR values are equal, the so-called EER (Equal Error Rate). Lower EER means a more accurate system. In an application where the speed of identification is more important than safety (e.g. hotel rooms), the high FAR value

<sup>2</sup> FAR (False Acceptance Rate); This can be expressed as a probability. For example, if FAR is 0.1 percent, it means that on average, one out of every 1000 impostors attempting to breach the system will be successful.

<sup>3</sup> FRR (False Rejection Rate); For example, if FRR is 0.05 percent, it means that on average, one out of every 2000 authorized persons attempting to access the system will not be recognized by that system.

can be allowed (Hicklin et al., 2005). Graphic presentation of both errors depending on the size of the error threshold of biometric system can be seen in Figure 2.

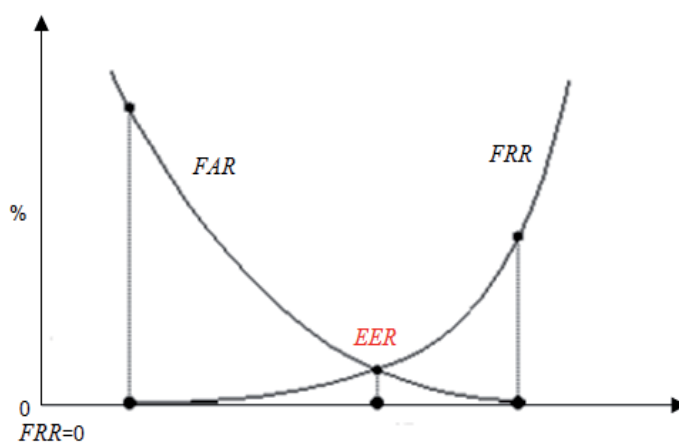


Fig. 2. Calculating EER from FAR – FRR intersection.

## 5.2 Usability and reliability characteristics of a biometric system reader

To fully understand user-centered design, it is essential to understand the features inherent in a usable system. Usability helps to ensure that systems and products are easy to learn, effective to use and enjoyable from the user's perspective. This is defined as: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." (ISO 13407:1999). Additional attributes of usability that may also be considered include:

- effective to use (effectiveness),
- efficient to use (efficiency),
- enjoyable to use (satisfaction),
- easy to learn (learnability) and
- easy to remember (memorability).

The table on the next page lists each of these usability goals and provides a short description of each, along with a few questions for biometric system designers to consider. Usability testing not only provides insights into users' behaviour, but it also allows project teams to quantitatively measure the success of a system, including capturing metrics such as error rates, successful performance on tasks, time to complete a task, etc. (NIST, 2008). For quantitative testing, many teams use the Common Industry Format (CIF) (ISO/IEC 25062:2006) to document the performance of the system. The CIF provides a standard way for organizations to present and report quantitative data gathered in a usability test, so that it can later be compared to the results gathered in subsequent tests.

A review of the literature and standards for design and anthropometric measurements provided guidance on proper angles for fingers or palm placement. Standards focus on line of sight and reach envelopes including sloping control panels for cockpits or nuclear power stations (NISTIR 7504).

To determine the reliability characteristics, we used the Weibull model, which is useful in cases where  $\lambda(t)$  cannot be illustrated by the constant function. For the resulting measurements we will take advantage of Weibull analysis, which provides a simple

graphical method. The analysis will be provided (with a reasonable error analysis) to obtain good estimates of parameters, despite the small sample size (in our case, thirty pieces of biometric modules). These solutions enable us to identify early signs of potential problems, so we can prevent more serious systemic failures and predict the maintenance cycle (increasing the availability of the system). The study was of a relatively small sample size also enabling cost-effective test curves. Testing is complete when the observed system fails (sudden failure) in each of the three groups (the first module of each series) biometric reader components and proceeds with the Weibull analysis.

Reliability of a biometric system depends on three factors (Chernomordik, 2002):

- uniqueness and repeatability, which means that the characteristic used should provide for different readings for different people, and the readings obtained for the same person at different times and under different conditions should be similar,
- reliability of the matching algorithm and
- quality of the reading device.

Failures, which we have taken into account in determining the characteristics of MTTF and MTTR of a biometric system (Table 1):

- failure of the software (the inability to read the sample),
- failure of hardware (biometric reader, PCBs) and
- errors due to sensor reading settings: FAR, FRR.

| Ser. No. | Time to first failure (days) | Time to second failure (days) | Time to third failure (days) | Average value (days) |
|----------|------------------------------|-------------------------------|------------------------------|----------------------|
| 36365    | 53                           | 89                            | 88                           | 88,8                 |
| 36359    | 60                           | 106                           | 73                           |                      |
| 36366    | 87                           | 106                           | 88                           |                      |
| 36364    | 86                           | 161                           | 88                           |                      |
| 36368    | 102                          | 130                           | 13                           |                      |
| 36369    | 99                           | 102                           | 98                           |                      |
| 36360    | 56                           | 150                           | 93                           |                      |
| 36345    | 57                           | 90                            | 126                          |                      |
| 36381    | 81                           | 52                            | 117                          |                      |
| 36384    | 90                           | 65                            | 105                          |                      |
| Ser. No. | Time to first repair (days)  | Time to second repair (days)  | Time to third repair (days)  | Average value (days) |
| 36365    | 1                            | 1                             | 1                            | 1,2                  |
| 36359    | 0                            | 1                             | 0                            |                      |
| 36366    | 1                            | 3                             | 2                            |                      |
| 36364    | 1                            | 0                             | 1                            |                      |
| 36368    | 1                            | 3                             | 1                            |                      |
| 36369    | 2                            | 1                             | 1                            |                      |
| 36360    | 2                            | 1                             | 1                            |                      |
| 36345    | 3                            | 1                             | 1                            |                      |
| 36381    | 2                            | 1                             | 1                            |                      |
| 36384    | 2                            | 0                             | 1                            |                      |

| Ser. No. | Time to restart (days) | Time to restart (days) | Time to restart (days) | Average value (days) |
|----------|------------------------|------------------------|------------------------|----------------------|
| 36365    | 56                     | 90                     | 89                     | 90,1                 |
| 36359    | 60                     | 105                    | 73                     |                      |
| 36366    | 88                     | 107                    | 90                     |                      |
| 36364    | 85                     | 161                    | 89                     |                      |
| 36368    | 103                    | 133                    | 16                     |                      |
| 36369    | 101                    | 103                    | 99                     |                      |
| 36360    | 58                     | 151                    | 96                     |                      |
| 36345    | 60                     | 91                     | 127                    |                      |
| 36381    | 83                     | 53                     | 118                    |                      |
| 36384    | 92                     | 65                     | 106                    |                      |

Table 1. Data for the MTTF, MTTR estimates determine for biometric system.

| TIME TO FIRST FAILURE (days)  |    |    |    |    |     |     |     |     |     |     |
|-------------------------------|----|----|----|----|-----|-----|-----|-----|-----|-----|
| $i$                           | 1  | 2  | 3  | 4  | 5   | 6   | 7   | 8   | 9   | 10  |
| $t_i$ (days)                  | 53 | 56 | 57 | 60 | 81  | 86  | 87  | 90  | 99  | 102 |
| $F_i$                         | 7  | 16 | 26 | 36 | 45  | 55  | 64  | 74  | 84  | 93  |
| TIME TO SECOND FAILURE (days) |    |    |    |    |     |     |     |     |     |     |
| $i$                           | 1  | 2  | 3  | 4  | 5   | 6   | 7   | 8   | 9   | 10  |
| $t_i$ (days)                  | 52 | 65 | 89 | 90 | 102 | 106 | 106 | 130 | 150 | 161 |
| $F_i$                         | 7  | 16 | 26 | 36 | 45  | 55  | 64  | 74  | 84  | 93  |
| TIME TO THIRD FAILURE (days)  |    |    |    |    |     |     |     |     |     |     |
| $i$                           | 1  | 2  | 3  | 4  | 5   | 6   | 7   | 8   | 9   | 10  |
| $t_i$ (days)                  | 13 | 73 | 88 | 88 | 88  | 93  | 98  | 105 | 117 | 126 |
| $F_i$                         | 7  | 16 | 26 | 36 | 45  | 55  | 64  | 74  | 84  | 93  |

Table 2. The times to failure and the associated estimates points for  $F(t)$  of the biometric system.

| TIME TO FIRST ACTIVE REPAIR (days)  |   |   |   |   |   |   |   |   |   |    |
|-------------------------------------|---|---|---|---|---|---|---|---|---|----|
| $i$                                 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $t_i$                               | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 3 | 2 | 2  |
| TIME TO SECOND ACTIVE REPAIR (days) |   |   |   |   |   |   |   |   |   |    |
| $i$                                 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $t_i$                               | 1 | 1 | 3 | 0 | 3 | 1 | 1 | 1 | 1 | 0  |
| TIME TO THIRD ACTIVE REPAIR (days)  |   |   |   |   |   |   |   |   |   |    |
| $i$                                 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $t_i$                               | 1 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1  |

Table 3. The times of active repairs for the biometric module.

Assuming that the times to failure in Tables 2 are exponentially distributed couples  $(t_i, F_i)$ . We join them together and rank them in Table 4 and estimate parameters  $\beta$  and  $\eta$  for a biometric system with the software Weibull++7.

Time to first failure is  $\beta = 4.6$  and  $\eta = 72$ , while they are behind the times to failure of another parameter  $\beta = 3.23$  and  $\eta = 117.2$ . For the third time to failure, the values of parameters  $\beta = 2$  and  $\eta = 101$ . Table 4 shows the ranking values of times to failure of biometric systems ( $t_i$ ;  $i=1,2,3$ ) and times to failure of the biometric identification system and the corresponding estimation point estimates of  $F(t)$ .

| biometric module |              |       |
|------------------|--------------|-------|
| $i$              | $t_i$ (days) | $F_i$ |
| 1                | 13           | 2,3   |
| 2                | 52           | 5,6   |
| 3                | 53           | 8,9   |
| 4                | 56           | 12,2  |
| 5                | 57           | 15,5  |
| 6                | 60           | 18,8  |
| 7                | 65           | 22,0  |
| 8                | 73           | 25,3  |
| 9                | 81           | 28,6  |
| 10               | 86           | 31,9  |
| 11               | 87           | 35,2  |
| 12               | 88           | 38,5  |
| 13               | 88           | 41,8  |
| 14               | 88           | 45,1  |
| 15               | 89           | 48,4  |
| 16               | 90           | 51,6  |
| 17               | 90           | 54,9  |
| 18               | 93           | 58,2  |
| 19               | 98           | 61,5  |
| 20               | 99           | 64,8  |
| 21               | 102          | 68,1  |
| 22               | 102          | 71,4  |
| 23               | 105          | 74,7  |
| 24               | 106          | 78,0  |
| 25               | 106          | 81,3  |
| 26               | 117          | 84,5  |
| 27               | 126          | 87,8  |
| 28               | 130          | 91,1  |
| 29               | 150          | 94,4  |
| 30               | 161          | 97,7  |

Table 4. Ranking times to failure for the biometric system and estimation point estimates of  $F(t)$ .



For the biometrics module we provide an estimated point of the average time of repairs:

$$M\hat{T}TR = \frac{1}{r} \sum_{i=1}^r t_i = \frac{1}{30} (15 + 12 + 10) = 1,2 \text{ days}$$

Estimates point for the availability of a biometric module for the period of observation is:

$$\hat{A} = \frac{MTTF}{MTTF + MTTR} = \frac{88,8}{88,8 + 1,2} = 0,987$$

With the Weibull++7 analysing tool we modeled probability density for time to failure of a biometric system with a distribution law and with the Weibull parameters  $\beta$  and  $\eta$  (Figure 3). At 100% probability, a failure of a biometric system, appears at  $\beta = 2.8$  and  $\eta = 101.7$ .

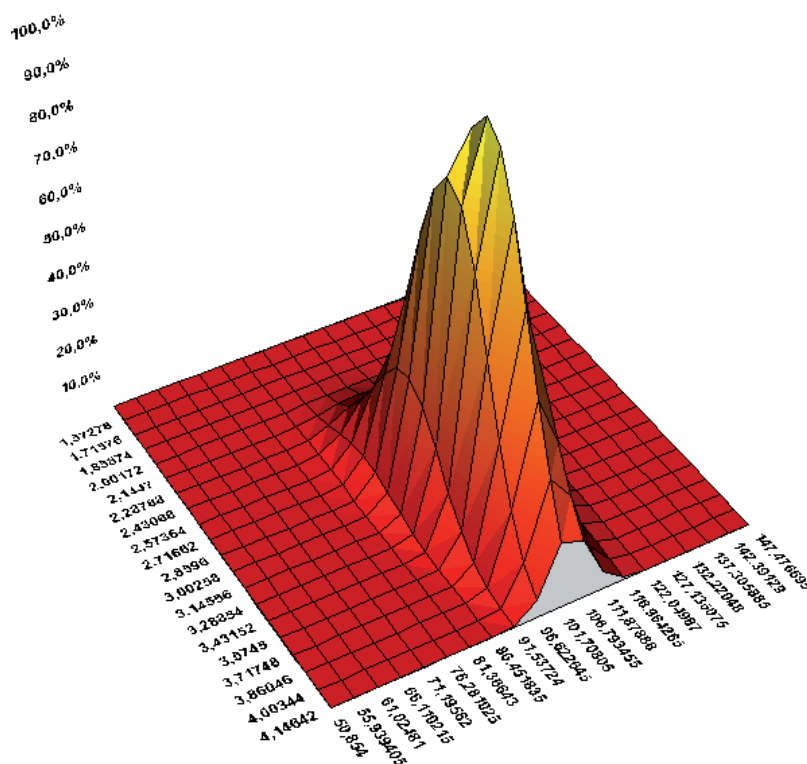


Fig. 3. Weibull model of failure appear for biometric system.

### 5.3 MTTF model calculation of system with two equivalent parts in parallel configuration

In practice, a request is made for the smooth functioning of the identification system, despite the likelihood of failure of a biometric card reader (airports, local government units, police stations, etc.). To increase the reliability of the biometric system and ensure the continuous operation despite the failure, we can associate two equivalent unit biometric module dynamic readings in the event of termination of the first reader to function by another

reader. We will show the probability graph for the biometric reader unit, which will be tied in parallel to achieve better reliability parameters of the identification system. Consider a system consisting of two equivalent units. From the failure rate  $\lambda$  of each dynamic reading module, the frequency of repairs and  $\mu$  conclusions we can construct a corresponding probability graph for reliability (Figure 4) and availability (Figure 5) in the passive parallel configuration with an absolutely reliable switch.

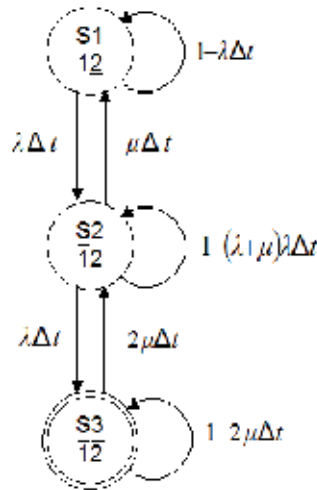


Fig. 4. Probability graph for the availability of two parallel biometric components.

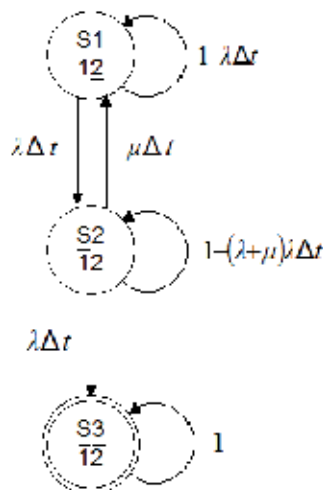


Fig. 5. Probability graph for the reliability of two parallel biometric components.

The probability graph for the availability of two parallel biometric components is shown in Figure 4. S3 state no longer abyss, the probability of transition from state S3 to state S2 is  $2\mu\Delta t$ . The probability graph for the reliability of two parallel biometric components is shown in Figure 5.

## 6. Summary and future work

The reliability and availability of assessing identification systems is an area that is very important and essential in choosing an access control system. In this article we have used statistical methods for assessing the effectiveness of biometric access by assessing the reliability and availability of all parts of the identification system with the Weibull model. The Weibull function of two variables well describes the characteristics of reliability of biometric identification systems. Data visualization using graphs give a clear correlation between the measurements and the Weibull distribution. The greater the slope of the line, which means a higher Weibull parameter  $\beta$ , the greater the reliability of the products and also the lower the risk (with the same parameter  $\eta$ ) that the identification system will terminate in shorter time. This is due to enhancing the value of the Weibull parameter leading to longer times to failure. In assessing the statistical parameters we must be aware that this appraisal is a deviation from actual values. It is clear that the expected interval of 30 data (with 90-percent confidence) for real values of the Weibull parameter allowing for a variation of about 10 percent of the calculated values of this parameter, while calculating the second parameter, the Weibull distribution is more reliable.

By calculating estimated times to failure and between failures of identification systems according to the Weibull methodology, we arrive at the following results for the assessment of the reliability and availability:

1. Estimated time to failure (reliability), of a biometric system by calculating the characteristics  $MTTF_S = 88.8$  days.
2. Estimated time to repair of a biometric system by calculating the characteristics  $MTTR_S = 1.2$  days.
3. Assessment of the availability of the biometric system by calculating the characteristics of  $A_S = 0.987$ .

During our research we carried out different models for estimating reliability and availability, which were designed using the Weibull approach. As a novelty in the field of design reliability estimates of the identification system, we also designed and applied a graphic Weibull model, which is independent of the calculated Weibull method and serves to check the calculations of the Weibull model. In the application model in the field of biometrics, we discussed the usefulness in a real domain.

The usefulness of biometric systems is shown in identification-logic environments where personal identification is needed. From this research it is evident that the ageing period of biometric systems begins relatively quickly. The results also show that the availability of biometric identification systems is therefore lower and maintenance costs are higher. The functional and ergonomic advantages of biometry are clear because there is neither the need for cards nor any other elements of identification in the identification process. The use of biometric systems will make identification simple and at the same time increase reliability due to non-transferability of identification elements (e.g. fingerprints) and prevent improper use.

It can be expected that Slovenia will attain biometric technology despite the doubts expressed by some institutions (Office for Personal Data Protection). Many open ethical questions arise, mostly regarding human personality, privacy and control. However research such as this on reliability and availability show, unequivocally, that biometric technology has an advantage both in practical use and data safety. Not only do usability improvements lead to better, easier-to-use products, they also lead to improved user

performance and satisfaction as well as substantial cost savings. By designing a biometric system with usability in mind, development teams can enhance ease of use, reduce system complexity, improve user performance and satisfaction, and reduce support and training costs.

Personal responsibility and accuracy in fields such as legislation, regulation adjustment, and production and supply chain management in global technical operations are more easily controlled using automated identification. With the automation of identification there are also possibilities for merging and comparing current process data with that from integral information systems (ERP, MRPII, etc.) or other business applications.

## 7. References

- Barnes, J. G. (2011). History, *The Fingerprint Sourcebook*. Retrieved 09.01.2011 on: <http://www.ncjrs.gov/pdffiles1/nij/225321.pdf>
- Daniel, G. (2006). *Biometrics - The Wave of the Future?* Retrieved 29.02.2011 on: [http://www.infosecwriters.com/text\\_resources/pdf/Biometrics\\_GDaniel.pdf](http://www.infosecwriters.com/text_resources/pdf/Biometrics_GDaniel.pdf)
- Hicklin, A.; Watson, C. & Ulery, B. (2005). The Myth of Goats: How many people have fingerprints that are hard to match?, *NIST Interagency Report 7271*.
- Hudoklin, A. & Rozman, V. (2004). *Reliability and availability of systems human-machine*. Publisher: Moderna organizacija, Kranj.
- Polajnar, A. (2005). Excellence of toolmaking firms : supplier - buyer - Toolmaker, *Collection of Conference consultation*, Portorose, 11.-13. October 2005.
- Polajnar, A. (2003) Exceed limits on new way : supplier - buyer - toolmaker, *Collection of Conference consultation*, Portorose, 14.-16. October 2003.
- MIL-HDBK-217, *Reliability Prediction of Electronic Equipment*. U.S. Department of Defense. Retrieved 09.02.2011 on: <http://www.itemuk.com/milhdbk217.html>
- FIDES (2006). *Nature of the Prediction*. Retrieved 29.01.2011 on: <http://fides-reliability.org/Default.aspx?tabid=94>
- Chernomordik, (2002). *Biometrics: Fingerprint based systems*. Retrieved 24.01.2011 on: [http://biometrica.ru/root/?Itemid=49&id=126&option=com\\_content&task=view&lang=en](http://biometrica.ru/root/?Itemid=49&id=126&option=com_content&task=view&lang=en)
- ISO 13407 (1999). *Human-centred design processes for interactive systems*. Retrieved 24.01.2011 on: <http://zonecours.hec.ca/documents/A2007-1-1395534.NormeISO13407.pdf>
- ISO/IEC 25062 (2006). *Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Common Industry Format (CIF) for usability test reports*. Retrieved 22.01.2011 on: [http://webstore.iec.ch/preview/info\\_isoiec25062%7Bed1.0%7Den.pdf](http://webstore.iec.ch/preview/info_isoiec25062%7Bed1.0%7Den.pdf)
- NIST (2008). *Ensuring Successful Biometric Systems, Usability & Biometrics*. Retrieved 22.01.2011 on: [http://zing.ncsl.nist.gov/biousa/docs/Usability\\_and\\_Biometrics\\_final2.pdf](http://zing.ncsl.nist.gov/biousa/docs/Usability_and_Biometrics_final2.pdf)
- NISTIR 7504 (2008). *Usability Testing of Height and Angles of Ten-Print Fingerprint Capture*. Retrieved 18.02.2011 on: <http://zing.ncsl.nist.gov/biousa/docs/NISTIR-7504%20height%20angle.pdf>

# Finger-Vein Recognition Based on Gabor Features

Jinfeng Yang, Yihua Shi and Renbiao Wu

*Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China  
China*

## 1. Introduction

Recently, a new biometric technology based on human finger-vein patterns has attracted the attention of biometrics-based identification research community. Compared with other traditional biometric characteristics (such as face, iris, fingerprint, etc.), finger vein exhibits some excellent advantages in application. For instance, apart from uniqueness, universality, permanence and measurability, finger-vein based personal identification systems hold the following merits:

- Immunity to counterfeit: Finger veins hiding underneath the skin surface make vein pattern duplication impossible in practice.
- Active liveness: Vein information disappears with musculature losing energy, which makes artificial veins unavailable in application.
- User friendliness: Finger-vein images can be captured noninvasively without the contagion and un-pleasant sensations.

Hence, the finger-vein recognition technology is widely considered as the most promising biometric technology in future.

The current available techniques for finger-vein recognition are mainly based on vein texture feature extraction (Miura et al., 2004; 2007; Mulyono and Horng, 2008; Zhang et al., 2006; Vlachos et al., 2008; Yang et al., , 2009a;b;c; Hwan et al., 2009; Liu et al., 2010; Yang et al., 2010). Although texture features are effective for finger-vein recognition, three inherent drawbacks remain unsolved. First, the current finger-vein ROI localization methods are sensitive to finger position variation, which inevitably increases intra-class variation of finger veins. Besides, the current finger-vein image enhancement methods are ineffective to improve the quality of finger-vein images, which is very unhelpful for feature information exploration. Most importantly, the current texture-based finger-vein extraction methods are impotent to reliably describe the properties of veins in orientation and diameter variations, which can directly impair the recognition accuracy.

For finger-vein recognition, a desirable finger-vein feature extraction approach should address ROI localization, image enhancement and oriented-scaled image analysis, respectively. Therefore, in this chapter, detailed descriptions on these aspects are given step by step. First, to localize finger-vein ROIs reliably, a simple but effective ROI segmentation method is proposed based on the physiological structure of a human finger. Second, haze removal method is used to improve the visibility of finger-vein images considering light scattering phenomenon in biological tissues. Third, a bank of even-symmetric Gabor filters is designed to exploit

finger-vein information in multi-scale and multi-orientation. Finally, to improve the reliability of identification, finger-vein features are extracted in Gabor transform domain, and a fusion scheme in decision level is adopted. Experimental results show that the proposed method performs well in personal identification.

## 2. Finger-vein imaging system

In anatomy, finger veins lie beneath epidermis, and form a network spreading along a finger in a high random manner. Since they are internal, visible lights usually are incapable of imaging them. Thus, illuminating the subcutaneous region of a finger properly is an important task of vein visualization. In medical applications, the NIR (near infrared) lights (760- 850nm) are often used in vein imaging because they can penetrate relatively deep into the skin as well as the radiation of lights can be absorbed greatly by the deoxyhemoglobin (Zharov et al., 2004).

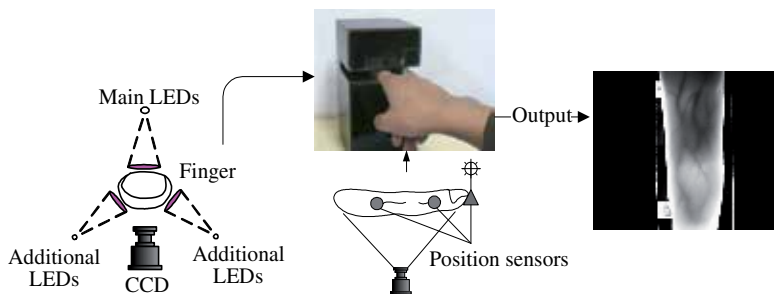


Fig. 1. The proposed principle of a homemade finger-vein imaging system.

In our application, a homemade finger-vein image acquisition system is designed and established as shown in Fig. 1. An open window with a fixed size centered in the width of CCD image plane is set for imaging. The luminaire contains main NIR light-emitting diodes (LEDs) and two additional LEDs at a wavelength of 760 nm, and a CCD sensor is placed underneath a finger. Here, the additional LEDs are only used for enhancing the contrast between veins and other tissues. Furthermore, to reduce the variations of imaging poses, two position sensors (denoted by two brighter cylinders in the right of Fig. 1) are set to light an indicator lamp when a finger is placed properly.

From the right of Fig. 1, we can see that the captured image contains not only the finger-vein region but also some uninformative parts. So, the original image needs to be preprocessed to localize a finger-vein region.

## 3. Finger-vein image preprocessing

### 3.1 Finger-vein ROI localization

It is well known that two phalangeal joints, as shown in Fig. 2(a), related with the middle phalanx of a finger make the finger activities possible. And, a functional interphalangeal joint organ is constituted by several components, as shown in Fig. 2(b). Obviously, the density of synovial fluid filling in the clearance between two cartilages is much lower than that of bones. This makes possible that more lights penetrate the clearance region when a near infrared LED array is placed over a finger. Thus, a brighter region may exist in the CCD image plane, as shown in Fig. 2(c). Actually, the clearance of a finger inter-phalangeal joint only is with 1.5-2 mm width. Hence, the brighter region can be substituted by a line with a pixel width. We

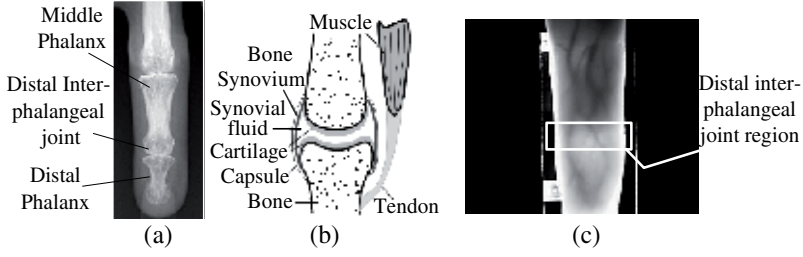


Fig. 2. Phalangeal joint prior. (a) A X-Ray finger image; (b) Phalangeal joint structure; (c) A possible region (white-rectangle) containing a phalangeal joint.

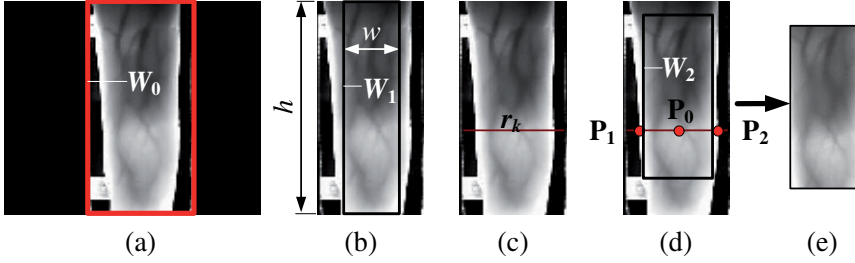


Fig. 3. Finger-vein ROI localization. (a) Finger-vein imaging window denoted by  $W_0$ ; (b) A subwindow  $W_1$  centered in the width of  $W_0$ ; (c) Inter-phalangeal joint position; (d) Finger-vein ROI region  $W_2$ ; (e) Finger-vein ROI image.

call the above observation the *interphalangeal joint prior*. This will be fully used in vein ROI localization.

According to the preceding observation, the idea resides in the use of the distal interphalangeal joint as the localized benchmark. In addition, Yang et al. found out that the only partial imagery of a human finger can deliver discriminating clues for vein recognition (Yang et al., 2009a). Likewise, we employ the similar subwindow scheme to achieve the description of vein images, since most of vein vessels actually disappear at the finger tip and boundaries. The specific procedure of vein ROI localization is as follows:

- A fixed window (denoted by  $W_0$  in Fig. 3(a)) same as finger-vein imaging window in size is used to crop a finger-vein candidate region in CCD imaging plane.
- A predefined  $w \times h$  window (denoted by  $W_1$ ) is used to locate a subregion in  $W_0$ . This can reduce the effect of uninformative background, as illustrated in Fig. 3(b);
- The pixel values at each row image are accumulated in the subregion  $W_1$ :

$$\Phi_i = \sum_{j=1}^w I(i, j), \quad i = 1, \dots, h; \quad (1)$$

- The maximum row-sum is pinpointed to approximately denote the position ( a line denoted by  $r_k$ ) of the distal interphalangeal joint, as displayed in Fig. 3(c):

$$r_k = \arg \max_{i \in [1, h]} (\Phi_i); \quad (2)$$

- Three exemplar points  $P_1$ ,  $P_2$ , and  $P_0$  are located along the detected baseline. The points  $P_1$  and  $P_2$  represent the intersection of the joint baseline and the finger borders, respectively. Meanwhile, the point  $P_0$  stands for the midpoint of the segment between  $P_1$  and  $P_2$ ;
- Based on point  $P_0$ , a window, denoted by  $W_2$  in Fig. 3(d), is used to crop a ROI image from the finger vein region as shown in Fig. 3(e). Note that the line  $r_k$  runs at  $2/3$  height of  $W_2$ .

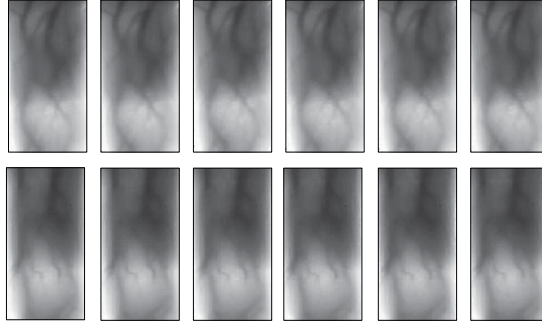


Fig. 4. Some samples ROI images from one subject at different sessions.

The fingers vary greatly in shape not only from different people but also from an identical individual, the cropped ROI by  $W_2$  therefore may be different in size. For reducing the aspect ratio variation of ROIs, all ROI images are normalized to  $180 \times 100$  pixels. Fig. 4 delineates some sample ROI images of one subject at different instants. We can note from Fig. 4 that the sample ROI images have little intra-class variation.

From Fig. 4, we can easily see that the contrast of finger-vein images usually is low and the separability is less between vascular and nonvascular regions. This brings a big challenge for finger-vein recognition, since the finger-vein patterns may be unreliable when feature extraction methods are weak in generalization.

### 3.2 Finger-vein image restoration

Researches in the medical domain reveal that the NIR lights penetrating through a human finger can be absorbed, reflected, scattered and refracted by such finger components as bones, muscles, blood vessels, and skin tissue (Delpy and Cope, 1997; Anderson and Parrish, 1981; Xu et al., 2002). This phenomenon is similar to the way of light scattering in fog (Sassaroli et al., 2004), which can greatly reduce the visibility of imaging scenes. Degraded finger-vein imageries therefore are nature products of the current available finger-vein imaging systems.

To remove the scattering effect from images, dehazing techniques currently are effective ways in many applications (Jean and Nicolas, 2009; Narasimhan and Nayar, 2003). Assume that  $I(x, y)$  is the captured image,  $R(x, y)$  is the original image free of haze,  $\rho(\lambda)$  denotes the extinction coefficient of the fog (scattering medium) and  $d(x, y)$  is the depth-map of the scene, the Koschmieder's law (Hautière et al., 2006) defined as the following often is used to restore the degraded image.

$$R(x, y) = I(x, y)e^{\rho(\lambda)d(x, y)} + I_v(1 - e^{\rho(\lambda)d(x, y)}), \quad (3)$$

where  $\lambda$  is wavelength of light and  $I_v$  is the luminance of the imaging environment. Approximatively, the Koschmieder's law can be transferred to solve the finger-vein image



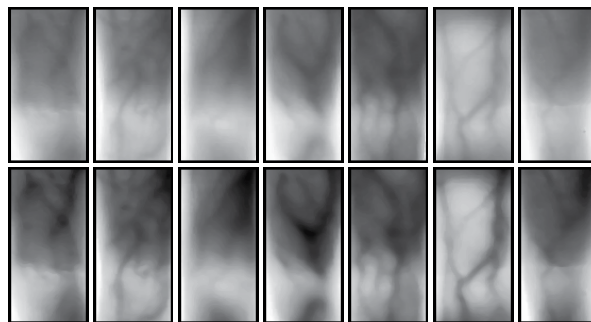


Fig. 5. Dehazing-based image restoration. Top: some original images; Bottom: restored images.

restoration problem since the fog and the biological tissues are two light scattering media with different extinction coefficients corresponding to different light wavelengths.

Inconveniently, it is difficult to obtain the exact  $\rho(\lambda)$ ,  $I_v$  and  $d(x, y)$  in practice, so a filter approach proposed in (Jean and Nicolas, 2009) is adopted here to estimate  $R(x, y)$ . This method can successfully implement visibility restoration from a single image with high speed. Fig. 5 shows some low-contrast, degraded finger-vein images and their restored versions. It can be seen from Fig. 5 that haze removal can improve image visibility apparently. However, it is also obvious that the contrast between venous region and nonvenous region is still low, and the brightness is nonuniform in nonvenous region. All these may affect the subsequent processing in feature extraction.

### 3.3 Finger-vein image enhancement

To further improve the contrast of a finger-vein image as well as compensate the nonuniform illumination in an automatic manner, a nonlinear method proposed in (Shi et al., 2007) is first used to correct pixels adaptively, then the illumination variations across the whole image are approximately estimated. From Fig. 5, we can see that venous regions are always darker than nonvenous regions in brightness due to NIR light absorption, which is not helpful for making venous region (object region) salient in practice. The negative version of a restored and corrected finger-vein image therefore is used for background illumination estimation, as shown in Fig. 6(b). Here, the average filter with a  $16 \times 16$  mask is used as a coarse estimator of the background illumination, as shown in Fig. 6(c).

Subtracting the estimated background illumination from the negative image, we can obtain an image with lighting variation compensation, as shown in Fig. 6(d). Then, we enhance the lighting corrected image by means of histogram equalization. Such processing compensates for the nonuniform illumination, as well as improves the contrast of the image. Fig. 6(e) and 6(f) show the enhanced results of some finger-vein images, from which we can clearly see that the finger-vein network characteristics become clearer than those in the top of Fig. 5. To reduce the noises generated by image operation, the median filter with a  $3 \times 3$  mask is used accordingly.

## 4. Finger-vein feature analysis

### 4.1 Even Gabor filter design

Gabor filters have been successfully employed in a wide range of image-analysis applications since they are tunable in scale and orientation (Jie et al., 2007; Ma et al., 2003; Jain et al., 2007; Laadjel et al., 2008; Lee, 1996; Yang et al., 2003; Zhu et al., 2007). Considering the variations

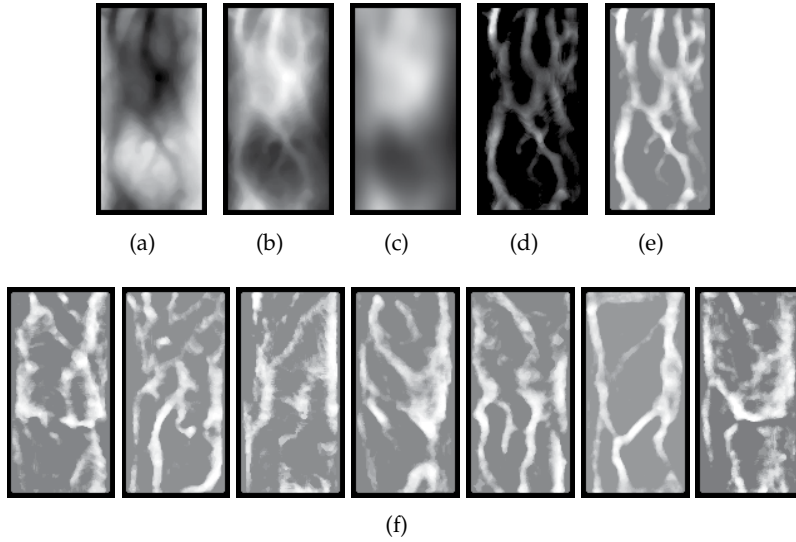


Fig. 6. Enhancement procedure. (a) Restored finger-vein image; (b) Negative version of the corrected image after restoration; (c) Estimated background illumination; (d) Image with background illumination subtraction; (e) Enhanced image; (f) Other enhanced results corresponding to the samples in Fig. 5.

of vessels in orientation and diameter along a finger, oriented Gabor filters in multiscale are therefore desirable for venous region texture analysis.

A two-dimensional Gabor filter is a function composed by a Gaussian-shaped function and a complex plane wave (Daugman, 1985), which is defined as

$$G(x, y) = \frac{\gamma}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2} \left( \frac{x_\theta^2 + \gamma^2 y_\theta^2}{\sigma^2} \right) \right\} \exp(j2\pi f_0 x_\theta), \quad (4)$$

where

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

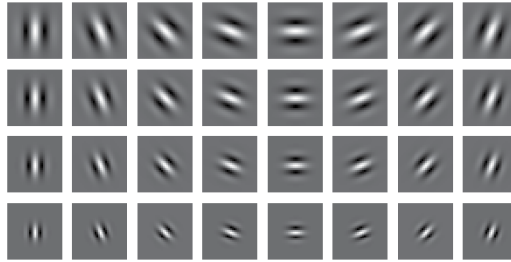
$\hat{j} = \sqrt{-1}$ ,  $\theta$  is the orientation of a Gabor filter,  $f_0$  denotes the filter center frequency,  $\sigma$  and  $\gamma$  respectively represent the standard deviation (often called scale) and aspect ratio of the elliptical Gaussian envelope,  $x_\theta$  and  $y_\theta$  are rotated versions of the coordinates  $x$  and  $y$ . Determining the values of the four parameters  $f_0$ ,  $\sigma$ ,  $\gamma$  and  $\theta$  usually play an important role in making Gabor filters suitable for some specific applications (Lee, 1996).

Using Euler formula, Gabor filter can be decomposed into a real part and an imaginary part. The real part, usually called even-symmetric Gabor filter (denoted by  $G^e(\cdot)$  in this paper), is suitable for ridge detection in an image (Yang et al., 2003), while the imaginary part, usually called odd-symmetric Gabor filter, is beneficial to edge detection (Zhu et al., 2007). Since the finger veins appear dark ridges in image plane, even-symmetric Gabor filter here is used to exploit the underlying features from the finger-vein network. To make even Gabor wavelets into admissible Gabor wavelets, the DC response should be compensated.

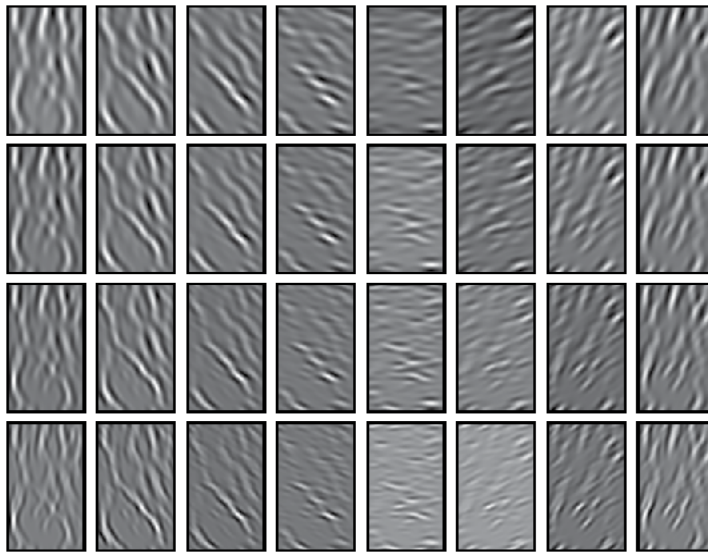
Based on Eq. 4, a bank of admissible even-symmetric Gabor filters subtracting the DC response can be expressed as (Lee, 1996)

$$G_{sk}^e(x, y) = \frac{\gamma}{2\pi\sigma_s^2} \exp \left\{ -\frac{1}{2} \left( \frac{x_{\theta_k}^2 + \gamma^2 y_{\theta_k}^2}{\sigma_s^2} \right) \right\} \times \left( \cos(2\pi f_s x_{\theta_k}) - \exp(-\frac{\nu^2}{2}) \right), \quad (5)$$

where  $s$  is the scale index,  $k$  is the orientation index and  $\nu$  is a factor determining DC response whose value is determined by  $\sqrt{2 \ln 2} (2^{\Delta\phi} + 1) / (2^{\Delta\phi} - 1)$ .



(a)



(b)

Fig. 7. Spatial filtering. (a) A bank of even-symmetric Gabor filters; (b) The 2D convolution results.

Since  $f_s$ ,  $\sigma_s$ ,  $\gamma$  and  $\theta$  usually govern the output of a Gabor filter, these parameters should be determined sensibly for finger-vein analysis application. Considering that vein vessels hold high random characteristics in diameter and orientation,  $\gamma$  is set equal to one (i.e., Gaussian function is isotropic) for reducing diameter deformation arising from elliptic Gaussian envelop,  $\theta$  varies from zero to  $\pi$  with a  $\pi/8$  interval (that is, the even-symmetric Gabor filters are embodied in eight channels). To determine the relation of  $\sigma_s$  and  $f_s$ , a scheme proposed in

(Daugman, 1985; Lee, 1996) is used here, which is defined as follow

$$\sigma_s f_s = \frac{1}{\pi} \sqrt{\frac{\ln 2}{2}} \cdot \frac{2^{\Delta\phi} + 1}{2^{\Delta\phi} - 1}, \quad (6)$$

where  $\Delta\phi (\in [0.5, 2.5])$  denotes the spatial frequency bandwidth (in octaves) of a Gabor filter. Let  $s = 1, \dots, 4$ ,  $\sigma_s = 8, 6, 4, 2$  and  $k = 1, \dots, 8$ , we can build a bank of even-symmetric Gabor filters with four scales and eight orientations, as shown in Fig. 7(a). Assume that  $F(x, y)$  denote a filtered  $R(x, y)$ , we can obtain

$$F_{sk}(x, y) = G_{sk}^e(x, y) * R(x, y), \quad (7)$$

where  $*$  denotes 2D image convolution operation. Thus, for a enhanced finger-vein image, 32 filtered images are generated by a bank of Gabor filters, as shown in Fig. 7(b). Noticeably, Gabor filters corresponding to the top row and the bottom row in Fig. 7(a), respectively, are undesirable for finger-vein information exploitation since they can result in losing a lot of vein information due to improper scales. The filtered images with two scales corresponding to the two rows in the middle of Fig. 7(b) therefore are used for finger-vein feature extraction.

#### 4.2 Finger-vein feature extraction

According to the above discussion, the outputs of Gabor filters at the  $s$ th scale forms an 8-dimensional vector at each point in  $R(x, y)$ . For a pixel, its corresponding vector therefore is able to represent its local characteristic. For dimension reduction, an 8-dimensional vector based on the statistical information in a  $10 \times 10$  small block of a filtered image is constructed instead of a pixel-based vector. Thus, for a certain scale, 180 ( $18 \times 10$ ) vectors can be extracted from the filtered images in Gabor transform domain. Assume that  $H_{18 \times 10}$  represent the block matrix of a filter image, the statistics based on a block  $H_{ij}$  (a component of  $H$  in the  $i$ th column and the  $j$ th row, where  $i = 1, 2, \dots, 10$  and  $j = 1, 2, \dots, 18$ ) can be computed. Here, the average absolute deviation from the mean (AAD) (Jain et al., 2000)  $\delta_{ij}^s$  of the magnitudes of  $F_{sk}(x, y)$  corresponding to  $H_{ij}$  is calculated as

$$\begin{cases} \delta_{ij}^{sk} = \frac{1}{K} \sum_{H_{ij}} |F_{sk}(x, y)| - \mu_{ij}^{sk} \\ \mu_{ij}^{sk} = \frac{1}{K} \sum_{H_{ij}} |F_{sk}(x, y)| \end{cases} \quad (8)$$

where  $K$  is the number of pixels in  $H_{ij}$ ,  $\mu_{ij}^{sk}$  is the mean of the magnitudes of  $F_{sk}(x, y)$  in  $H_{ij}$ . Based on Eq. 8, the local statistics of filtered images are shown in Fig. 8, where the statistical information in a red box is used for finger-vein feature analysis.

Thus, the vector matrix at the  $s$ th scale of Gabor filter can be represented by

$$\mathbf{V}_s = \begin{bmatrix} \vec{v}_{11}^s & \cdots & \vec{v}_{1N}^s \\ \vdots & \vec{v}_{ij}^s & \vdots \\ \vec{v}_{M1}^s & \cdots & \vec{v}_{MN}^s \end{bmatrix}_{18 \times 10}, \quad (9)$$

where  $\vec{v}_{ij}^s = [\delta_{ij}^{s1}, \dots, \delta_{ij}^{s8}]$ . According to Eq. 9, uniting all the vectors together can form a  $2880(180 \times 2 \times 8)$  dimensional vector, which is not beneficial for feature matching.

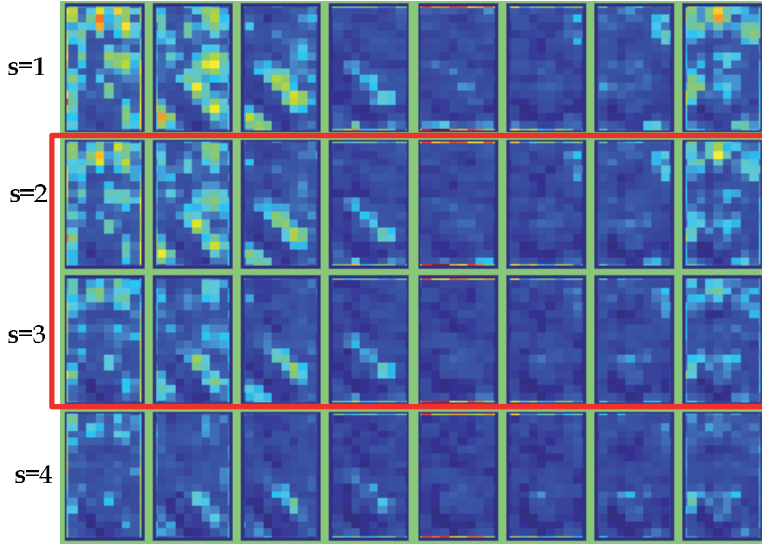


Fig. 8. The average absolute deviations (AADs) in  $[18 \times 10] \times 8$  blocks of the filtered finger-vein images in different scales and orientations.

Hence, based on  $\mathbf{V}_s$ , two new feature matrixes are constructed as

$$\mathbf{U}_s = \begin{bmatrix} \|\vec{v}_{11}^s\| & \cdots & \|\vec{v}_{1j}^s\| & \cdots & \|\vec{v}_{1N}^s\| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \|\vec{v}_{i1}^s\| & \cdots & \|\vec{v}_{ij}^s\| & \cdots & \|\vec{v}_{iN}^s\| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \|\vec{v}_{M1}^s\| & \cdots & \|\vec{v}_{Mj}^s\| & \cdots & \|\vec{v}_{MN}^s\| \end{bmatrix}, \quad (10)$$

and

$$\mathbf{Q}_s = \begin{bmatrix} \alpha_{1(1,2)}^s & \cdots & \alpha_{1(j,j+1)}^s & \cdots & \alpha_{1(N-1,N)}^s \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{i(1,2)}^s & \cdots & \alpha_{i(j,j+1)}^s & \cdots & \alpha_{i(N-1,N)}^s \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{M(1,2)}^s & \cdots & \alpha_{M(j,j+1)}^s & \cdots & \alpha_{M(N-1,N)}^s \end{bmatrix}, \quad (11)$$

where  $M = 18$ ,  $N = 10$ ,  $\|\bullet\|$  denotes the Euclidean norm of a vector,  $\alpha_{i(j,j+1)}^s$  is the angle of two adjacent vectors in the  $i$ th row. In this way, matrix  $\mathbf{U}_s$  is suitable for local feature representation, and matrix  $\mathbf{Q}_s$  is suitable for global feature representation. Hence, using  $\mathbf{U}_s$  and  $\mathbf{Q}_s$ , the local and global characteristics of a finger-vein image in the Gabor transform domain at the  $s$ th scale can be described sensibly and reliably. For convenience, the components of matrix  $\mathbf{U}_s$  and  $\mathbf{Q}_s$  are respectively rearranged by rows to form two 1D

feature vectors, here called FVCodes,

$$\begin{aligned} Z_s^U &= [\|\vec{v}_{11}^s\|, \dots, \|\vec{v}_{ij}^s\|, \dots, \|\vec{v}_{MN}^s\|]^T \\ Z_s^Q &= [\alpha_{1(1,2)}^s, \dots, \alpha_{i(j,j+1)}^s, \dots, \alpha_{M(N-1,N)}^s]^T. \end{aligned} \quad (12)$$

$(i = 1, 2, \dots, M; j = 1, 2, \dots, N; s = 2, 3)$

Since the components in  $U_s$  and  $Q_s$  are not of the same order of magnitude, it is not advisable to combine  $Z_s^U$  and  $Z_s^Q$  together for feature simplification in practices.

## 5. Finger-vein recognition

### 5.1 Finger-vein classification

As face, iris, and fingerprints recognition, finger-vein recognition is also based on pattern classification. Hence, the discriminability of the proposed FVCodes determines their reliability in personal identification. To test the discriminability of the extracted FVCodes at a certain scale, the cosine similarity measure classifier (CSMC) is adopted here for classification. The classifier is defined as

$$\begin{cases} \tau = \arg \min_{Z_s^k \in C_\kappa} \varphi(Z_s, Z_s^k) \\ \varphi(Z_s, Z_s^k) = 1 - \frac{Z_s^T Z_s^k}{\|Z_s\| \|Z_s^k\|} \end{cases} \quad (13)$$

where  $Z_s$  and  $Z_s^k$  respectively denote the feature vector of an unknown sample and the  $\kappa$ th class,  $C_\kappa$  is the total number of templates in the  $\kappa$ th class,  $\|\bullet\|$  indicates the Euclidean norm, and  $\varphi(Z_s, Z_s^k)$  is the cosine similarity measure. Using similarity measure  $\varphi(Z_s, Z_s^k)$ , the feature vector  $Z_s$  is classified into the  $\tau$ th class.

### 5.2 Decision-level fusion

According to section 4, we can see that the proposed FVCodes are different in content, dimension and scale for finger-vein feature description. Therefore, fusion of the matching results based on  $Z_s$  may improve the performance of identification. Nowadays, many approaches have been proposed in multi-biometrics fusion, such as Bayes algorithm, KNN classifier, OS-Rule, SVM classifier, decision templates algorithm, Dempster-Shafer (D-S) algorithm. Compared to other approaches, the D-S evidence theory works better in integrating multiple evidences for decision making. Details on D-S theory can be found in (Ren et al., 2009; Yager, 1987; Brunelli et al., 1995). Here, a overview of D-S theory is given briefly in the following.

Let  $\Theta = \{\theta_1, \dots, \theta_n\}$  be a frame of discernment, the power set  $2^\Theta$  be the set of  $2^n$  propositions (subsets) of  $\Theta$ . For an individual proposition  $A$  (or an evidence),  $m(A)$  is defined as basic belief assignment function (or mass function) if

$$\begin{cases} \sum_{A \in \Theta} m(A) = 1 \\ m(\emptyset) = 0 \end{cases} \quad (14)$$

For a subset  $A$  satisfying  $m(A) > 0$  is called focal element. Now, given two evidence sets  $E_1$  and  $E_2$  from  $\Theta$  with belief functions,  $m_1(\cdot)$  and  $m_2(\cdot)$ , let  $A_i$  and  $B_i$  be two focal elements respectively corresponding to  $E_1$  and  $E_2$ , the combination of the two evidences is given by

Dempster's orthogonal operator

$$m(A) = \begin{cases} \frac{\sum_{A_i \cap B_j = A} m_1(A_i)m_2(B_j)}{1 - K_a}, & A \neq \emptyset, \\ 0, & A = \emptyset \end{cases} \quad (15)$$

where

$$K_a = \sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j) \quad (16)$$

represents the conflict degree between two evidence sets. Traditionally,  $K_a$  usually leads to unimagined decision-making if it increases to a certain limit. Aiming to weaken the degree of evidence confliction, we have proposed an improved scheme in (Ren et al., 2009), and obtained better fusion results for fingerprint recognition. In view of finger-vein recognition application, a scheme based on D-S theory, as shown in Fig. 9, here is adopted to implement fusion in decision level.

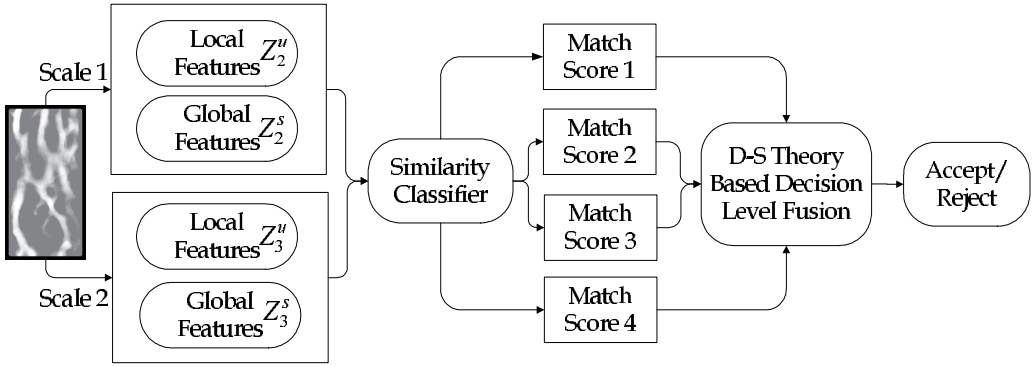


Fig. 9. The scheme of decision level fusion based on D-S theory.

First, for a certain extracted vector  $Z_s$ , match scores can be generated using CSMC. Based on the match scores, a basic belief assignment construction method proposed in (Ren et al., 2009) is then used for mass function formation. Thus, for a proposition  $A$ , mass function of each evidence is combined by

$$m(A) = (m_1 \oplus m_2 \oplus m_3 \oplus m_4)(A) \quad (17)$$

where  $\oplus$  represents the improved D-S combination rule proposed in (Ren et al., 2009), and  $m_1$ ,  $m_2$ ,  $m_3$  and  $m_4$  are the mass functions respectively computed from different evidence-match results using CSMC. The belief and plausibility committed to  $A$ ,  $Bel(A)$  and  $Pl(A)$ , can be computed as

$$\begin{cases} Bel(A) = \sum_{B \subset A} m(B) \\ Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \end{cases} \quad (18)$$

where  $Bel(A)$  represents the lower limit of probability and  $Pl(A)$  represents the upper limit. To give a reasonable decision, accept/reject, for incoming samples, an optimal threshold value related to the evidence mass functions should be found during training phase.

## 6. Experiments

### 6.1 Finger-vein image database

Because of the vacancy of common finger-vein image database for finger-vein recognition, we build an image database which contains 4500 finger-vein images from 100 individuals. Each individual contributes 45 finger-vein images from three different fingers: forefinger, middle finger and ring finger (15 images per finger) of the right hand. All images are captured using a homemade image acquisition system, as shown in Fig. 1. The captured finger-vein images are 8-bit gray images with a resolution of  $320 \times 240$ .

### 6.2 Performance evaluation of FVCode

Due to the high randomness of the finger-vein networks, the discriminability of the proposed FVCodes may embody not only in different individuals but also in different fingers of an identical individual. So, to investigate the differences among forefinger, middle finger and ring finger, 5 finger-vein images from one finger are selected as testing samples while the rest as training. Since the dimension of a FVCode is not high ( $\leq 180$ ), dimension reduction is not necessary for improving match efficiency. Moreover, the integrality of FVCodes describing finger-vein networks may be destroyed by dimension reduction. Therefore, the extracted FVCodes are directly used by CSMC for finger classification, some classification results are listed in Tables 1 and 2, where F\_finger, M\_finger and R\_finger, respectively represent forefinger, middle finger and ring finger, and FRR and FAR respectively represent false rejection rate and false acceptance rate.

| L-FVCodes |          | F-finger(500) | M-finger(500) | R-finger(500) | FAR(%) |
|-----------|----------|---------------|---------------|---------------|--------|
| s=2       | F-finger | 492(98.4%)    | 5             | 7             | 1.2    |
|           | M-finger | 6             | 489(97.8%)    | 10            | 1.6    |
|           | R-finger | 2             | 6             | 483(96.6%)    | 0.8    |
| s=3       | F-finger | 490(98.0%)    | 8             | 5             | 1.3    |
|           | M-finger | 6             | 486(97.2%)    | 13            | 1.9    |
|           | R-finger | 4             | 6             | 482(96.4%)    | 1.0    |
| FRR (%)   | s=2      | 1.6           | 2.2           | 3.4           |        |
|           | s=3      | 2.0           | 2.8           | 3.6           |        |

Table 1. Finger-vein image classification results using local FVCodes.

| G-FVCodes |          | F-finger(500) | M-finger(500) | R-finger(500) | FAR(%) |
|-----------|----------|---------------|---------------|---------------|--------|
| s=2       | F-finger | 488(97.6%)    | 10            | 8             | 1.8    |
|           | M-finger | 8             | 483(96.6%)    | 13            | 2.1    |
|           | R-finger | 4             | 7             | 479(95.8%)    | 1.1    |
| s=3       | F-finger | 492(98.4%)    | 4             | 8             | 1.2    |
|           | M-finger | 5             | 489(97.8%)    | 9             | 1.4    |
|           | R-finger | 3             | 7             | 483(96.6%)    | 1.0    |
| FRR (%)   | s=2      | 2.4           | 3.4           | 4.2           |        |
|           | s=3      | 1.6           | 2.2           | 3.4           |        |

Table 2. Finger-vein image classification results using global FVCodes.

From Tables 1 and 2, we can clearly see that forefingers hold the best capability in classification, while middle fingers appear better than ring fingers in correct classification rate (CCR) but lower than ring fingers in FAR. Moreover, for all test samples, local FVCodes can



achieve CCRs of 97.6% and 97.2% at two scales, which shows that the discriminability of local FVCodes decreases with scale ( $\sigma_2 > \sigma_3$ ). On the contrary, the discriminability of global FVCodes increases with scale since CCR increases from 96.67% at second scale ( $\sigma_2$ ) to 97.6% at third scale ( $\sigma_3$ ). These results show that 1) not only the proposed FVCode exhibits significant discriminability but also every finger is suitable for personal identification, and 2) fusion of local and global FVCodes in this two scales can improve the performance of finger-vein features in personal identification. Hence, the finger-vein images from different fingers can be viewed as from different individuals, and the proposed fusion scheme, as shown in Fig. 9, is desirable for recognition performance improvement. In the subsequent experiments, the database thus is expanded manually to 600 subjects and 15 finger-vein images per subject.

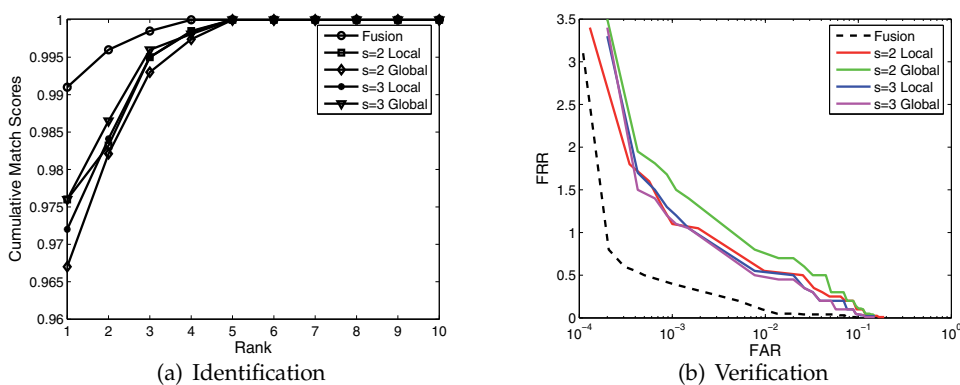


Fig. 10. The results of identification and verification

To obtain an unbiased estimate of the true recognition rate, a leave-one-out cross-validation method is used here. That is, leaving one example out sequentially and training on the rest accordingly, we conduct a classification of the omitted example. Consider that cumulative match scores (CMS) proposed in (Phillips et al., 2000) be more general in measuring classification performance, we therefore use it to evaluate the proposed finger-vein recognition algorithm. CMS can report the correct match probability (CMP) corresponding to the ranked  $n$  matches, and CCR is equivalent to the first CMP (rank = 1). Fig. 10 demonstrates the performance of the proposed method in identification and verification (for ranks up to 10). From Fig. 10, we can see that both local FVCodes and global FVCodes have their own merits in finger-vein recognition. Results from local FVCodes at the first scale ( $s = 2$ ) are somewhat better than those at the second ( $s = 3$ ), whereas this situation is reverse in results from global FVCodes. This is because, in the Gabor transform domain, feature variations are sufficiently represented locally at a big scale and globally at a small scale. Furthermore, the performance of both identification and verification for decision-level fusion is improved better, especially in FAR. This demonstrates that fusion of FVCodes at the two scales can improve the reliability of identification significantly. Hence, the finger-vein recognition technology is worthwhile to pay further attentions in security.

### 6.3 Comparison with existing methods

Unlike iris, face and fingerprints, research on finger-vein recognition is in an initial stage. For the purpose of comparison, we only implement some methods according to published papers (Miura et al., 2004; 2007; Mulyono and Horng, 2008; Zhang et al., 2006). Since it is difficult to obtain the detailed descriptions of the existing techniques in finger-vein recognition, only

algorithms on finger-vein feature extraction are implemented faithfully to the originals. The cosine similarity classifier here is used as a common measure to test the performance of the extracted finger-vein features. This is helpful to illustrate the capabilities of the existing finger-vein features in discriminability.

However, since some conditions may be uncertain in practice, our implemented versions may be inferior to the originals. Therefore, the comparison results only show the performance of previous methods approximately. Based on the expanded database (600 subjects, 15 images per subject), using 10 image of each subject as training samples and the rest as testing samples, we give the CCRs and FARs in Table 3.

| (%) | Miura('04) | Zhang('06) | Miura('07) | Lian('08) | L-FVCode(s = 2) |
|-----|------------|------------|------------|-----------|-----------------|
| CCR | 90.63      | 90.97      | 93.37      | 89.43     | 97.6            |
| FAR | 9.68       | 9.27       | 7.21       | 12.21     | 1.2             |

Table 3. Comparison results of the existing methods.

From Table 3, we can see that the proposed method is better than the previous in CCR and FAR. This is exciting indeed. Unfortunately, the results from the previous methods are significantly lower than those reported in the originals, especially in FARs. We think that two main reasons are responsible for this situation. First, the used image databases are different, which can directly lead to experimental deviations in practice. Second, the qualities of the used images may be different, that is to say, different image sensors generated different image qualities, which can significantly degrade different algorithms in finger-vein extraction. Hence, to make finger-vein recognition technology progress steadily, a standard finger-vein image database is indispensable in finger-vein based research community.

## 7. Conclusion

A method of personal identification based on finger-vein recognition has been discussed elaborately in this chapter. First, a stable finger-vein ROI localization method was introduced based on an interphalangeal joint prior. This is very important for finger-vein based practical application. Second, haze removal was adopted to restore the degraded finger-vein images, and background illumination was compensated from illumination estimation. Third, a bank of Gabor filters were designed to exploit the underlying finger-vein characteristics, and both local and global finger-vein features were extracted to form FVCodes. Finally, finger-vein classification was implemented using the cosine similarity classifier, and a fusion scheme in decision level was adopted to improve the reliability of identification. Experimental results have shown that the proposed method performed well in personal identification.

Undoubtedly, the method discussed in this chapter can not be optimal in accuracy and efficiency considering the development of finger-vein recognition technology. Therefore, this work should be just for reference to implement a finger-vein recognition task.

## 8. Acknowledgements

This work is jointly supported by NSFC (Grant No.61073143) and TJNSF (Grant No. 07ZCKFGX03700).

## 9. References

- Zharov, V.; Ferguson, S.; Eidt, J.; Howard, P.; Fink, L.; & Waner, M. (2004). Infrared Imaging of Subcutaneous Veins. *Lasers in Surgery and Medicine*, Vol.34, No.1, (January 2004), pp.56-61, ISSN 1096-9101

- Miura, N. & Nagasaka, A. (2004). Feature Extraction of Finger-Vein Pattern Based on Repeated Line Tracking and Its Application to Personal Identification. *Sachine Vision and Applications*, Vol.15, No.4, (October 2004), pp.194-203, ISSN 0932-8092
- Miura, N., Nagasaka, A. & Miyatake, T. (2007). Extraction of Finger-Vein Patterns Using Maximum Curvature Points in Image Profiles. *IEICE - Transaction on Information and Systems*, Vol.E90-D, No.8, (August 2007), pp.1185-1194, ISSN 0916-8532
- Mulyono, D. & Horng S.J. (2008). A Study of Finger Vein Biometric for Personal Identification. *Proceedings of International Symposium on Biometrics and Security Technologies*, pp.1-8, ISBN 978-1-4244-2427-6, Islamabad, Pakistan, April 23-24, 2008
- Zhang, Z.; Ma, S. & Han, X. (2006). Multiscale Feature Extraction of Finger-Vein Patterns Based on Curvelets and Local Interconnection Structure Neural Network. *Proceedings of 18th International Conference on Pattern Recognition*, pp.145-148, ISBN 0-7695-2521-0, Hongkong, China, August 20-24, 2006
- Vlachos, M. & Dermatas, E. (2008). Vein Segmentation in Infrared Images Using Compound Enhancing and Crisp Clustering. *Lecture Notes in Computer Science*, Vol.5008, (May 2008), pp.393-402, ISSN 0302-9743
- Jie, Z.; Ji, Q. & Nagy, G. (2007). A Comparative Study of Local Matching Approach for Face Recognition. *IEEE Transaction on Image Processing*, Vol.16, No.10, (October 2007), pp.2617-2628, ISSN 1057-7149
- Ma, L.; Tan, T.; Wang, Y. & Zhang, D. (2003). Personal Identification Based on Iris Texture Analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.25, No.12, (December 2003), pp.1519-1533, ISSN 0162-8828
- Jain, A.K.; Chen, Y. & Demirkus, M. (2007). Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.29, No.1, (Janary 2007) pp.15-27, ISSN 0162-8828
- Laadjel, M.; Bouridane, A.; Kurugollu, F. & Boussakta, S. (2008). Palmprint Recognition Using Fisher-Gabor Reature Extraction. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp.1709-1712, ISSN 1520-6149, Las Vegas, USA, March 31-April 4, 2008
- Shi, Y.H.; Yang, J.F. & Wu, R.B. (2007). Reducing Illumination Based on Nonlinear Gamma Correction. *Proceedings of IEEE International Conference on Image Processing*, pp.529-532, ISBN 978-1-4244-1437-6, San Antonio, Texas, USA, September 16-19, 2007
- Yang, J.F.; Shi, Y.H. & Yang, J.L. (2009). Finger-vein Recognition Based on a Bank of Gabor Filters. *Proceedings of Acian Conference on Computer Vision*, pp.374-383, ISBN 978-3-642-12306-1, Xi'an, China, September 23-27, 2009
- Yang, J.F.; Yang, J.L. & Shi, Y.H. (2009). Finger-vein segmentation based on multi-channel even-symmetric gabor filters. *Proceedings of IEEE International Conference on Intelligent Computing and Intelligent Systems*, pp.500-503, ISBN 978-1-4244-4754-1 Shanghai, China, November 20-22, 2009
- Yang, J.F.; Yang, J.L. & Shi, Y.H. (2009). Combination of gabor wavelets and circular gabor filter for finger-vein extraction. *Lecture Notes in Computer Science*, Vol.5754, (September 2009), pp.346-354, ISSN 0302-9743
- Choi, J.H.; Song, W.; Kim, T.; Lee, S. & Kim, H.C. (2009). Finger veinextraction using gradient normalization and principal curvature. *Proceedings of the SPIE*, Vol.7251, pp.1-9, ISBN 978-0-8194-7501-5, San Jose, USA, January 21, 2009
- Liu, Z.; Yin, Y.; Wang, H.; Song, S. & Li, Q. (2010). Finger Vein Recognition with Minifold Learning. *Journal of Network and Computer Application*, Vol.33, No.3, (May 2010), pp.275-282, ISSN 1084-8045

- Yang, J.F. & Li, X. (2010). Efficient Finger Vein Localization and Recognition. *Proceedings of 20th International Conference on Pattern Recognition*, pp.1148-1151, ISBN: 978-1-4244-7542-1, Istanbul, Turkey, Aug. 23-26, 2010.
- Ren, X.H.; Yang, J.F.; Li, H.H. & Wu, R.B. (2009). Multi-fingerprint Information Fusion for Personal Identification Based on Improved Dempster-Shafer Evidence Theory. *Proceedings of IEEE International Conference on Electronic Computer Technology*, pp.281-285, ISBN 978-0-7695-3559-3, Macau, China, February 20-22, 2009
- Yager, R. (1987). On the Dempster-Shafer Framework and New Combination Rules, *Information Sciences*, Vol.41, No.2, (March 1987), pp.93-138, ISSN 0020-0255
- Brunelli, R. & Falavigna, D. (1995). Person Identification Using Multiple Cues. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.17, No.10, (October 1995), pp.955-966, ISSN 0162-8828
- Daugman, J.G. (1985). Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by 2D Visual Cortical Filters. *Journal of the Optical Society of America*, Vol.2, No.7, (July 1985), pp.1160-1169, ISSN 1520-8532
- Lee, T.S. (1996). Image Representation Using 2D Gabor Wavelets, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.18, No.10, (October 1996), pp.1-13, ISSN 0162-8828
- Yang, J.; Liu, L.; Jiang, T. & Fan, Y. (2003). A Modified Gabor Filter Design Method for Fingerprint Image Enhancement. *Pattern Recognition Letters*, Vol.24, No.12, (August 2003), pp.1805-1817, ISSN 0167-8655
- Zhu, Z.; Lu, H. & Zhao, Y. (2007). Scale Multiplication in Odd Gabor Transform Domain for Edge Detection. *Journal of Visual Communication and Image Representation*, Vol.18, No.1, (February 2007), pp.68-80, ISSN 1047-3203
- Jain, A. K.; Prabhakar, S.; Hong, L.; Pankanti, S. (2000). Filterbank-Based Fingerprint Matching. *IEEE Transaction on Image Processing*, Vol.9, No.5, (May 2000), pp.846-859, ISSN 1057-7149
- Phillips, J.; Moon, H.; Rizvi, S. & Raue, P. (2000). The FERET Evaluation Methodology for Face Recognition Algorithms. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.22, No.10, (October 2000), pp.1090-1104, ISSN 0162-8828
- Delpy, D.T. & Cope, M. (1997). Quantification in Tissue Near-Infrared Spectroscopy. *Philosophical Transactions of the Royal Society of London, B*, Vol.352, No.1354, (June 1997), pp.649-659, ISSN 1471-2970
- Anderson, R. & Parrish, J. (1981). The Optics of Human Skin. *Journal of Investigative Dermatology*, Vol.77, no.1, (July 1981), pp.13-19, ISSN 0022-202X
- Xu, J.; Wei, H.; Li, X.; Wu, G. & Li, D. (2002). Optical Characteristics of Human Veins Tissue in Kubelka-Munk Model at He-Ne Laser in Vitro. *Journal of Optoelectronics Laser*, Vol.13, No.3, (March 2002), pp.401-404, ISSN 1005-0086
- Sassaroli, A. et al. Near-infrared spectroscopy for the study of biological tissue. Available from <http://ase.tufts.edu/biomedical/research/fantini/researchAreas/NearInfraredSpectroscopy.pdf>
- Hautière, N.; Tarel, J.P.; Lavenant, J. & Aubert, D. (2006). Automatic Fog Detection and Estimation of Visibility Distance through Use of an Onboard Camera. *Machine Vision and Applications*, Vol.17, No.1, (January 2006), pp.8-20, ISSN 0932-8092
- Jean-Philippe, T. & Nicolas, H. (2009). Fast Visibility Restoration from a Single Color or Gray Level Image. *Proceedings of International Conference on Computer Vision*, pp.2201-2208, ISBN 978-1-4244-4420-5, Kyoto, Japan, September 29-October 2, 2009
- Narasimhan, S.G. & Nayar, S.K. (2003). Contrast Restoration of Weather Degraded Images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.25, No.6, (June 2003), pp.713-724, ISSN 0162-8828

# Efficient Fingerprint Recognition Through Improvement of Feature Level Clustering, Indexing and Matching Using Discrete Cosine Transform

D. Indradevi

*Indra Ganesan College of Engineering, Tiruchirappalli,  
India*

## 1. Introduction

Fingerprint recognition refers to the automated method of verifying a match between two human fingerprints. Fingerprints are one of many forms of biometrics used to identify an individual and verify the identity. Because of their uniqueness and consistency over time, fingerprints have been used for over a century, more recently becoming automated biometric due to advancement in computing capabilities. Fingerprint identification is popular because of the inherent ease in acquisition, the numerous sources available for collection. In order to design a Fingerprint recognition system, the choice of feature extractor is very crucial and extraction of pertinent features from two-dimensional images of human finger plays an important role. A major challenge in Fingerprint recognition today is to select the low dimensional representative features and to reduce the search space for identification process.

The framework for the Fingerprint recognition system, as shown in Fig.1.1 consists of three phases. i) Feature extraction and representation phase, ii) Featurelevel Clustering, iii) Indexing and Fingerprint Matching. The three phases of fingerprint recognition framework are detailed in the following sections.

### 1.1 Fingerprint feature extraction

To detect the machine-readable representation completely capture the invariant and discriminatory information in the input measurements is the most challenging problem in representing fingerprint data. This representation issue constitutes the essence of system design and has far reaching implications on the design of the rest of the system. The unprocessed measurement values are typically not invariant over the time of capture and there is a need to determine salient features of the input measurement which both discriminate between the identities as well as remain invariant for a given individual. Thus, the problem of representation is to determine a measurement(feature) space which is invariant for input signals belonging to the same identity and which differ maximally for those belonging to different identities (higher interclass) variation and low interclass variation.

In [1], fingerprint features are classified into three classes. Level 1 features show macro details of the ridge flow shaped, Level 2 features (minutiae point) are discriminative enough for recognition, and Level 3 features (pores) complement the uniqueness of Level 2 features. The popular fingerprint representation scheme have evolved from an intuitive system developed by forensic experts who visually match the fingerprints. These schemes are either based on predominantly local landmarks (e.g. minutiae-based fingerprint matching systems [1]) or exclusively global information (fingerprint classification based on the Henry System [2]). The minutiae-based automatic identification techniques first locate the minutiae points and then match their relative placement in a given finger and the stored template. The global representation of fingerprints (e.g. whorl, Left loop, right loop, arch, and tented arch) is typically used for indexing, and does not offer good individual discrimination. The global representation schemes of the fingerprint used for classification can be broadly categorized into four main categories: (i) knowledge-based, (ii) structure-based, (iii) frequency-based, and (iv) syntactic. The different type of fingerprint patterns are described below.

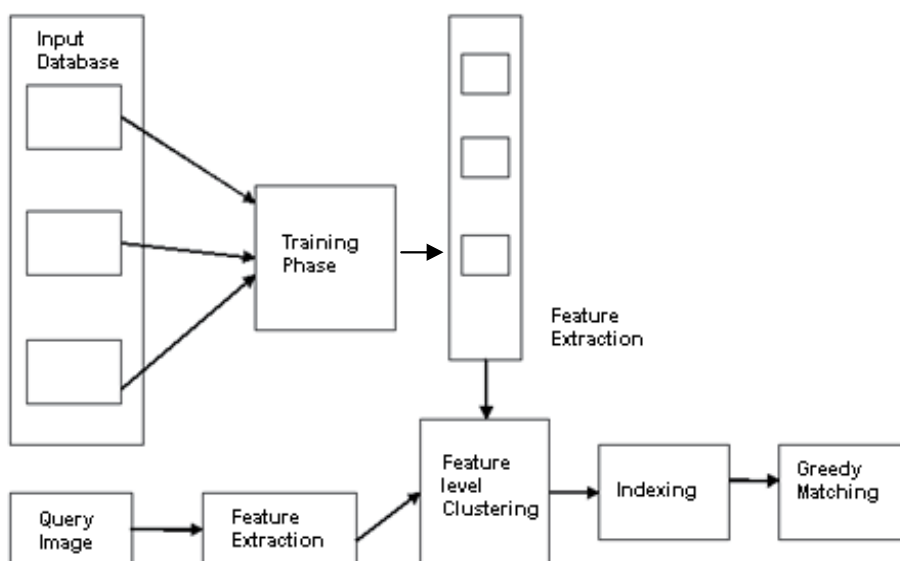


Fig. 1.1. Fingerprint Recognition System.

### 1.1.1 Global ridge pattern

A fingerprint is a pattern of alternating convex skin called ridges and concave skin called valleys with a spiral-curve-like line shape (Fig. 1.2). There are two types of ridge flows: the pseudo-parallel ridge flows and high-curvature ridge flows which are located around the core point and/or delta points. This representation relies on the ridge structure, global landmarks and ridge pattern characteristics. The commonly used global fingerprint features are:

- Singular points - discontinuities in the orientation field. There are two types of singular point. A core is the uppermost of the innermost curving ridge [3], and a delta point is the junction point where three ridge flows meet. They are usually used for fingerprint registration, fingerprint classification.

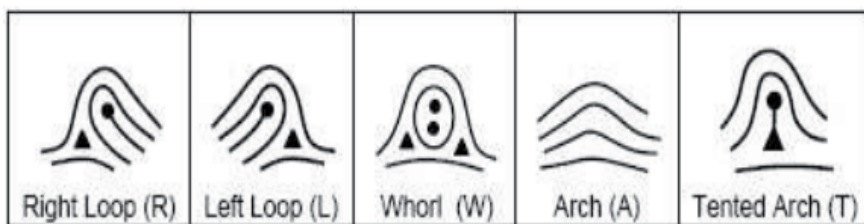


Fig. 1.2. Global Fingerprint Ridge Flow Patterns.

- Ridge orientation map - local direction of the ridge-valley structure. It is commonly utilized for classification, image enhancement, minutia feature verification and filtering.
- Ridge frequency map - the reciprocal of the ridge distance in the direction perpendicular to local ridge orientation. It is formally defined in [4] and is extensively utilized for contextual filtering of fingerprint images.

This representation is sensitive to the quality of the fingerprint images. However, the discriminative abilities of this representation are limited due to absence of singular points.

### 1.1.2 Local ridge pattern

This is the most widely used and studied fingerprint representation. Local ridge details are the discontinuities of local ridge structure referred to as minutiae. Sir Francis Galton was the first person who observed the structures and permanence of minutiae. Therefore, minutiae are also called „Galton details“. They are used by forensic experts to match two fingerprints. There are about 150 different types of minutiae categorized based on their configuration. Among these minutiae types, ridge ending and ridge bifurcation are the most commonly used, since all the other types of minutiae can be seen as combinations of ridge ending and ridge bifurcations. Some minutiae are illustrated in Fig.1.3.

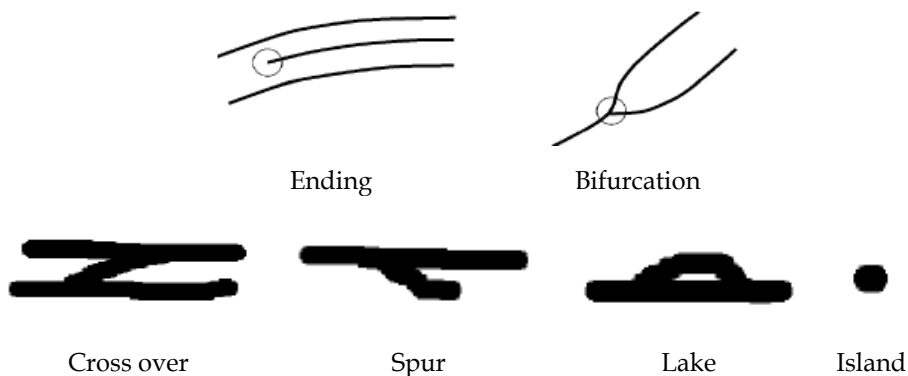


Fig. 1.3. Common Minutiae Types.

### 1.1.3 Intra-ridge detail

On every ridge of the finger epidermis, there are many tiny sweat pores and other permanent details. Pores are considered to be highly distinctive in terms of their number, position, and shape. However, extracting pores is feasible only in high-resolution fingerprint

images and with very high image quality. Therefore, this kind of representation is not adopted by currently deployed automatic fingerprint identification systems. Some fingerprint identification algorithm (such as FFT) may require so much computation. Discrete Cosine Transform based algorithm may be the key to making a low cost fingerprint identification system.

### **1.2 Feature level clustering**

With the increase in the size of the Fingerprint database, reliability and scalability issues become the bottleneck for low response time, high search and retrieval efficiency in addition to accuracy. Fingerprint classification refers to the problem of assigning fingerprints to one of several prespecified classes. It is an important stage in automatic fingerprint identification systems (AIFS) because it significantly reduces the time taken in identification of fingerprints, especially where accuracy and speed are critical. Traditionally Fingerprint identification systems claim identity of an individual by searching templates of all users enrolled in the database. These comparisons increase the data retrieval time along with the error rates. Thus a size reduction technique must be applied to reduce the search space and thus improve the efficiency. Conventionally databases are indexed numerically or alphabetically to increase the efficiency of retrieval. However, Fingerprint databases do not possess a natural order of arrangement which negates the idea to index them alphabetically/numerically. Reduction of search space in databases thus remains a challenging problem. Considering the classification issues [5], several methods have been proposed in the past couple of years to address the fingerprint classification issues. Most of these methods classify the images based on the ridges, local feature (i.e. minutiae) and global features (i.e. singular points). Model based approaches based on the global features (singular points) of the fingerprints have been found more effective in classifying the fingerprints into different known classes. Structure-based approaches based on the estimated orientation field in a fingerprint image can be found capable to classify the images into one of the five classes. The role of the estimated orientation field for fingerprint classification is generic. However, if the images are of poor quality then the orientation field estimation could not be done properly. Also, in such case difficulties encountered during extraction of other features like minutiae, finger code, Poincare index for singular points detection etc. Exclusive fingerprint classification is a traditional approach that has been widely investigated in the literature [5-7]. It classifies each fingerprint exclusively into one of the predefined classes such as Henry classes. Although it has some advantages such as human-interpretability, fast retrieval and rigid database partitioning, most automated classification algorithms are able to classify fingerprints into only four or five classes. Moreover, fingerprints are not evenly distributed in these classes. Thus, the exclusive classification cannot sufficiently narrow down the search of database. Most of the existing fingerprint classification approaches make use of the orientation image [8]. The main drawback of classification is that it is the supervised method where number of classes has to be known in advance. Further the data within each class is not uniformly distributed so the time required to search some classes is comparatively large. The limitations of classification can be addressed with unsupervised approach known as Clustering. It involves the task of dividing data points into homogeneous classes or clusters so that items in the same class are as similar as possible and items in different classes are as dissimilar as possible. Intuitively it can be visualized as a form of data compression, where a large number of samples are converted into a small number of representative prototypes.



### 1.3 Fingerprint database indexing

Identification refers to determine the identity of an individual from the database of persons available to the system. With the increase in the size of database, the number of false acceptances grows geometrically. Further, the time required to claim an identification is directly proportional to the size of the database. Thus efficiency of such systems can be improved either by minimising the error rates or by reducing the search space. The former is dependent on the efficiency of the algorithm and cannot be reduced considerably. Hence more emphasis can be laid on reducing the search space for improving performance as it is not possible to retrieve each element from probe set and compare with all elements of gallery set to determine the identity. There already exist few indexing schemes to partition the biometric database.

Rajiv Mukherjee et al [9] reported that iris database indexing generated from mean vector constructed from each row of the Iris Code and iris feature vector. Unsupervised clustering is performed on these index vectors. Puhan et al [10] proposed iris indexing based on the iris color with chrominance components is generated as index code. Amit et al. [11] reported Indexing hand geometry database by pyramid technique. Gupta et al. [12] proposed an efficient indexing scheme for binary feature template with the help of B+ tree. Umarani et al. [13] employed a modified B+ tree biometric database indexing. The higher dimensional feature vector is projected to lower dimension and the reduced dimensional feature vector is used to index the database by forming B+ tree. All of the aforementioned tree-like data structures lead to "curse of dimensionality". While they work reasonably well in a 2-D or 3-D space, as the dimensionality of the data increases, the query time and data storage would exhibit an exponential increase.

### 1.4 Fingerprint matching

In order to improve fingerprint recognition performance, many techniques have been designed. The most popular matching strategy for fingerprint verification is minutiae matching. The simplest pattern of the minutiae-based representation consists of a set of minutiae, including ridge endings and bifurcation defined by their spatial coordinates. Each minutiae is described by its spatial location associated with the direction and minutiae type. In correlation based matching, correlation between the corresponding pixels is computed for different alignments. In feature based matching features of the fingerprint ridge pattern, texture information may be extracted more reliably than minutiae.

## 2. Significant fingerprint feature extraction with discrete cosine transform

### 2.1 Discrete cosine transform

In this section the discrete cosine transform that can be used for analyzing the Fingerprint image and to extract the local features in the transform domain is described. The Discrete cosine transform that has already been well established by Chaur-Heh for image compression [14] is extended in this model to extract the deterministic fingerprint feature. In order to extract the features of fingerprint image, block DCT-based transformation is employed. Each image is divided into sub blocks with  $(N \times N)$  size. There are  $n \times n$  coefficients in each block after DCT is applied. Only some of the DCT coefficients are to be computed for feature extraction, and they are enough to represent the information that are needed from the block. The equation used for the DCT calculation for each pixel is given as follows:

$$F(u,v) = \alpha(u)\alpha(v) \frac{2}{n} \sum_{y=0}^{n-1} \sum_{x=0}^{n-1} f(x,y) \cos \frac{(2x+1)u\pi}{2n} \cos \frac{(2y+1)v\pi}{2n} \quad (2.1)$$

$$\alpha(u) = \frac{1}{\sqrt{2}} \quad \text{if } u=0 \quad (2.2)$$

$$\alpha(u) = 1 \quad \text{otherwise}$$

$F(u,v)$  is the DCT domain representation of  $f(x,y)$  image and  $u,v$  represent vertical and horizontal frequencies that have values ranges from 0 to 7. The DCT coefficients reflect the compact energy of different frequencies. The first coefficient (0, 0), called DC, is the mean of visual gray scale value of pixels of a block. And the other values denoted as AC coefficients, representing each spatial characteristic including vertical, horizontal and diagonal patterns. Having described the discrete cosine transform, the feature extraction technique is presented in the next section.

## 2.2 Statistical feature extraction

The feature set is based on a measure of DCT coefficients excluding DC (0, 0). The features derived from the DCT computation are limited to an array of summed spectral energies within a block in frequency domain. The given input image is applied to DCT by considering non-overlapping  $N$  by  $N$  blocks (where  $N=8$ ).

To ensure adequate representation of the image, each block non-overlaps horizontally and vertically with the neighboring blocks, thus for an image which has  $N_R$  rows and  $N_C$  columns, there are  $N_B$  blocks found by following formula:

$$N_B = \left[ \frac{N_R}{N} \times \frac{N_C}{N} \right] \quad (2.3)$$

The input image is divided into sub-image blocks:  $F(u, v)$ ,  $u = 1$  to  $N$  and  $v = 1$  to  $N$ , and then the DCT is performed independently on the sub-image blocks. For each sub-block considering AC coefficients, the DCT feature set is extracted. The proposed feature descriptor  $F$  consisting of the following features are denoted as feature set.

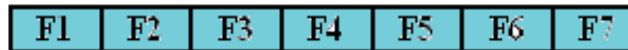


Fig. 2.1. Feature Set.

The features are denoted as Signal Energy, Edge Orientation, Ridge Frequency Estimation, Ridge Orientation Estimation, Non-Coherence Factor, Angular Bandwidth and Directional Strength Orientation respectively.

### 2.2.1 Signal energy

DCT is a reversible transform which obey the "Energy Preservation Theorem" - Total energy in pre-transform domain is equal to total energy in post transform domain [15]. The energy is one of the image properties using signal processing technique, and it means

characteristics of images. From the result of DCT transformed from image  $F(u,v)$  of size  $N \times N$ , the energy using DCT coefficients is defined as following equation [16].

$$F1 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |F(u,v)|^2 \quad (2.4)$$

### 2.2.2 Edge orientation

Edge is a strong feature for characterizing an image, which can be used to construct an important clue to understand the content of fingerprint image. The edge information from an image can be directly extracted by some simple measurements on the AC coefficients of each block in the compressed domain. Thus the following relationship determines the presence of edge in a block. Accurate edge orientation information is computed as follows:

$$F2 = \tan \theta = \frac{\sum_{v=1}^{N-1} F_{(0,v)}}{\sum_{u=1}^{N-1} F_{(0,u)}} \quad (2.5)$$

### 2.2.3 Ridge frequency estimation

The ridge frequency estimates the curve dominant flow in a block. The highest DCT peak of highest frequency spectrum is given by [16]:

$$F3 = \sqrt{u_0^2 + v_0^2} \quad (2.6)$$

where  $u_0, v_0$  is the coordinate of the highest DCT peak value.

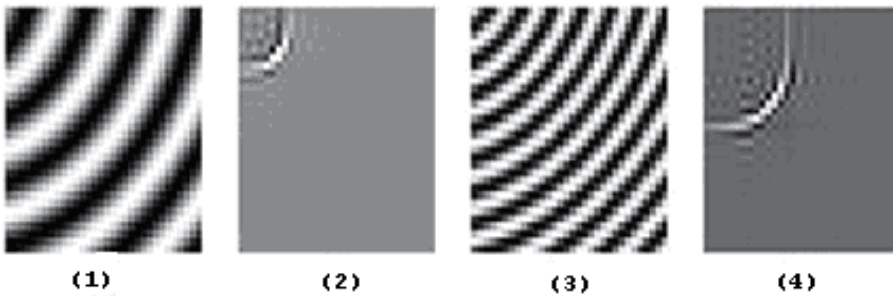


Fig. 2.2. 1) and 3) Represents blocks of a fingerprint image with different frequency, 2) and 4) DCT coefficients of (1) and (3) respectively.

### 2.2.4 Ridge orientation estimation

The dominant orientation of parallel ridges, are closely related to a peak-angle in DCT coefficients, where  $\phi$  is measured counterclockwise (if  $\phi > 0$ ) from the horizontal axis to the terminal side of the highest spectrum peak of highest frequency (DC is excluded). However  $\phi$  and  $\theta_0$  relationship is not one-to-one mapping. The ridge orientation which

varies in the range of 0 to  $\pi/2$  is projected into the peak-angle  $\theta$  which varies in the range of 0 to  $\pi$ . Relationship between ridge orientation  $\phi$  in spatial domain and peak angle  $\phi_0$  in frequency domain is given by:

$$F4 = \tan^{-1} \frac{v_0}{u_0} \quad (2.7)$$

Let F4 be  $\phi_0$ , then  $\phi_0 = \left| \frac{\pi}{2} - \theta_o \right|$  where  $0 \leq \theta_o \leq \pi$

### 2.2.5 Non-coherence factor

The factor represents how wide ridge orientation can be in the block that has more than one dominant orientation. This factor is in the range of 0 to 1, where 1 represents highly non-coherence or highly curved region and 0 represents orientation region. The non-coherence factor is given by:

$$F5 = \frac{\sum_{i,j \in N} \left| \sin(\theta_{u_c, v_c} - \theta_{u_i, v_j}) \right|}{N \times N} \quad (2.8)$$

where  $u_c, v_c$  is the center position of the block,  $u_i, v_j$  is the  $i^{\text{th}}$  and  $j^{\text{th}}$  position of neighbor blocks within  $N \times N$ .

### 2.2.6 Angular bandwidth

High ridge curvature occurs where the local ridge orientation changes rapidly, i.e. near cores and deltas. Away from these singular points, ridge curvature tends asymptotically to zero. In regions of higher curvature, a wider range of orientations is present. Away from singular points, the angular bandwidth is  $\pi/8$ . The angular bandwidth must equal to  $\pi$  is calculated by the following equation:

$$F6 = \sin^{-1}(F5) \quad (2.9)$$



Fig. 2.3. Core Region.

### 2.2.7 Directional strength orientation

To identify the quadrant and avoid influence of interference, two 2-D perpendiculars diagonal vectors  $D_1$  and  $D_2$  are formed with size of  $5 \times 3$  pixels center at the peak position.

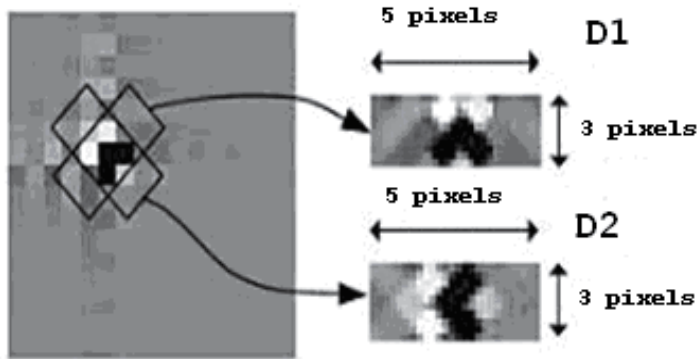


Fig. 2.4. 2-D perpendicular diagonal vectors at 45° and 135° .

The average directional strengths of each vector  $DS_i$

$$Max_{n=-1,0,1} \left| \frac{\sum_{m=-2}^2 D_{u_0+m, v_0+n}}{5} \right| = DS_i, \text{ where } i=1,2 \quad (2.10)$$

Then the quadrant can be classified and the actual fingerprint ridge orientation can be identified as [16]:

$$F7 = \begin{cases} \pi / 2 - F4 & \text{where } DS_1 \geq DS_2 \\ \pi - (\pi / 2 - F4) & \text{otherwise} \end{cases} \quad (2.11)$$

where  $u_0, v_0$  is the coordinate of the highest DCT peak value,  $u_c, v_c$  is the center position of the block,  $u_i, v_j$  is the  $i^{th}$  and  $j^{th}$  position of neighbor blocks within  $N \times N$ , perpendicular diagonal set  $D_1, D_2$  with size  $5 \times 3$  pixels and  $DS_1, DS_2$  average directional strength respectively.

The feature is extracted for each  $N \times N$  block and its average of the total number of blocks of that image is finally derived. The Non-Coherence Factor represents ridge orientation of how wide can be in that block. The maximum value occurred on a block represents highly curved region which is considered as a reference block for indexing. The significant statistical feature values are summarized in the following table.

| Feature1 | Feature2 | Feature3 | Feature4 | Feature5 | Feature6 | Feature7 | IndexCode  |
|----------|----------|----------|----------|----------|----------|----------|------------|
| 680      | 60       | 83       | 2.0965   | 43       | 0.6839   | 21       | 9200002114 |
| 1205     | 62       | 92       | 1.9354   | 42       | 0.7269   | 20       | 1400000711 |
| 1089     | 65       | 83       | 2.2276   | 37       | 0.6879   | 20       | 1896304005 |
| 1116     | 64       | 87       | 2.4107   | 41       | 0.6758   | 21       | 2120284181 |
| 1127     | 64       | 88       | 2.2014   | 43       | 0.6962   | 19       | 4783545190 |

Table 2.1. Significant Statistical Feature Set.

### 3. Clustering categorical fingerprint features

#### 3.1 Robust clustering (ROCK)

The main aim of cluster analysis is to assign objects into groups (clusters) in such a way that two objects from the same cluster are more similar than two objects from different clusters. There are different ways of classifying the cluster analysis methods. It can distinguish partitioning methods (denoted as flat) which optimize assignment of the objects into a certain number of clusters, and methods of hierarchical cluster analysis with graphical outputs which make assignment of objects into different numbers of clusters possible. The partitioning method includes k-means, fuzzy c mean, k-medians, k-medoids. The hierarchical method can be agglomerative (step-by-step clustering of objects and groups to larger groups) or divisive (step-by-step splitting of the whole set of objects into the smaller subsets and individual objects).

The ROCK [17], algorithm is an agglomerative hierarchical clustering algorithm for clustering categorical data. ROCK is based on links between pairs of data objects and the agglomerative process of merging clusters terminates either when there is no pair of clusters with links between them or when the required number clusters is obtained. Informally, the number of links between two tuples is the number of common neighbors they have in the dataset. After an initial computation of the number of links between the data objects, the algorithm starts with each cluster being a single object and keeps merging clusters based on a goodness measure for merging. The merging is continued till one of the following two criteria is reached.

- A specified number of clusters are obtained (or)
- No links remain between the clusters.

Instead of working on the whole dataset, clusters a sample randomly drawn from the dataset and then partitions the entire dataset based on the clusters from the sample.

#### 3.2 Categorical fingerprint feature representation

In the cluster analysis phase, a cluster representative is generated to characterize the clustering result. However, in the categorical domain, there is no common way to decide cluster representative. Based on the assumption, the numerical data are converted into categorical data [22] as described below:

| Orientation Range (in degrees) | Angle (in degrees) | Label |
|--------------------------------|--------------------|-------|
| 0 to 22.5 & 157.5 to 180       | 0                  | Q1    |
| 22.5 to 67.5                   | 45                 | Q2    |
| 67.5 to 112.5                  | 90                 | Q3    |
| 112.5 to 157.5                 | 135                | Q4    |

Table 3.1. Orientation Label.

| Range | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
|-------|----|----|----|----|----|----|----|----|----|
| Label | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 |

Table 3.2. Frequency Label.

Feature 1: Signal Energy

By considering the first 5 minimum values from the training test, image is set to low else high.

$$F1 = \{high, low\}$$

Feature 2: Edge Orientation

$$F2 = \{Q1, Q2, Q3, Q4\}$$

Feature 3: Ridge Frequency

$$F3 = \{F1, F2, F3, F4, F5, F6, F7, F8, F9\}$$

Feature 4: Ridge Orientation

$$F4 = \{Q1, Q2, Q3\}$$

Algorithm: **Categorical Data**

Input: Numerical Data  $\{N_1, N_2, N_3, \dots, N_n\}$

Output: Categorical Data  $\{C_1, C_2, C_3, \dots, C_n\}$

Step 1: Read the dataset N

Step 2: Convert Numerical Data into Categorical Data based on condition

Step 3: Assign Label to each numerical data

| IMAGE NAME | SIGNAL ENERGY | EDGE ORIENTATION | RIDGE FREQUENCY | RIDGE ORIENTATION |
|------------|---------------|------------------|-----------------|-------------------|
| Image 1    | High          | Q3               | F1              | Q2                |
| Image 2    | High          | Q2               | F2              | Q3                |
| Image 3    | High          | Q4               | F2              | Q1                |
| Image 4    | High          | Q1               | F3              | Q1                |
| Image 5    | High          | Q2               | F4              | Q2                |

Table 3.3. Categorical Data.

### 3.3 Feature level clustering algorithm

#### 3.3.1 Robust clustering

Algorithm 1: ROCK

Procedure: Cluster(S,k)

Let S = {set of n sampled feature}

k = {number of clusters}

Input: Feature vector Fv

Output: Clusters C1, C2....., Cn

---

**Step1:** Read Fv

**Step2:** Compute links

link:= compute\_links(S)

**Step3:** for each s  $\in$  S do

q[s]:=build\_local\_heap(link,s)

---

---

**Step4:** Q:=build\_global\_heap(S,q)

**Step5:** Find the best cluster to merge

while size(Q)>k

do

u:=max(Q)

v:=max(q[u])

delete(Q,v)

w:=merge(u,v)

for each x ∈ q(u) U q(v)

do

link[x,w]:=link[x,u]+link[x,v]

delete(q[x],u,v)

insert(q[x],w,g[x,w])

insert(q[w],x,g[x,w])

update(Q,x,q[x])

insert(Q,w,q[w])

deallocate()

**Step 6:** End

---

### 3.3.2 K-Mean clustering for categorical data

Let  $X = (F_1, F_2, F_3, F_4, F_5)$  be a feature vector where the scalars  $F_1, F_2, F_3, F_4, F_5$  represent the individual features of each object. Let  $V = (v_1, v_2, \dots, v_k)$  be the set of  $k$  clusters where  $v_1, v_2, \dots, v_k$  represent the centroid of each cluster [18].

The centroid of a cluster is defined as:

$$v_k = \frac{\sum_{i=1}^N F_i \in v_k}{|v_k|} \quad (3.1)$$

The above formula simply computes the mean or the average of the values of features belonging to the particular cluster.

The objective function to be minimized is defined below:

$$J(X, V) = \min \sum_{k=1}^K \sum_{i=1}^{N_k} d(F_i, v_k) \quad (3.2)$$

where Euclidean Distance:

$$d(v_k, F_i) = \sqrt{\sum_{j=1}^N (F_{ij} - v_{kj})^2}$$

For categorical attributes, the similarity measure is defined as:

$$d(F_i, v_k) = \sum_{i=1}^N (d_{ij}^n - v_{kj}^n)^2 + wt \times \sum_{i=1}^C \delta(d_{ij}^c, v_{kj}^c) \quad (3.3)$$



where  $\delta(a,b) = 0$  for  $a=b$  and  $\delta(a,b) = 1$  for  $a \neq b$ .

$d_{ij}^n$  and  $v_{kj}^n$  are values of numeric attributes and  $d_{ij}^c$  and  $v_{kj}^c$  are values of categorical attributes for the object  $o$  and cluster  $k$ .  $N_k$  is the number of elements in cluster  $k$ . The equation can be rewritten as:

$$J(X, V) = \sum_{k=1}^p \sum_{i=1}^{N_k} \left( \sum_{j=1}^N (d_{ij}^{na} - v_{kj}^{na})^2 + wt \times \sum_{j=1}^C \delta(d_{ij}^c, v_{kj}^c) \right) \quad (3.4)$$

### Algorithm 2: K Means- for Categorical Data

#### Inputs:

$I = \{i_1, i_2, \dots, i_k\}$  (Instances to be clustered)

$N$  (Number of clusters)

#### Outputs:

$C = \{c_1, \dots, c_n\}$  (Cluster Centroids)

$m: I \rightarrow C$  (Cluster Membership)

---

**Step 1:** Set  $C$  to initial value (e.g. random selection of  $I$ )

**Step 2:** For each  $i_j \in I$

$$m(i_j) = \arg \min_{k \in \{1..n\}} d(i_j, c_k) + wt \times \sum_{k=1}^C \delta(d_{ij}^c, C_k^c)$$

End

**Step 3:** While  $m$  has changed

For each  $j \in \{1..n\}$

Recompute  $i_j$  as the centroid of  $\{i \mid m(i) = j\}$

End

Return  $C$

---

### 3.3.3 Fuzzy C-means clustering for categorical data

A variant of K-Means that performs fuzzy clustering is called Fuzzy C-Means (FCM). The output of a fuzzy clustering is not a partition but still a clustering with each object having a certain degree of membership to a particular cluster [20]. The objective function being minimized in FCM is given below:

$$J(X, V) = \min \sum_{k=1}^K \sum_{i=1}^{N_k} (U_{ki}^m) d(F_i, v_k) \quad (3.5)$$

where  $d(v_k, F_i)$  is the distance function similar to the one discussed above and  $u_{ki}$  is the degree of membership of object  $i$  to cluster  $k$  subject to the following constraints:

$$u_{ki} \in [0, 1] \text{ and } \sum_{k=1}^K u_{ki} = 1, \forall i,$$

which means that the sum of the membership values of the objects to all of the fuzzy clusters must be one.  $m$  is the fuzzifier which determines the degree of fuzziness of the resulting

clusters. The objective function can be minimized by using Lagrange multipliers. By taking the first derivatives of  $J$  with respect to  $u_{ik}$  and  $v_k$  and setting them to zero results in two necessary but not sufficient conditions for  $J$  to be at the local extrema.

The result of the derivation is given below:

$$v_k = \frac{\sum_{i=1}^N u_{ki}^m x_i}{\sum_{i=1}^N u_{ki}^m} \quad (3.6)$$

which computes the centroid of cluster  $k$ .

$$u_{ki} = \frac{1}{\sum_{j=1}^K (d(v_k, x_i) / d(v_k, x_j))^{m-1}} \quad (3.7)$$

computes the membership of object  $i$  to cluster  $k$ . For categorical data, Let  $X = (F_1, F_2, F_3, F_4, F_5)$  be a set of categorical objects.

Let  $F_k = [F_{k1}, F_{k2}, \dots, F_{kp}]$  and  $F_l = [F_{l1}, F_{l2}, \dots, F_{lp}]$  be two categorical objects.

The matching dissimilarity between them is defined as:

$$D(F_k, F_l) = \sum_{j=1}^p \delta(F_{kj}, F_{lj}) \quad (3.8)$$

where  $\delta(a, b) = 0$  for  $a=b$  and  $\delta(a, b) = 1$  for  $a \neq b$ .

### Algorithm 3: Fuzzy C Mean- for Categorical Data

**Inputs:**

$I = \{i_1, i_2, \dots, i_n\}$  (Instances to be clustered)

$N$  (Number of clusters)

**Outputs:**

$C = \{c_1, \dots, c_n\}$  (Cluster Centroids)

---

Step 1: Select  $m$  and  $\epsilon$  to initial value.

Step 2: Initialize the modes  $v_i$  ( $1 \leq i \leq c$ )

Step 3: Calculate the membership degrees  $u_{ik}$  and determine  $J_m(U, V)$

Step 4: Set  $J_m^{old}(U, V) = J_m(U, V)$

Step 5: Update the cluster centers  $v_i$  ( $1 \leq i \leq c$ )

Step 6: Calculate the membership degrees  $u_{ik}$  ( $1 \leq i \leq c, 1 \leq k \leq n$ )

Step 7: Update the  $J_m(U, V)$

Step 8: If  $|J_m(U, V) - J_m^{old}(U, V)| \leq \epsilon$  Go to Step 4

---

### 3.4 Similarity measures

The algorithm finds all the combinations of feature values in an object, which represent a subset of all the attribute values, and then groups the database using [21] the similarity of these combinations by constructing the Jaccard coefficient matrix. Objects in a cluster have not only similar attribute value sets but also strongly associated attribute values (feature). The similarity measure of the ROCK algorithm is found using Jaccard Coefficient

$$sim(o_i, o_j) = \frac{o_i \cap o_j}{o_i \cup o_j}$$

The measure is computed using Single, Complete and Average Linkage.

|           |           |            |            |            |            |            |
|-----------|-----------|------------|------------|------------|------------|------------|
|           | <b>I1</b> | <b>I2</b>  | <b>I3</b>  | <b>I4</b>  | <b>I5</b>  | <b>I6</b>  |
| <b>I1</b> | 1         | <b>0.4</b> | <b>0.4</b> | <b>0.6</b> | <b>0.8</b> | <b>0.4</b> |
| <b>I2</b> | 0.4       | 1          | <b>0.2</b> | <b>0.4</b> | <b>0.4</b> | <b>0.8</b> |
| <b>I3</b> | 0.4       | 0.2        | 1          | <b>0.6</b> | <b>0.4</b> | <b>0.8</b> |
| <b>I4</b> | 0.6       | 0.4        | 0.6        | 1          | <b>0.2</b> | <b>0.4</b> |
| <b>I5</b> | 0.8       | 0.4        | 0.4        | 0.2        | 1          | <b>0.6</b> |
| <b>I6</b> | 0.4       | 0.8        | 0.8        | 0.4        | 0.6        | 1          |

|              |           |              |              |           |
|--------------|-----------|--------------|--------------|-----------|
|              | <b>I1</b> | <b>I2/I3</b> | <b>I4/I5</b> | <b>I6</b> |
| <b>I1</b>    | 1         | 0.4          | 0.6          | 0.4       |
|              |           |              |              |           |
| <b>I2/I3</b> | 0.4       | 1            | 0.2          | 0.4       |
| <b>I4/I5</b> | 0.6       | 0.2          | 1            | 0.4       |
| <b>I6</b>    | 0.4       | 0.4          | 0.4          | 1         |

|                    |           |                     |           |
|--------------------|-----------|---------------------|-----------|
|                    | <b>I1</b> | <b>I2/I3/ I4/I5</b> | <b>I6</b> |
| <b>I1</b>          | 1         | 0.2                 | 0.4       |
| <b>I2/I3/I4/I5</b> | 0.2       | 1                   | 0.4       |
| <b>I6</b>          | 0.4       | 0.4                 | 1         |

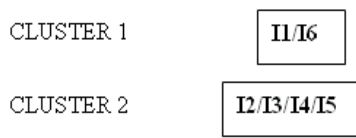


Table 3.4. Jaccard Coefficient Matrix: Single Linkage.

An exhaustive analysis of the proposed method in terms of false acceptance rate (FAR) vs classification accuracy was also carried out FVC datasets. The results and its comparison with its other counterparts are reported in Fig. 3.1.

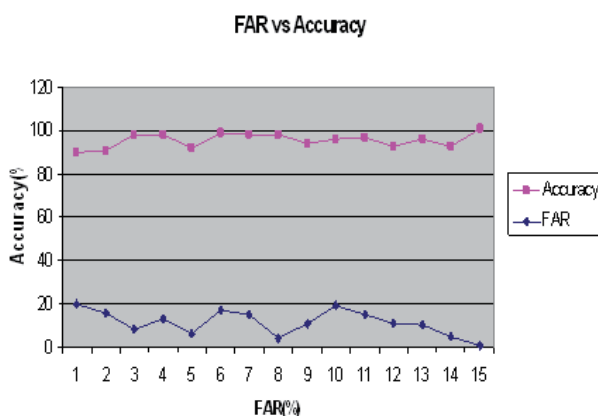


Fig. 3.1. FAR Vs Accuracy for fingerprint classification.

An algorithm performs well compared to another one has higher intra cluster similarity and lower inter cluster similarity compared to other algorithm. From the graphs (Fig 3.2 & 3.3) it is obvious that ROCK algorithm consistently performs better than other algorithms in both intra cluster and inter cluster similarity measure. For  $k=2$  and 3 bisecting K Means performs slightly better than ROCK algorithm in intra cluster similarity measure but overall ROCK algorithm outperforms other algorithms.

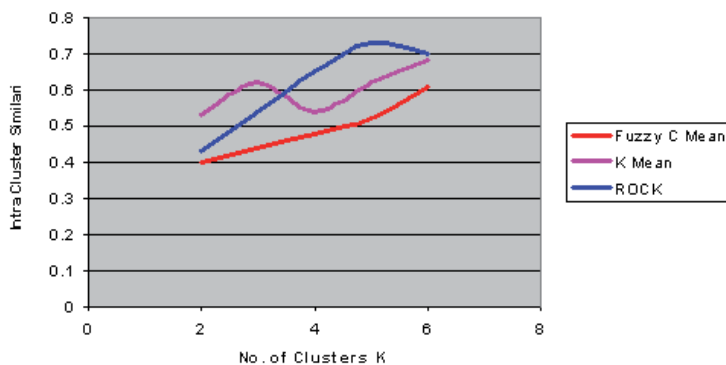


Fig 3.2. Intra Cluster Similarity for various algorithms on fingerprint database.

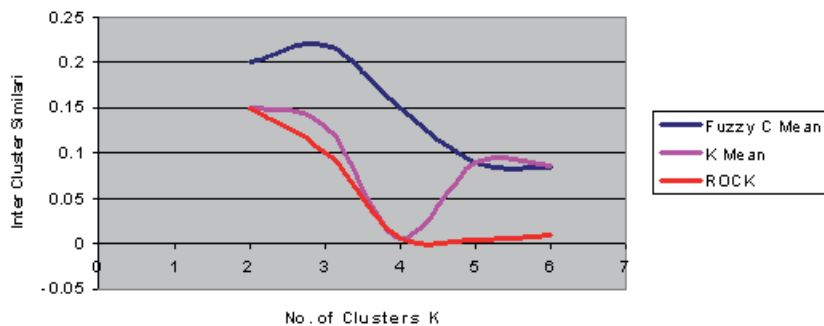


Fig 3.3. Inter Cluster Similarity for various algorithms on fingerprint database.

The FAR and FRR curve as claimed by the algorithm is shown in (Fig. 3.4). To evaluate the matching performance of the algorithm on database of FVC 2004, experiments have been conducted. The experiment considers all fingerprints in the database, leading to 95 matches and 5 non-matches. In this case, FAR and FRR values were 30-35% approximately and the equal error rate is 0.36.

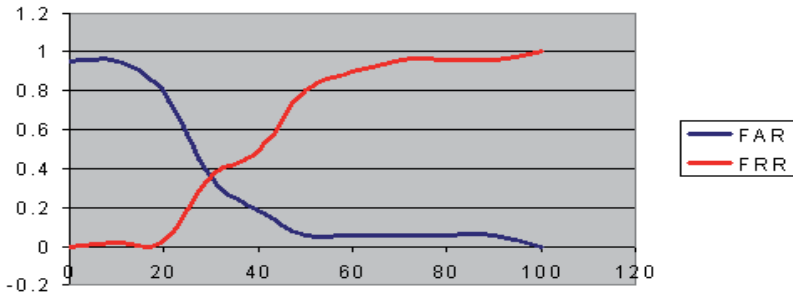


Fig. 3.4. FAR and FRR Curve.

## 4. Indexing and matching

### 4.1 Indexing scheme

It is expected that the query response time should depend upon the templates similar to the query template and not the total number of templates in the database. Thus the database should be logically partitioned such that images having similar patterns are indexed together. To search the large visual databases, content based image indexing and retrieval mechanism based on sub block of DCT coefficients are used. The scheme provides fast image retrievals. In this indexing technique feature vector which comprises of global and local features extracted from offline fingerprint databases are used by Robust clustering technique to partition the database. As biometric features possess no natural order of sorting, thus it is difficult to index them alphabetically or numerically. Hence, indexing is required to partition the search space. At the time of identification the fuzziness criterion is introduced to find the nearest clusters for declaring the identity of query sample. The system is tested using bin-miss rate and performs better in comparison to traditional k-means approach.

This method relies on the use of a small set of DCT coefficients for indexing a fingerprint image. By considering a reference block extracted using non-coherence factor, we only extract  $F_{0,0}, F_{0,1}, F_{1,0}, F_{2,0}, F_{1,1}, F_{0,2}, F_{0,3}, F_{1,2}, F_{3,0}$  for indexing and retrieval. All these coefficients are quantized finally.

The fingerprint image associated with an identity is subjected to the procedure above in order to generate the index code. The index code is stored in the fingerprint database along with the identity of the associated fingerprint. When a query image is presented to the system, initially image get categorized to which cluster belong to, in turn search with index code reference. During training phase, the images taken are grouped under various bins based on their index key. During testing phase, the index key is generated for the query image and the respective bin is located. The fingerprint template of the test image is

compared only with the fingerprint templates matched bin. The accuracy analysis of the recognition system with and without indexing (Exhaustive search) is presented in Table 4.1.

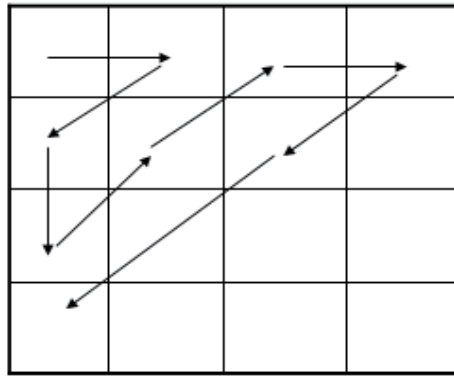


Fig. 4.1. Index Code.

| Method       | Without indexing | With indexing |
|--------------|------------------|---------------|
| Accuracy (%) | 91               | 97            |

Table 4.1. Accuracy of recognition system with and without indexing scheme.

Table 4.1 shows that the system with indexing provides 97% accuracy; whereas the system without indexing scheme provides only 91% of accuracy. This indexing mechanism considerably increases the recognition accuracy of the proposed feature extraction scheme. Also the number of comparisons made during the testing phase is reduced in the proposed recognition system with the help of indexing than the system using exhaustive search.

## 4.2 Greedy matching

The approach used is based on the general combination technique [22]. The function called greedy matching algorithm which integrates different matching criteria based on heterogeneous features. Different representations are often related to different features of the fingerprints in determining appropriate matching criteria. Generally, matching algorithm has to solve two problems: correspondence and similarity computation. For the correspondence problem, assign two descriptors: edge-based and ridge-based, and use an alignment-based greedy matching algorithm to establish the correspondences between blocks. From the similarity computation, a matching score is computed.

### 4.2.1 Alignment

Alignment is a crucial step for the proposed algorithm, as misalignment of two fingerprints of the same finger certainly produces a false matching result. In the proposed approach, two fingerprints are aligned using the top  $n$  most similar orientation pair. If none of the  $n$  pairs is correct, a misalignment occurs. The test is conducted using orientation-based, frequency-based and combined descriptors. It can be concluded that alignment based on combined descriptors is very reliable.

### 4.2.2 Block pairing

For matching, the fingerprint should be cropped to  $64 \times 64$  from the obtained reference block. With the registered orientation field with respect to ridge and edge, the procedure identifies the corresponding orientation block pairs is straightforward. Let denote the corresponding orientation block pair,  $p_k$  from test fingerprint, block  $q_k$  from template fingerprint respectively. The similarity degree  $S(p_k, q_k)$  of the two blocks is calculated as follows:

To compute the similarity between two sampling blocks, first compute the similarity of orientation and the similarity of frequency as follows:

$$s_o(p_k, q_k) = e^{-|\lambda(\alpha_k - \beta_k)|/(\pi/16)}$$

$$s_f(p_k, q_k) = e^{-|v_k - w_k|/3}$$

Then the similarity is defined as

$$s_t(p, q) = \omega \cdot s_o(p, q) + (1 - \omega) s_f(p, q) \quad (4.1)$$

where  $\omega$  is set to 0.5.

To compute the similarity between two sampling blocks, the similarity of non-coherence as follows:

$$s_m = \frac{m_p + 1}{M_p + 1} \cdot \frac{m_q + 1}{M_q + 1} \quad (4.2)$$

where  $m_p$  and  $m_q$  represent the number of matching factor of  $N(p)$  and  $N(q)$ , respectively, and  $M_p$  and  $M_q$  represent the total number of non-coherence of  $N(p)$  and  $N(q)$  that should be matching respectively. All terms plus 1 means that two central blocks  $p$  and  $q$  are regarded as matching. Here  $m_p$  and  $m_q$  are different because do not establish one-to-one correspondence. Since orientation-based descriptors and frequency-based descriptors capture contemporary information, further improve the discriminating ability of descriptors by combining two descriptors using the product rule:

$$s_c = s_t \cdot s_m \quad (4.3)$$

$$L = \{(S_c(p, q))\}_{p=1 \dots N1; j=1 \dots N2}$$

A list is used to store the normalized similarity degrees and indices of all block pairs. Elements in  $L$  are sorted in decreasing order with respect to  $S_c(p, q)$ . The first block pair  $(p1, q1)$  in  $L$  is used as the initial block pair, and two blocks are aligned using the initial pair. A pair of block is said to be matchable, if they are close in position and the difference of direction is small. The greedy matching algorithm is used to find match. Two arrays *flag1* and *flag2* are used to mark block that have been matched, in order that no block can be matched to more than one block. As the initial pair is crucial to the matching algorithm, the block pairs of the top  $Na$  elements in  $L$  are used as the initial pairs, and for each of them, a matching attempt is made. Totally  $Na$  attempts are performed,  $Na$  scores are computed and the highest one is used as the matching score between two fingerprints.

### 4.3 Matching process

**Input:** Orientation Value, Frequency Value, Non-coherence Factor ( $i_0, i_1, \dots, i_n$ ), ( $j_1, j_2, \dots, j_n$ )

**Output:** Matching Score (MS)

**function GreedyMatch**( $i_0, j_0$ )

**Step1:** Initialize  $flag1$  and  $flag2$  with 0;

**Step 2:**  $flag1[i_0] = 1;$

$flag2[j_0] = 1;$

**Step 3:** for  $m = 1$  to  $N1 \times N2$

$i = L(m).i; j = L(m).j;$

if ( $flag1[i]=0$ ) & ( $flag2[j]=0$ ) & ( $p_i$  and  $q_j$  are matchable)

**Step 4:** Insert ( $i, j$ ) into MP;

**Step 5:**  $flag1[i] = 1;$

$flag2[j] = 1;$

end if

end for

### 4.4 Conclusion

Fingerprint identification is one of the most well-known and publicized biometrics. Because of their uniqueness and consistency over time, fingerprints have been used for identification for over a century, more recently becoming automated due to advancements in computing capabilities. The fingerprint image is divided into sub-blocks and allow evaluating the statistical features from the DCT Coefficients. The feature space emphasizes meaningful global information by considering AC coefficients. The presented approach requires only a small number of parameters for global ridge pattern. The numerical data transformed into categorical data to produce the discrimination between the intra and inter class cluster with ROCK based algorithm. The system is performing comparatively superior as compared to K-Means and Fuzzy C Mean clustering technique for categorical attributes. The performance of the ROCK method for classification based on complete linkage is better than the other linkage measures. The matching results based on the FVC database using testing data sets for images with different orientations and translation shows the system outperforms better in recognition rate and requires short time for both the training and querying process.

### 5. References

- [1] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle. "An identity authentication system using fingerprints". Proc. IEEE, 85(9):1365–1388, 1997.
- [2] A. K. Jain, S. Prabhakar, and L. Hong, "A Multichannel Approach to Fingerprint Classification", IEEE Trans. Pattern Anal. and Machine Intell., Vol. 21, No. 4, pp. 348-359, 1999.
- [3] L. Hong, Y. Wan, and A. K. Jain. "Fingerprint image enhancement: Algorithms and performance evaluation". IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(8):777–789, 1998.
- [4] J. I.T. "Principle Component Analysis". Springer-Verlag, New York, 1986.



- [5] Ballan Meltem, Sakarya Ayhan, and Evans Brian, "A Fingerprint Classification Technique Using Directional Images," *Mathematical and Computational Applications*, pp. 91-97, 1998.
- [6] Cho Byoung-Ho, Kim Jeung-Seop, Bae Jae-Hyung, Bae In-Gu, and Yoo Kee-Young. "Fingerprint Image Classification by Core Analysis," *Proceedings of ICSP*, 2000.
- [7] Jain Anil, Prabhakar Salil, and Hong Lin, "A Multichannel Approach to Fingerprint Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 348-359, 1999.
- [8] Ji Luping, and Yi Zhang, "SVM-based Fingerprint Classification Using Orientation Field," *3<sup>rd</sup> International conference on Natural Computation*, vol. 2, pp. 724- 727, 2007.
- [9] Rajiv Mukherjee and Arun Ross, "Indexing iris database," *International Conference on Pattern Recognition*, Vol.54, pp. 45-50, 2008.
- [10] Puhan N B and Sudha, "A novel iris database indexing method using the iris color," *IEEE Conference on Industrial Electronics and Applications*, Vol.4, pp. 1886-1891, 2008.
- [11] Amit Mhatre, Sharat Chikkerur and Venu Govindaraju, "Indexing Biometric Databases Using Pyramid Technique," *International Conference on Audio-Video Biometric Person Authentication*, Vol. 354, pp.39-46, 2005.
- [12] Gupta P, Sana A, Mehrotra H and Jinshong Hwang C, "An efficient indexing scheme for binary feature based biometric database," *SPIE Biometric Technology for Human Identification*, Vol. 2, pp. 200-208, 2007.
- [13] Umarani Jayaraman, Surya Prakash and Phalguni Gupta, "An efficient technique for indexing multimodal biometric databases," *International Journal of Biometrics*, Vol.1, No.4, pp.418-441, 2009.
- [14] Chaur-Heh Hsieh, "DCT - Based codebook design for vector quantization of images," *IEEE Transaction on circuits and systems for video technology*, vol.2, pp. 456- 470, 1992.
- [15] Tienwei Tsai, Yo-Ping Huang and Te-Wei Chiang, "Dominant Feature Extraction in Block-DCT Domain", *IEEE Trans.*, 2006.
- [16] Suksan Jirachaweng and Vuctipong Areekul, "Fingerprint Enhancement Based on Discrete Cosine Transform", *Springer Berlin / Heidelberg*, Volume 4642/2007
- [17] Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim, "ROCK: A Robust Clustering Algorithm for Categorical Attributes" *Proc. Of the 15th Int'nl Conference on Data Engineering*, 2000.
- [18] S K Gupta, K.Sambasiva and Vasuda Bhatnagar, "K-means clustering algorithm for categorical attributes", *Springer-Verlag*
- [19] Kudov.P, Rezankova.H, Husek.D and Snasel.V, "Categorical Data Clustering Using Statistical Methods and Neural Networks"
- [20] George E. Tsekouras, Dimitris Papageorgiou, Sotiris Kotsiantis, Christos Kalloniatis, and Panagiotis Pintelas, "Fuzzy Clustering of Categorical Attributes and its Use in Analyzing Cultural Data", *World Academy of Science, Engineering and Technology* 2005.
- [21] Shyam Boriah, Varun Chandola and Vipin Kumar, "Similarity Measures for Categorical Data: A Comparative Evaluation", *In Proceedings of 2008 SIAM Data Mining Conference*, April 2008.

- [22] W. Sheng, G.Howell, M.C.Fairhur, F.Derav, K.Harmer, "Consensus fingerprint matching with genetically optimised approach", Pattern Recognition, Elsevier 2009.

## **Part 2**

### **Face Recognition**



# Facial Identification Based on Transform Domains for Images and Videos

Carlos M. Travieso-González, Marcos del Pozo-Baños and Jesús B. Alonso  
*University of Las Palmas de Gran Canaria, Signals and Communication Department,  
Institute for Technological Development and Innovation in Communications,  
Campus Universitario de Tafiera, Las Palmas de Gran Canaria,  
Spain*

## 1. Introduction

For the last decade, researchers of many fields have pursued the creation of systems capable of human abilities. One of the most admired humans' qualities is the vision sense, something that looks so easy to us, but it has not been fully understood yet. In the scientific literature, face recognition has been extensively studied and; in some cases, successfully simulated.

According to the Biometric International Group, nowadays Biometrics represent not only a main security application, but an expanding business according to Fig. 1a. Besides, as it can be seen in Fig. 1b, facial identification has been pointed out as one of the most important biometric modalities (Biometric International Group 2010).

However, face recognition is not an easy task. Systems can be trained to recognize subjects in a given case. But along the time, characteristics of the scenario (light, face perspective, quality) can change and mislead the system. In fact, the own subject's face varies along the time (glasses, hats, stubble). These are major problems with which face recognition systems have to deal using different techniques.

Since a face can appear in whatever position within a picture, the first step is to place it. However, this is far from the end of the problem, since within that location, a face can present a number of orientations. An approach to solve these problems is to normalize space position; variation of translation, and rotation degree; variation of rotation, by analyzing specific face reference points (Liun & He, 2008).

There are plenty of publications about gender classification, combining different techniques and models trying to increase the state of the art performance. For example, (Chenamma et al., 2010) presented the problem of face or person identification from heavily altered facial images and manipulated faces generated by face transformation software tools available online. They proposed SIFT features for efficient face identification. Their dataset consisted on 100 face images downloaded from <http://www.thesmokinggun.com/mugshots>, reaching an identification rate up to 92 %. In (Chen-Chung & Shiuian-You Chin, 2010), the RGB images are transformed into the YIQ domain. As a first step, (Chen-Chung & Shiuian-You Chin, 2010) took the Y component and applied wavelet transformation. Then, the binary two dimensional principal components (B2DPC) were extracted. Finally, SVM was used as classifier. On a database of 24 subjects, with 6 samples per user, (Chen-Chung & Shiuian-You Chin, 2010) achieved an average identification rate between 96.37% and 100%.

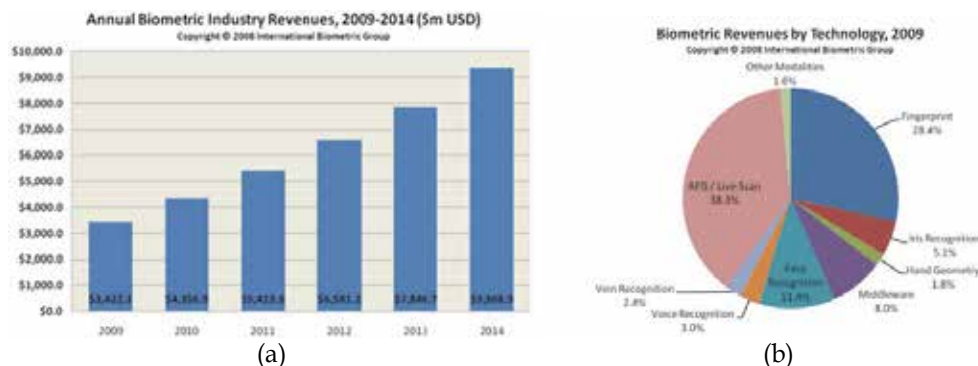


Fig. 1. Evolution on Biometric Market and Modalities according to International Biometric Group.

In (Zhou & Sadka, 2010) approach, while diffusion distance is computed over a pair of human face images, shape descriptions of these images were built using Gabor filters consisting of a number of scales and levels. (Zhou & Sadka, 2010) used the Sheffield Database, which is composed of 564 face images from 20 individual persons, mixed race/gender/appearance, along with the MIT-CBCL face recognition database that contains images of ten subjects and 200 images per user. They run experiments comparing the proposed approach against several competing methods and the proposed Gabor diffusion distance plus k-means classification (“GDD-KM”), reaching a success rate over 80%. (Bouchaffra, 2010) presented a machine learning paradigm that extends an HMM state-transition graph to account for local structures as well as their shapes. This projection of a discrete structure onto a Euclidean space is needed in several pattern recognition tasks. In order to compare the COHMMs approach proposed with the standard HMMs, they made a set of experiments using different wavelet filters for feature extraction with HMMs-based face identification. GeorgiaTech, Essex Faces95 and AT&T-ORL Databases were used. Besides, FERET Database was used on evaluation or test. Identification accuracies over 92.2% by CHMM-HMM, and 98.5% by DWT/COHMM were achieved.

In (Kisku et al., 2007), the database and query face images are matched by finding the corresponding feature points using two constraints to deal with false pair assignments and optimal feature sets. BANCA database was used. For this experiment, the Matched Controlled (MC) protocol was followed. Computing the weighted Error Rate prior EER on G1 and G2 for the two methods: gallery image based match constraint, and reduced point based match constraint, the error reached was between 8.52% and 4.29%. In (Akbari et al., 2010), a recognition algorithm based on feature vectors of Legendre moments was introduced as an attempt to solve the single image problem. A subset of 200 images from FERET database and 100 images from AR database were used in their experiments. The results achieved 91% and 89.5% accuracy for AR and FERET, respectively.

In (Khairul & Osamu, 2009) the implementation of moment invariants in an infrared-based face identification system was presented to develop a face identification system in thermal spectrum. A hierarchical minimum distance measurement method for classification was used. The performance of this system is encouraging with 87% of correct identification rate for test to registered image ratio of 2:1 and 84% of correct identification rate for test to registered image ratio of 4:1. Terravic facial IR database was used on this work. (Shreve et al., 2010) presented a method for face identification under adverse conditions by combining

regular, frontal face images with facial strain maps using score-level fusion. Strain maps were generated by calculating the central difference method of the optical flow field obtained from each subject's face during the open mouth expression. Extended Yale B database was used on this work, only the P00A+000E+00 image of each of the 38 subjects were used for the gallery and the remaining 2376 images were used as probes to test the accuracy of the identification system. Success rates achieved between 88% and 98% depending of the level of difficulty.

(Sang-Il et al., 2011) presented an approach that can simultaneously handle illumination and pose variations to enhance face recognition rate (Sang-Il et al., 2011). The proposed method consists of three parts which are pose estimation (projects a probe image into a low-dimensional subspace and geometrical distribution of facial components), shadow compensation (first they calculate the light direction) and face identification. This work was implemented under CMU-PIE and Yale B Databases, reaching a success rate between 99.5% and 99.9%. In (Yong et al., 2010) a three solution schemes for LPP (mathematical), using a k-nearest-neighbour as classifier was proposed and checked with ORL, AR, and FERET Databases. The success rates were 90%, 76% and 51%, respectively. (Chaari et al., 2009) developed a face identification system and used the reference algorithms of Eigenfaces and Fisherfaces in order to extract different features describing each identity. They built a database partitioning with clustering methods which split the gallery by bringing together identities with similar features and separating dissimilar features in different bins. Given a facial image that they want to establish its identity, they computed its partitioning feature that they compared to the centre of each bin. The searched identity is potentially in the nearest bin. However, by choosing identities that belong only to this partition, they increased the probability of an error discard of the searched identity. XM2VTS database was used reaching over 99% of success rate.

Techniques introduced in (Kurutach et al., 2010) were composed of two parts. The first one was the detection of facial features by using the concepts of Trace Transform and Fourier transform. Then, in the second part, the Hausdorff distance was employed to measure and determine of similarity between the models and tested images. Finally, their method was evaluated with experiments on the AR, ORL, Yale and XM2VTS face databases. The average of accuracy rate of face recognition was higher than 88%. In (Wen-Sheng et al, 2011), as each image set was represented by a kernel subspace, they formulate a KDT matrix that maximizes the similarities of within-kernel subspaces, and simultaneously minimizes those of between kernel subspace. Yale Face database B, Labeled Faces in the Wild and a self-compiled database. Success rates were achieved between 98.7% and 97.9%.

In (Zana & Cesar, 2006), polar frequencies descriptors were extracted from face images by Fourier-Bessel transform (FBT). Next, the Euclidean distance between all images was computed and each image was then represented by its dissimilarity to the other images. A pseudo-Fisher linear discriminant was built on this dissimilarity space. The performance of discrete Fourier transform (DFT) descriptors and a combination of both feature types was also evaluated. The algorithms were tested on a 40- and 1196-subjects face database (ORL and FERET, respectively). With five images per subject in the training and test datasets, error rate on the ORL database was 3.8, 1.25, and 0.2% for the FBT, DFT, and the combined classifier, respectively, as compared to 2.6% achieved by the best previous algorithm. In (Zhao et al., 2009), feature extraction was carried out on face images respectively through conventional methods of wavelet transform, Fourier transform, DCT, etc. Then, these image transform methods are combined to process the face images. Nearest-neighbour classifiers

using Euclidean distance and correlation coefficients used as similarity are adopted to recognize transformed face images. By this method, five face images from a face database (ORL database) were selected as training samples, and the rest as testing samples. Success recognition rate was 97%. When five face images were from Yale face database, the correct recognition rate was up to 94.5%.

The methodology proposed in (Azam et al., 2010) is a hybrid approach to face recognition. DCT was applied to hexagonally converted images for dimensionality reduction and feature extraction. These features were stored in a database for recognition purpose. Artificial Neural Network (ANN) was used for recognition. Experiments and testing were conducted over ORL, Yale and FERET databases. Recognition rates were 92.77% (Yale), 83.31% (FERET), and 98.01% (ORL). In (Chaudhari & Kale, 2010) the process was divided in two steps: 1) Detect the position of pupils in the face image using geometric relation between the face and the eyes and normalizes the orientation of the face image. Normalized and non normalized face images are given to holistic face recognition approach. 2) Select features manually. Then determine the distance between these features in the face image and apply graph isomorphism rule for face recognition. Then apply a Gabor filter on the selected features. This Algorithm takes into account Gabor coefficient as well as Euclidean distance between features for face recognition. Brightness normalized and non normalized face images were given to feature based approach face recognition methods. ORL database was used, reaching over 99.5% for the best model. (Chi Ho Chan & Kittler, 2010) combined sparse representation with a multi-resolution histogram face descriptor to create a powerful representation method for face recognition. The recognition was performed using the nearest neighbour classifier. Yale Face Database B and the extended Yale Face Database B were used, achieving a success rate up to 99.78%.

In this chapter, we present a model to identify subjects from TV video sequences and images from different public databases. This model basically works with three steps: detect faces within the frames, undergo face recognition for each extracted face, and for videos, use information redundancy to increase the recognition rate. The main goal is the use of transform domains in order to reach good results for facial identification. Therefore, this chapter aims to show that our best proposal, applied on face images, can be extended to be used on videos, in particular on TV videos, developing for this purpose a real application. Results presented on the experiment sections show the robustness of our proposal on illumination, size and lighting variations of facial images.. Finally, we have used the inter-frame information in order to improve our approach on its use for video mode. Therefore, the whole system presents a good innovation.

The first experimental setting is based on images from ORL and Yale databases; in order to determinate the best transform domains under illumination changes and without it. Besides, those databases have been used widely; therefore, we show our good results vs other approaches. We have applied different Transform Domains Feature Extractions, as Discriminative Common Vectors (DCV), Discrete Wavelet Transform (DWT), Independent Component Analysis (ICA), Discrete Cosine Transform (DCT), Linear Discriminant Analysis (LDA), and Principal Components Analysis (PCA). For the supervised classification, we have used Support Vector Machine (SVM), Neural Network (NN) and Euclidean Distance methods. Our proposal adjusts both parameterization and classification steps, and our best approach is finally applied to our TV video database (V-DDBB), composed of 40 videos.



## 2. Facial pre-processing

The first block is the pre-processing block. It gets the faces samples ready for the forthcoming blocks, reducing the noise and even transforming the original signal in a more readable one. An important property of this block is that it tries to reduce lightning variations among pictures. In this case, samples are first resized to standard dimensions of 20x20 pixels. This ensures that the training time will not reach unviable levels. Then images' grey scale histograms are equalized. Fig. 2 female face (left) shows the effect of apply this process to a given sample.



Fig. 2. Example of Pre-processed image.

Finally, for a new set of experiments, a local normalization function is added at the end of the block (Xiong, 2005). This function is based on a double Gaussian filtering, and makes the local mean and variance uniform along the picture. The effect of this new tool is dramatic and it can be seen in Fig. 2 male face (right).

## 3. Feature extraction

Once the samples are ready, the feature extractor block transforms them in order to obtain the best suited information for the classification step. Six types of feature extraction techniques are shown in this section.

### 3.1 Principal Component Analysis (PCA)

Principal components analysis (PCA) is a technique used to reduce multidimensional data sets to lower dimensions for analysis (Banu & Nagaveni, 2009). The applications include exploratory data analysis and generating predictive models. PCA involves the computation of the eigenvalue decomposition or singular value decomposition of a data set, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of scores and loadings. This process applied to face recognition is named blind source separation, where there are fewer sources than input channels.

The blind source separation consists in several sources that are mixed in a system, these mixtures are recorded together and they have to be separated to obtain the estimations of the original sources. The following figure shows the mixing system;

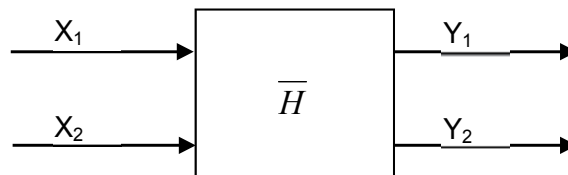


Fig. 3. Two Sources in a two mixtures system.

Generally, there are  $n$  source signals statistically independent  $s(t) = [s_1(t), \dots, s_n(t)]$ , and  $m$  observed mixtures that are linear and instantaneous combinations of the previous signals  $x(t) = [x_1(t), \dots, x_m(t)]$ . Beginning with the linear case, the simplest one, we have that the mixtures are:

$$x_i(t) = \sum_{j=1}^n h_{ij} \cdot s_j(t) \quad (1)$$

Now, we need to recover  $s(t)$  from  $x(t)$ . It is necessary to estimate the inverse matrix of  $H$ , where  $h_{ij}$  are contained. Once we have this matrix:

$$y(t) = \overline{W} \cdot x(t) \quad (2)$$

where  $y(t)$  contains the estimations of the original source signals, and is the inverse mixing matrix. Now we have defined the simplest case, it is time to explain the general case that involves convolute mixtures. The process is defined as follows:

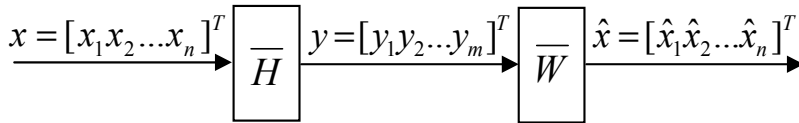


Fig. 4. BSS General Problem.

where  $\overline{H}$  is the mixed system;

$$\overline{H} = \begin{bmatrix} h_{11} & \dots & h_{1n} \\ \dots & \dots & \dots \\ h_{m1} & \dots & h_{mn} \end{bmatrix} \quad (3)$$

The  $h_{ij}$  are FIR filters, each one represents a signal transference multi-path function from source,  $i$ , to sensor,  $j$ .  $i$  and  $j$  represent the number of sources and sensors.

### 3.2 Discrete Wavelet Transform (DWT)

The Wavelet transform is another preprocessing and feature extraction technique which can be directly applied to face images. The Discrete Wavelet Transform (DWT) (González & Woods, 2002) is defined as follows:

$$C[j, k] = \sum_{n \in \mathbb{Z}} f[n] \psi_{j,k}[n] \quad (4)$$

where  $\psi_{j,k}$  is the transform function:

$$\psi_{j,k}[n] = 2^{\frac{-j}{2}} \cdot \psi[2^{-j}n - k] \quad (5)$$

The application of different mother families on pre-processing (artefacts elimination) and on the feature extraction has got a set of good and discriminate parameters.

### 3.3 Linear Discriminate Analysis (LDA)

The objective of LDA is the reduce of sample dimensionality while preserving all the information among possible classes (Huang et al., 2003). As opposed to the components analysis, the discriminate analysis seeks chiefly a projection that separates the data in the way of the quadratic error. Therefore in LDA, a different point of view is taken with respect to PCA.

LDA is a popular mapping technique when the labels of the different classes are known. This is an advantage since it can make use of this previous information, which gives a description of the disposition of the data.

It is obtained the "with dispersion" and "without dispersion" matrixes between classes, maximizing the first one and diminishing the second. This is carried out with the measure of the ratio among the projection of the "with dispersion" matrix determinant between classes and the projection of the "without dispersion" matrix determinant, which does that the projection among separation of classes be maximum.

The "without dispersion" matrix between classes is defined as:

$$S_w = \sum_{i=1}^c S_i \quad (6)$$

where  $c$  is the number of classes and:

$$S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^T \quad (7)$$

for each set of  $D_i$ , where:

$$m_i = \frac{1}{n_i} \sum_{x \in D_i} x \quad (8)$$

On the other hand, the "with dispersion" matrix between classes is defined as:

$$S_b = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T \quad (9)$$

where  $n_i$  is the number of samples for each class  $i$  and  $m$  is the vector of total mean.

As in the case of PCA, where it is not necessary to separate the classes, the total dispersion matrix is defined as:

$$S_t = \sum_{x \in D} (x - m)(x - m)^T \quad (10)$$

where  $D$  is the total matrix of data.

Then, in  $S_t = S_w + S_b$  from (6) and (9), it is searched a projection, which satisfies:

$$\begin{aligned} W &= \arg \max_w \frac{|W^{*T} S_b W^*|}{|W^{*T} S_w W^*|} \\ &= [w_1 \quad w_2 \quad \cdots \quad w_d] \end{aligned} \quad (11)$$

where  $W^*$  are the auto-values of  $S_w^{-1}S_b$ .

Considering  $S_w$  as nonsingular, the base-vectors of  $W$  correspond to the  $M$  greater eigenvalues of  $S_w^{-1}S_b=W^*$ . The representation of  $M$ -dimensional space is obtained with the projection of the original data on the sub-space  $W$ , with  $M$  eigen-vectors; where  $M$  is lower than  $n$ .

This technique provides a tool for the classification that permits to diminish in a considerable way the calculation of the characteristics of the different samples. Moreover, it conserves the information among each class, what provides a greater level of discrimination among the samples.

### 3.4 Discrete Cosine Transform (DCT)

We have applied Discrete Cosine Transform (DCT) for noise elimination and details of high frequency (González & Woods, 2002). Besides, this transform has a good energy compaction property that produces uncorrelated coefficients, where the base vectors of the DCT depend only on the order of the transformation selected, and not of the statistical properties of the input data.

Another important aspect of the DCT is its capacity to quantify the coefficients utilizing quantification values, which are chosen of visual way. This transformation has had a great acceptance inside the image digital processing, because there is a high correlation among elements for the data of a conventional image.

### 3.5 Independent Component Analysis (ICA)

The main objective of the blind source separation (BSS) is to obtain, from a number of observations, the different signals that compose these observations. This objective can be reached using either a spatial or a statistical approach. The former is based on a microphone array and depends on the position and separation of them. It also uses the directions of arrival (DOA) from the different audio signals.

On the other hand, the statistical separation supposes that the signals are statistically independent, that they are mixed in a linear way and that it is possible to get the mixtures with the right sensors (Hyvärinen et al., 2001) (Parra, 2002). This technique is the newest and it is in a continuous development. It is used in different fields such as real life applications (Saruwatari et al., 2003) (Saruwatari et al., 2001), natural language processing (Murata et al., 2001), bioinformatics, image processing (Cichocki & Amari, 2001), etc.

Two main types of BSS problem can be differentiate: with linear and the nonlinear signal mixture. The former presents linear mixtures where the data is mixed without echoes or reverberations, while the mixtures of the latter are convolutive and they are not totally independent due to the propagation of the signal through dynamic environments. This more complex case is known as the "Cocktail party problem".

Depending on the mixtures, there are several methods to solve the BSS problem. The first case can be seen as a simplification of the second one.

In this work the statistical approach named Independent Component Analysis (ICA) is studied. ICA comes from the previously introduced PCA (Hyvärinen et al., 2001) (Smith, 2006). The BBS based on ICA is also divided into three groups; the first one are those methods that works in the time domain, the second are those who works in the frequency domain and the last group are those methods that combine frequency and time domain methods.

### 3.6 Discriminative Common Vectors (DCV)

The DCV approach was first introduced as a solution for the small sample size problem. This tool performs a double transformation on the training set. One transformation relies in the within-class properties, while the other plays in role in the inter-class domain.

In the first step, the DCV approach extracts the common properties of each class, making within-class samples more alike (Cevikalp et al., 2005) (Travieso et al., 2009). Let  $C$  denotes the number of classes,  $N_C$  the number of samples of each class, and  $x_n^c$  a  $d$ -dimensional column vector that represents the  $n$ th samples of  $c$ th class. Assume that the small sample size problem is present, thus  $d > M - C$ . In this case, the within-class  $S_W$ , between-class  $S_B$ , and total  $S_T$  projection matrixes can be defined as:

$$S_W = \sum_{c=1}^C \sum_{n=1}^{N_C} (x_n^c - \mu_c)(x_n^c - \mu_c)^T \quad (12)$$

$$S_B = \sum_{c=1}^C (\mu_c - \mu)(\mu_c - \mu)^T \quad (13)$$

$$S_T = \sum_{c=1}^C \sum_{n=1}^{N_C} (x_n^c - \mu)(x_n^c - \mu)^T \quad (14)$$

where  $\mu$  is the mean of all samples, and  $\mu_c$  is the mean of samples in the  $i$ th class. In order to obtain a common vector for each class, it is interesting to know that every sample can be decomposed as follow:

$$x_n^c = y_n^c + z_n^c \quad (15)$$

where  $y_n^c$  denotes the rank of  $S_W$  and  $z_n^c$  the null space of  $S_W$ . Therefore, the common vectors will be obtained in the null space as:

$$z_n^c = x_n^c - y_n^c \quad (16)$$

and the rank space can be obtained using the eigenvectors corresponding to the no null eigenvalues of  $S_W$  as a projecting matrix  $Q$ .

$$z_n^c = x_n^c - QQ^T x_n^c \quad (17)$$

It can be proven [8] that this formula drives to one unique vector for each class, the common vector.

$$x_{com}^c = z_n^c, \quad \forall n \quad (18)$$

And the within-class dispersion matrix of these new common vectors is null:

$$S_W^{com} = 0 \quad (19)$$

In the second transformation, the differences between classes are magnified in order to increase the distance between common vectors, and therefore increase the inter-class dispersion. To achieve this, it is necessary to project into the rank of  $S_W^{com}$ , or similarly into

the rank space of  $S_{W^{com}}$  (remember equations (14) and (19)). This can be accomplished by an eigenanalysis of  $S_{W^{com}}$ . Using eigenvectors corresponding to no null eigenvalues to form a new projecting matrix  $W$ , and apply it to the common vectors:

$$\beta^C = W^T x_{com}^c \quad (20)$$

Because all common vectors are the same for each class, the  $S_T^{com}$  function is calculated using only one common vector for each class.

This approach maximizes the DCV criterion:

$$J_{DCV}(W_{op}) = \max_{|W^T S_T W| \neq 0} |W^T S_T W| \quad (21)$$

#### 4. Classification system

For this work, we have studied three supervised classification systems, based on Euclidean Distance (lineal classification), Neural Network (Bishop, 1991), and Support Vector Machines (Hamel, 2009) (Cristianini & Shawe-Taylor, 2000). With this set of classifiers, we can observe the behaviour of parameterization techniques versus the illumination conditions of a facial image.

The first classifier is a linear classifier; in particular, it is a Euclidean Distance. Its expression is,

$$\|x - y\|_e = \left[ \sum_{i=1}^D |x_i - y_i|^2 \right]^{1/2} \quad (22)$$

Another classifier is the Neural Network, in particular the perceptron score. The perceptron of a simple layer establishes its correspondence between classes with a lineal discrimination rule. However, it is possible to define discriminations for not lineally separable classes utilizing multilayer perceptrons, which are networks without refreshing (feed-forward) with one or more layers of nodes between the input layer and exit layer. These additional layers contain hidden neurons or nodes, which are directly connected to the input and output layer.

A neural network multilayer perceptron (NN-MLP) of three layers is shown in the figure 5 (Bishop, 1991), with one hidden layer. Each neuron is associated with a weight and a bias, these weights and biases of the connections of the network will be trained to make their values suitable for the classification between the different classes.

Particularly, the neural network used in the experiments is a Multilayer Perceptron (MLP) Feed-Forward with Back-Propagation training algorithm, and with only one hidden layer.

The third classifier is a SVM light. SVM light is an implementation of Vapnik's Support Vector Machine (Hamel, 2009) (Cristianini & Shawe-Taylor, 2000) for the problems of pattern recognition, regression, and learning a ranking function. The optimization algorithms used in SVM light are described in (Hamel, 2009) (Cristianini & Shawe-Taylor, 2000). The algorithm has scalable memory requirements and can handle problems with many thousands of support vectors efficiently. For this reason is very interesting for this applications, because we are going through thousands of parameters. The program is free for scientific use (Joachims, 1999) (svmlight, 2007).

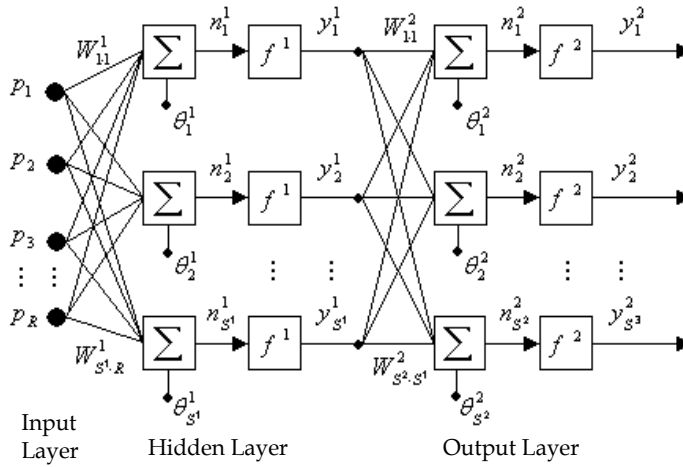


Fig. 5. Multilayer Perceptron.

In Fig. 6, we can see the detection of support vectors and the creation of boundary, one per each class, because this is a bi-class classifier. In our implementation, we have built a multi-classes classification module, from this SVM light.

## 5. Experimental settings

Our experiments are done on images and videos. The first step has been to study the effect of Transform Domain on facial images, and the best results to study on facial videos. In this work, six feature extractions and three classifiers have been used. All the experiments were done under automatic supervised classification.

### 5.1 Image and video face databases

Two data databases have been used for the first experiments, allowing to study the effects of illumination conditions on different techniques. These databases are ORL and Yale databases (AT&T Laboratories Cambridge, 2002) (Face Recognition Homepage, 2011). ORL is composed by 40 faces with 10 samples by face, having a total of 400 samples. The images are in grey scale (8 bits), and all the images have a dimension of  $92 \times 112$  pixels. This database is independent of illumination conditions, because its light focus is fixed. Yale is composed by 15 faces with 11 samples by face, in total, it has 165 images. Again, the database is in grey scale with 8 bits, and the size of the images is  $243 \times 320$  pixels. Each sample of this database has a different focus of illuminations. Samples from both databases are showed in Fig. 7.

We have also acquired a video database (V-DDBB) composed by 40 videos; 10 for each 4 subjects. This database is divided in two sets regarding the quality of the picture. Videos from the lower resolution V-DDBB present 32 frames per second with dimensions  $208 \times 160$  and variable data speed between 39 and 46 kbps, while videos from the higher resolution V-DDBB present 29 frames per second with dimensions  $320 \times 240$  and a constant speed rate of 2000 kbps. Information about length, size, and used can be find in tables 1, 2, 3, and 4, where 'Tr/Test' means that the use of the video can be either training or test depending on the

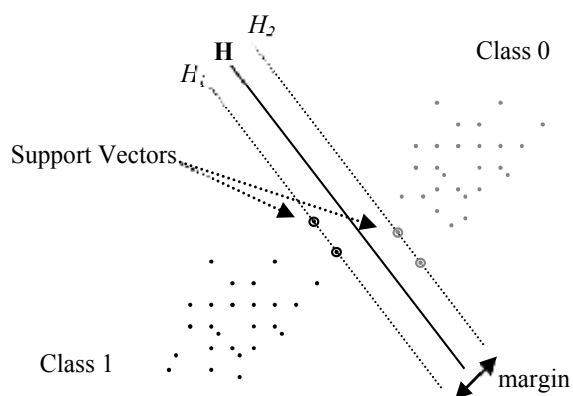


Fig. 6. Separate lineal Hyperplane in SVM.



Fig. 7. Samples from both databases. (a) Yale Database. (b) ORL Database.

experiment. In any case, it is important to clarify that videos used for tests were not used for training the system.

Videos were originally captured from regular analogical TV emissions. However, the capture process was not the same for every video. Even though the aspect ratio range of AVI files remains almost constant, picture's resolution goes from a fuzzy image to an acceptable one. Moreover, lightning and background are very changing aspects from both inter-subject and intra-subject V-DDBB since each video came from a different TV program.

| SUBJECT 1                  |        |         |          |                             |        |          |          |
|----------------------------|--------|---------|----------|-----------------------------|--------|----------|----------|
| Lower Definition Data Base |        |         |          | Higher Definition Data Base |        |          |          |
| Video                      | Length | Size    | Used for | Video                       | Length | Size     | Used for |
| 1                          | 02:35  | 5.75 MB | Training | 1                           | 01:17  | 12,40 MB | Training |
| 2                          | 01:01  | 2,51 MB | Training | 2                           | 01:41  | 14,00 MB | Training |
| 3                          | 01:21  | 3,81 MB | Training | 3                           | 01:03  | 4,80 MB  | Training |
| 4                          | 01:43  | 4,56 MB | Tr/Test  | 4                           | 02:45  | 28,80 MB | Tr/Test  |
| 5                          | 01:40  | 4,43 MB | Test     | 5                           | 01:21  | 12,40 MB | Test     |

Table 1. V-DDBB information for subject 1.



| SUBJECT 2                  |        |         |          |                             |        |         |          |
|----------------------------|--------|---------|----------|-----------------------------|--------|---------|----------|
| Lower Definition Data Base |        |         |          | Higher Definition Data Base |        |         |          |
| Video                      | Length | Size    | Used for | Video                       | Length | Size    | Used for |
| 1                          | 01:35  | 4,09 MB | Training | 1                           | 01:03  | 5,51 MB | Training |
| 2                          | 01:42  | 4,06 MB | Training | 2                           | 00:58  | 4,45 MB | Training |
| 3                          | 02:00  | 4,70 MB | Training | 3                           | 01:06  | 4,49 MB | Training |
| 4                          | 02:07  | 4,95 MB | Tr/Test  | 4                           | 01:40  | 12,0 MB | Tr/Test  |
| 5                          | 01:27  | 3,47 MB | Test     | 5                           | 01:40  | 14,0 MB | Test     |

Table 2. V-DDBB information for subject 2.

| SUBJECT 3                  |        |         |          |                             |        |         |          |
|----------------------------|--------|---------|----------|-----------------------------|--------|---------|----------|
| Lower Definition Data Base |        |         |          | Higher Definition Data Base |        |         |          |
| Video                      | Length | Size    | Used for | Video                       | Length | Size    | Used for |
| 1                          | 01:33  | 3,87 MB | Training | 1                           | 02:42  | 14,5 MB | Training |
| 2                          | 03:28  | 10,2 MB | Training | 2                           | 02:30  | 16,0 MB | Training |
| 3                          | 02:10  | 5,53 MB | Training | 3                           | 10:03  | 40,3 MB | Training |
| 4                          | 02:39  | 7,26 MB | Tr/Test  | 4                           | 01:24  | 10,5 MB | Tr/Test  |
| 5                          | 03:01  | 9,27 MB | Test     | 5                           | 00:46  | 2,38 MB | Test     |

Table 3. V-DDBB information for subject 3.

| SUBJECT 4                  |        |         |          |                             |        |         |          |
|----------------------------|--------|---------|----------|-----------------------------|--------|---------|----------|
| Lower Definition Data Base |        |         |          | Higher Definition Data Base |        |         |          |
| Video                      | Length | Size    | Used for | Video                       | Length | Size    | Used for |
| 1                          | 01:54  | 4,59 MB | Training | 1                           | 01:37  | 10,3 MB | Training |
| 2                          | 01:01  | 2,58 MB | Training | 2                           | 00:59  | 6,52 MB | Training |
| 3                          | 02:10  | 5,38 MB | Training | 3                           | 02:26  | 12,1 MB | Training |
| 4                          | 02:20  | 5,46 MB | Tr/Test  | 4                           | 00:20  | 2,65 MB | Tr/Test  |
| 5                          | 01:00  | 2,58 MB | Test     | 5                           | 00:56  | 5,05 MB | Test     |

Table 4. V-DDBB information for subject 4.

Focus was a major issue too as subjects were not always presented in the same way. Thus our system has to be tuned to recognize characters only from close up pictures. On Fig. 8 different examples with lower and higher definition are shown.

For the video database, OpenCV library were used to build up the Face Extractor Block (FEB), a library of programming functions mainly aimed at real time computer vision. For face detection, it uses Haar-like features to encode the contrasts exhibited by a human face and their spacial relationships. Basically, a classifier is first trained and subsequently applied to a region of interest (Bradski, 2011).



Fig. 8. (a) Pictures from two different subjects of the higher quality Video Database. (b) Pictures from two different subjects of the lower quality Video Database.

The FEB receives an AVI XVID MPEG-4 video file and a sample rate parameter (SR) as inputs. It checks frames for faces every SR seconds. If any face is found, the FEB extracts and saves it as a JPEG picture. The number of the studied frame sequence is saved for the future time analysis. In order to delimit the quality of our future face database we imposed an aspect ratio of the extracted face and a minimum face size.

## 5.2 Experiments on facial images

The goal of this present work is to study and search a good identification system for different illumination conditions. Therefore, we have tested all parameterization tools with different classification techniques. In particular, our classification system has been tested with three different methods, Euclidean Distance (lineal classification), Neural Network (Bishop, 1991), and Support Vector Machines (Hamel, 2009) (Cristianini & Shawe-Taylor, 2000).

These classification systems have been used with supervised classification; and therefore, we have developed two modes in this process. The first mode is the training mode, where the system is trained with 50% of the database in use, while the remainder is used during the test mode. Our experiments have been repeated 10 times, and therefore, our results are showed as mean and variance.

Table 5 shows the results achieved for each classifier with different parameterization systems. It can be observed that the best system is that using SVM, for all parameterizations and both databases. Therefore, the best classifier with independency of illumination conditions is a SVM, based on Radial Basis Function (RBF) kernel. However, the same cannot be said for the parameterization technique, as there is not one dominant tool along different scenarios (see table 5). Besides, from this table 5 we can observe that in general the obtained results are quite robust.

About the computational time, we have studied the mean value with the best classifier. With the train mode, it has been obtained 30 seconds and with test mode, 0.5 seconds. Those computational times have been reached with MATLAB language (Matlab, 2011), which is an interpretative programming language. On future works, we expect to decrease the computational times by 5 to 7 times by migrating to C++ programming language.

| Type of parameters | ORL Database    |                  |                  | Yale Database    |                  |                 |
|--------------------|-----------------|------------------|------------------|------------------|------------------|-----------------|
|                    | EDC             | NN               | SVM              | EDC              | NN               | SVM             |
| DWT - Haar         | 95,30%<br>±2,84 | 86,50%<br>±14,95 | 97,20%<br>±0,96  | 97,73%<br>±3,58  | < 50%            | 98,67%<br>±1,58 |
| DWT - Bior         | 96,80%<br>±1,23 | < 50%            | 98,90%<br>±0,79  | 98,27%<br>±9,17  | < 50%            | 99,07%<br>±1,60 |
| LDA                | 95,00%<br>±6,83 | 90,90%±5,10      | 96,10%<br>±3,16  | 98,80%<br>±1,36  | 94,13%<br>±17,55 | 99,20%<br>±0,87 |
| PCA                | 95,90%<br>±3,04 | 86,40%±11,82     | 96,15%<br>±2,89  | 98,13%<br>±4,43  | 89,60%<br>±10,59 | 98,80%<br>±2,15 |
| DCT                | 94,90%<br>±3,04 | 95,55%<br>±1,69  | 96,96%<br>±1,19  | 97,87%<br>±3,64  | 78,67%<br>±19,40 | 99,07%<br>±1,60 |
| ICA                | 87,85%<br>±6,95 | < 50%            | 92,50%<br>±4,94  | 97,60%<br>±12,82 | 70,67%<br>±19,27 | 99,33%<br>±0,89 |
| DCV                | 95,47%<br>±1,23 | < 50%            | 98,21% ±<br>1,67 | 95,34%<br>±4,93  | 88,25%<br>±7,32  | 98,12%<br>±2,93 |

Table 5. Success Rates for ORL and Yale Databases.

From the Table 5, it is observed that DWT gives good success for image without illumination changes using SVM (ORL Database). For data with illumination changes (Yale Database), we can see ICA gives the better modelling. Therefore, those two techniques will be used on the next section.

### 5.3 Experiments on facial videos

As it is specified in tables 1, 2, 3, and 4, different videos were used for either training or test. A number of models with different setups were created in order to find the best configuration. Results for each experiment are presented in the following tables. Some of these videos present public events, therefore doing the pool of subjects immeasurable. Both for ICA and DWT experiments different set ups were used, varying the training / test ratio. Moreover, for ICA experiments (see Table 6) the number of principal components used were also a parameter of configuration; ranging from 5 to 40. Higher classification rates of 76,87 % and 83,12 % were obtained for lower and higher resolution databases respectively.

For DWT experiments (see Table 7) the number of iterations ranged from 1 to 3. On the other hand, both 'bior 4.4' and 'haar' filters were used. In this case, higher classification rates of 86,25% and 98,75 % were obtained for lower and higher DDBBs respectively.

In order to further understand the results, a few more experiments were done reducing the registered subjects from 4 to 3 and 2. Here, it was shown that the identification rate increases while the number of subjects decreases. However, an exception occurs for DWT experiments and higher quality DDBB. This exception appears due to a lower performance of the system when particular subjects were modelled to be detected. In other words, the system appeared to have problems differencing between two specific subjects. When these two subjects were presented in the experiment, the system performs extremely badly, and therefore decreasing the mean identification rate.

| ICA                  |                               |                               |                               |                               |
|----------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Principal Components | Lower Resolution              |                               | Higher Resolution             |                               |
|                      | 80 for training / 20 for test | 60 for training / 40 for test | 80 for training / 20 for test | 60 for training / 40 for test |
| 5                    | 5 % ± 0                       | 30,62 % ± 0                   | 41 % ± 4,76                   | 58,12 % ± 0                   |
| 10                   | 47,08 % ± 4                   | 51,25 % ± 0                   | 33,33 % ± 3,14                | 62,29 % ± 0.36                |
| 20                   | 58,75 % ± 0                   | 46,25 % ± 0                   | 25,83 % ± 5,05                | 64,37 % ± 0                   |
| 30                   | 57,5 % ± 0                    | 64,37 % ± 0                   | 17,5 % ± 0                    | 83,12 % ± 0                   |
| 40                   | 58,75 % ± 0                   | 76,87 % ± 0                   | 15 % ± 0                      | 79,37 % ± 0                   |
| Subjects             | Using the best combination    |                               | Using the best combination    |                               |
| 3 subjects           | 77,08 % ± 4,34                |                               | 84,56 % ± 1,99                |                               |
| 2 subjects           | 80 % ± 15,41                  |                               | 89,68 % ± 13,2                |                               |

Table 6. Identification results for ICA experiments and lower and higher Resolution V-DDBB.

| DWT                                 |                               |                               |                               |                               |
|-------------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Using 'bior 4.4'                    |                               |                               |                               |                               |
| Principal Iterations / Coefficients | Lower Resolution              |                               | Higher Resolution             |                               |
|                                     | Lower Resolution              |                               | Higher Resolution             |                               |
|                                     | 80 for training / 20 for test | 60 for training / 40 for test | 80 for training / 20 for test | 60 for training / 40 for test |
| 1 / 1764                            | 85 % ± 0                      | 75,62 % ± 0                   | 97,5 % ± 0                    | 76,25 % ± 0                   |
| 2 / 638                             | 76,25 % ± 0                   | 66,25 % ± 0                   | 98,75 % ± 0                   | 75,62 % ± 0                   |
| 3 / 285                             | 45 % ± 0                      | 40,625 % ± 0                  | 93,75 % ± 0                   | 71,87 % ± 0                   |
| Using 'haar'                        |                               |                               |                               |                               |
| Iterations / Coefficients           | Lower Resolution              |                               | Higher Resolution             |                               |
|                                     | Lower Resolution              |                               | Higher Resolution             |                               |
|                                     | 80 for training / 20 for test | 60 for training / 40 for test | 80 for training / 20 for test | 60 for training / 40 for test |
| 1 / 1440                            | 86,25 % ± 0                   | 81,25 % ± 0                   | 93,75 % ± 0                   | 74,37 % ± 0                   |
| 2 / 368                             | 86,25 % ± 0                   | 80,62 % ± 0                   | 90 % ± 0                      | 76,25 % ± 0                   |
| 3 / 96                              | 81,25 % ± 0                   | 76,25 % ± 0                   | 86,25 % ± 0                   | 74,37 % ± 0                   |
| Subjects                            | Using the best combination    |                               | Using the best combination    |                               |
| 3 subjects                          | 87,08 % ± 10,3                |                               | 95,41 % ± 5,16                |                               |
| 2 subjects                          | 93,12 % ± 6,88                |                               | 98,12 % ± 3,75                |                               |

Table 7. Identification results for DWT experiments and lower and higher Resolution V-DDBB.

Finally, Testing Process, these are not the final identification rates of the whole system. The results coming out of the SVM section are processed in the TU. After applying the TU's double condition the identification rate is dramatically increased. The new probability of success is  $P(X \geq 2)$ ; two or more recognitions in 3 seconds.

$$P(X \geq 2) = 1 - [P(X=0) + P(X=1)] \quad (23)$$

For our case, with 6 faces analyzed in 3 seconds, (4) can be expressed as:

$$P(X \geq 2) = 1 - [(1 - p)^6 + 6(1 - p)^5 p] \quad (24)$$

Here, 'p' is the probability of success in one attempt (percentages shown in Tables 6 and 7). Applying (5) to the results obtained with best combinations of ICA and DWT experiments with both lower and higher resolution DDBBs, really encouraging performances are obtained (see Table 8).

| SYSTEM'S PERFORMANCE  |         |
|---|---------|
| Lower Resolution  |         |
| DWT 'haar': 1 Iteration / 1440 Coefficients<br>80 for training / 20 for test        | 99,97 % |
| ICA: 40 Principal Components<br>60 for training / 40 for test                       | 99,68 % |
| Higher Resolution   |         |
| DWT 'bior 4.4': 2 Iterations / 838<br>Coefficients<br>80 for training / 20 for test | 100 %   |
| ICA: 30 Principal Components<br>60 for training / 40 for test                       | 99,93 % |

Table 8. Identification results for the whole system using best configuration.

## 6. Discussions and conclusions

A number of experiments have been done using different databases and configurations. In general terms, the results show very high standard deviations. This fact points out that results are highly dependent on the set of samples used for training and test. However, this may not be due system's instability, but due the experimental procedure. This refers to the fact that the same number of samples has been used to build up both positive and negative classes in an unbalance problem, where different classes need different number of training samples.

For experiments based on image databases, we have obtained a classification system, which can be used for arbitrary illumination conditions. We have searched a classifier with a good efficacy for fixed and dynamic illumination conditions, though the parameterization has to be different for reaching a better success rate. In particular, ICA has been used for a database with arbitrary illumination conditions, and DWT- Bior has been used for a database with fixed illumination conditions. The results are upper 98.9% for the studied cases.

For experiments based on videos databases, we have used the DWT and SVM classifier; we have created a system implemented in Matlab which is able to detect a subject in a video

sequence with an accuracy of 100%. The major errors detected here were maximum delay around 2 seconds. Even though for our V-DDBB subjects were always detected, more tests with a wider DDBB is needed before came to a conclusion in this aspect, and ensure that the system performance in perfect. For example, as it has been said before, the FEB does not perform perfectly, which means that the system does not always have 6 pictures of the subject's face in 3 seconds. Obviously, this plays against system's accuracy rate.

However, computing time has turned up to be a handicap of the resulted system. With and actual processing time of 5 times the length of the video, more research is needed in order to speed it up. One solution could be sharp the Face Extractor Block in order to increase its accuracy, reducing the number of false face founds. Tuning Training and Testing Blocks are always an interesting point, and along with the Face Extractor Block improvement the number of analyzed faces per second could be reduced, and therefore the whole processing time will be reduced to without decrease the system's accuracy.

Finally, processing time will be shorted again once the Matlab code has been translated to C++ and run as a normal application.

## 7. Acknowledgment

This work has been partially supported by "Cátedra Telefónica ULPGC 2009-10", and by the Spanish Government under funds from MCINN, with the research project reference "TEC2009-14123-C04-01".

## 8. References

- Akbari, R.; Bahaghighat, M.K. & Mohammadi, J. (2010). "Legendre moments for face identification based on single image per person", *2nd International Conference on Signal Processing Systems (ICSPS)*, Vo. 1, pp. 248-252, 5-7 July 2010
- AT&T Laboratories Cambridge (January 2002). ORL Face Database, 14-03-2011, Available from <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- Azam, M.; Anjum, M.A. & Javed, M.Y. (2010). "Discrete cosine transform (DCT) based face recognition in hexagonal images," *2nd International Conference on Computer and Automation Engineering (ICCAE)*, Vo. 2, pp. 474-479, 26-28 Feb. 2010
- Banu, R.V. & Nagaveni, N. (2009), "Preservation of Data Privacy Using PCA Based Transformation", *International Conference on Advances in Recent Technologies in Communication and Computing*, ARTCom '09., pp. 439 - 443
- Biometric International Group, (January 2011), 14.03.2011, Available from <http://www.ibgweb.com/>
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press.
- Bouchaffra, D. (2010). "Conformation-Based Hidden Markov Models: Application to Human Face Identification", *IEEE Transactions on Neural Networks*, vol.21, no.4, (April 2010), pp.595-608
- Bradski G., (March 2011), OpenCVWiki, 18-03-2011, Available from <http://opencv.willowgarage.com>
- Cevikalp, H., Neamtu, M., Wilkes, M. & Barkana, A. (2005). "Discriminative common vectors for face recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, No.1 pp. 4-13, (Jan. 2005)

- Chaari, A.; Lelandais, S. & Ben Ahmed, M. (2009). "A Pruning Approach Improving Face Identification Systems," *Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009*, pp.85-90, 2-4 Sept. 2009
- Chaudhari, S.T. & Kale, A. (2010), "Face Normalization: Enhancing Face Recognition," *3rd International Conference on Emerging Trends in Engineering and Technology (ICETET)*, pp. 520-525, 19-21 Nov. 2010
- Chen-Chung, Liu & Shiu-an-You, Chin, (2010). "A face identification algorithm using support vector machine based on binary two dimensional principal component analysis," *International Conference on Frontier Computing. Theory, Technologies and Applications, (2010 IET)*, pp. 193-198, 4-6 Aug. 2010
- Chennamma, H.R.; Rangarajan, Lalitha & Veerabhadrapa (2010). "Face Identification from Manipulated Facial Images Using SIFT," *3rd International Conference on Emerging Trends in Engineering and Technology (ICETET)*, pp. 192-195, 19-21 Nov. 2010
- Chi Ho, Chan; Kittler, J. (2005). "Sparse representation of (Multiscale) histograms for face recognition robust to registration and illumination problems," *17th IEEE International Conference on Image Processing (ICIP)*, pp. 2441-2444, 26-29 Sept. 2010
- Cichocki, A., Amari, S. I. Adaptive Blind Signal and Image Processing. John Wiley & Sons, 2002. In proceedings EUROSPEECH2001, 2603 - 2606, 2001.
- Cristianini, N. & Shawe-Taylor, J. (2000). *An introduction to support vector machines*, Cambridge University Press, 2000
- Face Recognition Homepage, (March 2011). Yale Database, 14-03-2011, Available from <http://www.face-rec.org/databases/>
- González R. & Woods R. (2002). *Digital Image Processing*, Ed. Prentice Hall
- Hamel, L. H. (2009). *Knowledge Discovery with Support Vector Machines*, Wiley & Sons
- Huang J.; Yuen, P.C.; Wen-Sheng Chen & Lai, J.H. (2003). "Component-based LDA method for face recognition with one training sample ", *IEEE International Workshop on Analysis and Modeling, of Faces and Gestures*, pp. 120 - 126
- Hyvärinen, A. ; Karhunen, J. & Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sons
- Joachims, T. (1999). "Making large-Scale SVM Learning Practica"l. *Advances in Kernel Methods - Support Vector Learning*, B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 25-07-2007, Available from : <http://svmlight.joachims.org/>
- Khairul, H.A. & Osamu, O. (2009). "Infrared-based face identification system via Multi-Thermal moment invariants distribution", *3rd International Conference on Signals, Circuits and Systems (SCS)*, pp.1-5, 6-8 Nov. 2009
- Kisku, D.R.; Rattani, A.; Grosso, E. & Tistarelli, M. (2007). "Face Identification by SIFT-based Complete Graph Topology," *IEEE Workshop on Automatic Identification Advanced Technologies*, pp.63-68, 7-8 June 2007
- Kurutach, Werasak; Fooprateepsiri, Rerkchai & Phoomvuthisarn, Suronapee (2010). "A Highly Robust Approach Face Recognition Using Hausdorff-Trace Transformation ", *Neural Information Processing. Models and Applications, Lecture Notes in Computer Science*, Springer Berlin / Heidelberg ; Vo. 6444, pp. 549-556, 2010
- Liun, R. & He, K. (2008). "Video Face Recognition on Invariable Moment". *International Conference on Embedded Software and Systems Symposia '08*, pp : 502 - 506, 29-31 July 2008
- MathWorks, (January 2011), MATLAB, 14.03.2011, Available from <http://www.mathworks.com/products/matlab/>

- Murata, N. ; Ikeda, S. & Ziehe, A. (2001). "An approach to blind source separation based on temporal structure of speech signals". *Neurocomputational*, Vo. 41, pp. 1–24
- Parra, L.(2002). *Tutorial on Blind Source Separation and Independent Component Analysis*. Adaptive Image & Signal Processing Group, Sarnoff Corporation
- Sang-Il, Choi ; Chong-Ho, Choi & Nojun Kwak, "Face recognition based on 2D images under illumination and pose variations", *Pattern Recognition Letters*, Vo. 32, No. 4, pp. 561-571, (March 2011)
- Saruwatari, H. ; Kawamura T. & Shikano, K. (2001). "Blind Source Separation for speech based on fast convergence algorithm with ICA and beamforming. 7th European Conference on Speech Communication and Technology ", In *EUROSPEECH-2001*, pp. 2603-2606, September 3-7, 2001
- Saruwatari, H. ; Sawai, K. ; Lee A. ; Kawamura T. ; Sakata M. & Shikano, K. (2003). "Speech enhancement and recognition in car environment using Blind Source Separation and subband elimination processing", *International Workshop on Independent Component Analysis and Signal Separation*, pp. 367 - 37
- Shreve, M.; Manohar, V.; Goldgof, D. & Sarkar, S. (2010), "Face recognition under camouflage and adverse illumination", *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp.1-6, 27-29 Sept. 2010
- Smith, L. I. (2006), "A tutorial on Principal Component Analysis", 01-12-2006, Available from :  
[http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf),
- Travieso, C.M. ; Botella, P. ; Alonso, J.B. & Ferrer M.A. (2009). "Discriminative common vector for face identification". *43rd Annual 2009 International Carnahan Conference on Security Technology, 2009*, pp.134-138, 5-8 Oct. 2009.
- Wen-Sheng, Chu ; Ju-Chin, Chen & Jenn-Jier, James Lien (2006). "Kernel discriminant transformation for image set-based face recognition" , *Pattern Recognition*, In Press, Corrected Proof, Available online 16 February 2011
- Xiong. G. (January 2005), The "localnormalize", 14.03.2011, Available from <http://www.mathworks.com/matlabcentral/fileexchange/8303-local-normalization>
- Yong, Xu ; Aini, Zhong ; Jian, Yang & Zhang, David (2010). "LPP solution schemes for use with face recognition", *Pattern Recognition*, Vo.43, No. 12, (December 2010), pp 4165-4176
- Yossi, Zana & Roberto M. Cesar, Jr. (2006). "Face recognition based on polar frequency features." *ACM Trans. Appl. Percept.* 3, 1 (January 2006), pp. 62-82
- Zhao Lihong; Zhang Cheng; Zhang Xili; Song Ying; Zhu Yushi (2009), "Face Recognition Based on Image Transformation", *WRI Global Congress on Intelligent Systems, 2009*. GCIS '09., vol.4, pp. 418-421, 19-21 May 2009
- Zhou, H. & Sadka, A. H. (2010). "Combining Perceptual Features With Diffusion Distance for Face Recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Review*, No. 99, (junio 2010) pp. 1-12



# Towards Unconstrained Face Recognition Using 3D Face Model

Zahid Riaz<sup>1</sup>, M. Saquib Sarfraz<sup>2</sup> and Michael Beetz<sup>1</sup>

<sup>1</sup> *Intelligent Autonomous Systems (IAS), Technical University of Munich, Garching*

<sup>2</sup> *Computer Vision Research Group, COMSATS Institute of Information*

*Technology, Lahore*

<sup>1</sup>*Germany*

<sup>2</sup>*Pakistan*

## 1. Introduction

Over the last couple of decades, many commercial systems are available to identify human faces. However, face recognition is still an outstanding challenge against different kinds of real world variations especially facial poses, non-uniform lightings and facial expressions. Meanwhile the face recognition technology has extended its role from biometrics and security applications to human robot interaction (HRI). Person identity is one of the key tasks while interacting with intelligent machines/robots, exploiting the non intrusive system security and authentication of the human interacting with the system. This capability further helps machines to learn person dependent traits and interaction behavior to utilize this knowledge for tasks manipulation. In such scenarios acquired face images contain large variations which demands an unconstrained face recognition system.

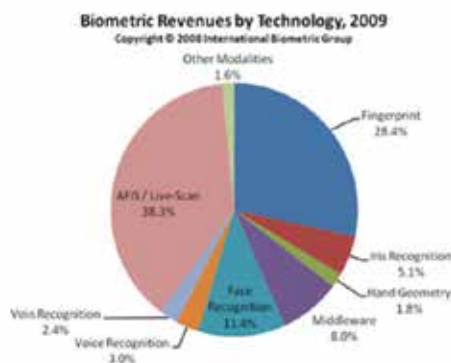
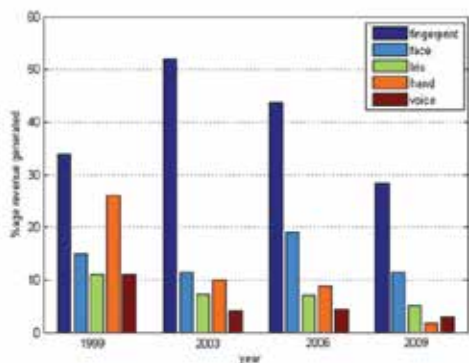


Fig. 1. Biometric analysis of past few years has been shown in figure showing the contribution of revenue generated by various biometrics. Although AFIS are getting popular in current biometric industry but faces are still considered as one of the widely used biometrics.

By nature, face recognition systems are widely used due to high universality, collectability and acceptability. However this attribute has natural challenges like uniqueness, performance and circumvention. Although several commercial biometrics systems use face recognition as a key tool but most of them perform in constrained environment. Real world challenges to face recognition problems involve varying lighting, poses, facial expressions, aging effects, occlusions which also include make up, facial scars or cuts, facial hairs, low resolution images and most recently spoofing. Due to their non intrusiveness faces are the most important biometrics to be employed in real life systems. Figure 1 shows the contribution of the face recognition system in biometrics market. Over the last decade, faces have been consistently used after the usage of fingerprints and AFIS (automated fingerprints identification system)/live scans. Face recognition finds its major applications in document classification, security access control, surveillance, web application and human robot interaction. However fingerprints can easily be spoofed (research is still in progress to deal with this issue using live scans), intrusive, aging effects are more in the sense of damages and finally template security and updation issues.

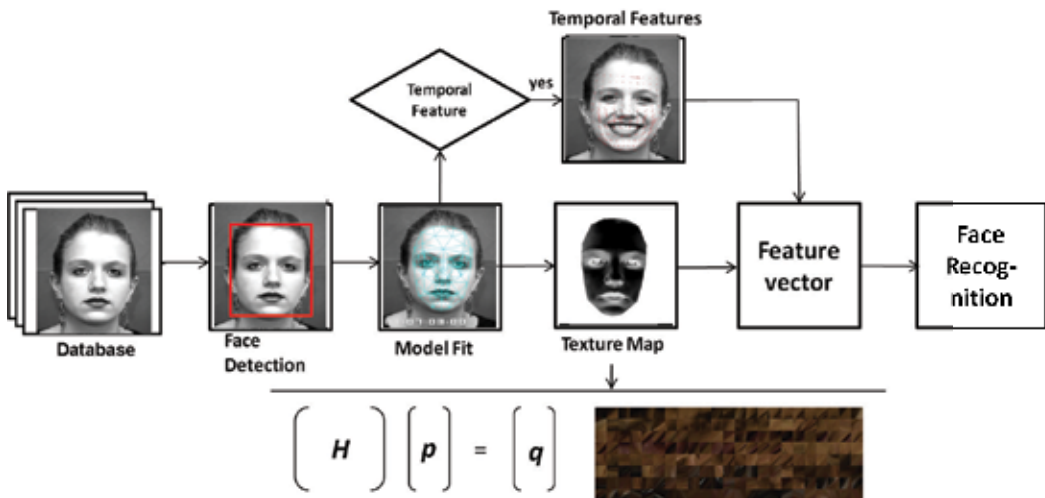


Fig. 2. Detailed process for model based face recognition system. Texture map is generated by using homogeneous 2D point  $p$  in texture coordinates and a projection  $q$  of a general 3D point in homogeneous coordinates by using transformation  $H \times p = q$ . Where  $H$  is the homography matrix Riaz et al. (2010).

In this chapter we focus on an unconstrained face recognition system and also study other soft biometric traits of human faces which can be useful for man machine interaction applications. The proposed system can be applied in major biometrics application equally well. The goal of this chapter is two folded. Firstly, it serves as self-contained and compact tutorial for face recognition systems from design to its development. It describes a brief history of face recognition systems, major approaches used and challenges. Secondly, it describes in detail an approach towards the development of a robust face recognition system against varying poses, facial expressions and lighting. This approach uses a 3D wireframe model which is useful for such applications. In order to proceed, a comprehensive overview of the face models currently used in the area of computer vision applications is provided. This covers an overview about deformable models, point distribution models, photorealistic models and finally wireframe

models. In order to provide an automatic generation of a realistic human face model, a 3D wireframe model called Candide-III Ahlberg (2001) is used. This model has benefit over the others models in the sense of its speed, textures realization and well-defined of facial animation units. On the other hand realizing this model is more challenging since it is less detailed as compared to 3D morphable model which consists of a dense point distribution of facial points acquired from laser scans. Candide-III consists of a coarse mesh containing only 184 triangulations and defines facial action coding system (FACS) Ekman & Friesen (1978), which can easily be used for facial animations. Figure 2 show the overview of different modules working in our designed system.

## 2. System overview

The system mainly comprise of different modules contributing towards the final feature extraction. It starts with a face detection module followed by a face model fitting. Candide-III is fitted to face images using robust objective functions Wimmer et al. (2006). This fitting algorithm is compared with other algorithms and has been adopted here due to their efficiency in real time and robustness. The model is fitted to any face image with arbitrary pose, however pre-requisite for this module is face detection. If face detection module fails to find any face inside the image then system searches again for a face in next image. After model fitting, extracted structural information from this image contains shape and pose information. The fitted model is used for extracting the texture from given image. We use graphic tools to render texture to model surface and warp the texture to standard template. This standard template is a block based texture map shown in Figure 2. Texture variations are majorly caused by illuminations and shadows, which are dealt with textural feature and image filtering. We use principal component analysis (PCA) to parameterize extracted textures. Cognitive science explains that temporal information is an essential constituent for human perception and learning Sinha et al. (2006). In this regard, we choose local descriptors on the face image and track their motion to find temporal features. This motion provides local variations and deformations in the face image caused by facial expressions. Finally we construct a feature vector set by concatenating these three different features for any given image.

$$u = (b_s, b_g, b_t) \quad (1)$$

Where  $b_s$ ,  $b_g$  and  $b_t$  are the structural, textural and temporal features. For texture extraction we comparatively perform different methods including discrete cosine transform (DCT), principal components analysis (PCA) and local binary patterns (LBP). From these three types of textural feature, it is observed that PCA outperforms other two feature set and hence used for further experimentation. The results are given in detail in section 10.

## 3. Related work

Face recognition has always been a focus of attention for the researchers and has been addressed in different ways by using holistic and heuristic features, dense and sparse feature representation, parametric and non-parametric models and content- and context-aware methods Zhao et al. (2003)Zhao & Chellappa (2005). Due to the universality, collectability non-intrusiveness, face recognition is currently used as a baseline algorithm in several biometric systems either as a standalone technology or together with other biometrics, called multibiometric systems. In 3D space, faces are complex objects consisting of a regularized

structure and pigmentation and mostly observed performing various action and conveying a set meaningful information O'Toole (2009). Besides these meaningful set of information, faces convey several challenges which are under consideration by the research community. Human facial recognition system is unconstrained and provides stability against varying poses, facial expressions, changing illuminations, partial occlusions (including facial hair, scars and make ups) and temporal effects (aging factors).

Traditional recognition systems have the abilities to recognize the human using various techniques like feature based recognition, face geometry based recognition, classifier design and model based methods. In Zhao et al. (2003) the authors give a comprehensive survey of face recognition and some commercially available face recognition software. Subspace projection method like PCA was firstly used by Sirovich and Kirby Sirovich & Kirby (1987), which were latterly adopted by M. Turk and A. Pentland introducing the famous idea of eigenfaces Turk & Pentland (1991). This chapter focuses on the modeling of human face using a three dimensional model for shape model fitting, texture and temporal information extraction and then low dimensional parameters for recognition purposes. The model using shape and texture parameters is called Active Appearance Model (AAMs), introduced by Cootes et. al. Cootes et al. (1998) Edwards, Taylor & Cootes (1998). For face recognition using AAM, Edwards et al Edwards, Cootes & Taylor (1998) use weighted distance classifier called Mahalanobis distance. In Edwards et al. (1996) the authors used separate information for shape and gray level texture. They isolate the sources of variation by maximizing the interclass variations using discriminant analysis, similar to Linear Discriminant Analysis (LDA), the technique which was used for Fisherfaces representation Belhumeur et al. (1997). Fisherface approach is similar to the eigenface approach however outperforms in the presence of illuminations. In Wimmer et al. (2009) the authors have utilized shape and temporal features collectively to form a feature vector for facial expressions recognition. These models utilize the shape information based on a point distribution of various landmarks points marked on the face image. Blanz et al. Blanz & Vetter (2003) use state-of-the-art morphable model from laser scanner data for face recognition by synthesizing 3D face. This model is not as efficient as AAM but more realistic. In our approach a wireframe model known as Candide-III Ahlberg (2001) has been utilized. In order to perform face recognition applications many researchers have applied model based approach. Riaz et al Riaz, Mayer, Wimmer, Beetz & Radig (2009) apply similar features for explaining face recognition using 2D model. They use expression invariant technique for face recognition, which is also used in 3D scenarios by Bronstein et al Bronstein et al. (2004) without 3D reconstruction of the faces and using geodesic distance. Park et. al. Park & Jain (2007) apply 3D model for face recognition on videos from CMU Face in Action (FIA) database. They reconstruct a 3D model acquiring views from 2D model fitting to the images.

In Riaz, Mayer, Beetz & Radig (2009a) author introduced spatio-temporal feature for expressions invariant face recognition. This chapter is an extended version of the similar approach but with improved texture realization and texture descriptors. Further the feature set used in this chapter are stable against facial poses.

#### **4. Spatio-temporal Multifeatures (STMF)**

We address the problem in which a 3D model can extract a common feature set automatically from face images and performs unconstrained face recognition. This system can not only be used for biometric applications but also for soft biometric traits document classification,

|                   | Identity | Expressions | Gender   | Age       | Ethnicity |
|-------------------|----------|-------------|----------|-----------|-----------|
| <b>Structural</b> | <i>p</i> | <i>p</i>    | <i>p</i> | <i>p</i>  | <i>p</i>  |
| <b>Textural</b>   | <i>p</i> | <i>s</i>    | <i>p</i> | <i>p</i>  | <i>p</i>  |
| <b>Temporal</b>   | <i>s</i> | <i>p</i>    | <i>s</i> | <i>na</i> | <i>na</i> |

Table 1. Contribution of different components to our proposed feature set, *p* = primary, *s* = secondary, *na* = not applicable.

access control and medical care. In such applications an automatic and efficient feature extraction technique is necessary to interpret maximum available information from the faces image sequences. Further, the extracted feature set should be robust enough to be directly applied in real world applications. In such scenarios faces are seen from different views under varying facial deformations and poses. This issue is treated by using 3D modeling of the faces. The invariance to facial poses and expressions is discussed in detail in section 10.

For face recognition textural information plays a key role as features Zhao & Chellappa (2005) Li & Jain (2005) whereas facial expressions are mostly person independent and require motion and structural components Fasel & Luetten (2003). Similar facial structure and texture vary significantly between gender classes. On the basis of this knowledge and literature survey we categorize three major features as primary and secondary contributor to three different facial classifications. Table 1 summarizes the significance of these constituents of the feature vector with their primary and secondary contribution towards the feature set formation. Since our feature set consists of all three kinds of information hence it can successfully represent facial identity, expressions and gender. The results are discussed in detail in section 10. Model parameters are obtained in an optimal way to maximize information within the face region in the presence of different facial pose and expressions. We use a 3D wireframe model however, any other comparable model can be used here.

## 5. Model fitting and structural features

Our proposed algorithm is initialized by applying a face locator in the given image. We use Viola and Jones face detector Viola & Jones (2004). If a face is found then the system proceeds towards face model fitting. For model fitting, local objective functions are calculated using haar-like features. An objective function is a cost function which is given by the equation 2. A fitting algorithm searches for the optimal parameters which minimizes the value of the objective function. For a given image  $I$ , if  $E(I, c_i(\mathbf{p}))$  represents the magnitude of the edge at point  $c_i(\mathbf{p})$ , where  $\mathbf{p}$  represents set of parameters describing the model, then objective function is given by:

$$f(I, \mathbf{p}) = \frac{1}{n} \sum_{i=1}^n f_i(I, c_i(\mathbf{p})) = \frac{1}{n} \sum_{i=1}^n (1 - E(I, c_i(\mathbf{p}))) \quad (2)$$

Where  $n = 1, \dots, 113$  is the number of vertices  $c_i$  describing the face model. This approach is less prone to errors because of better quality of annotated images which are provided to the system for training. Further, this approach is less laborious because the objective function design is replaced with automated learning. For details we refer to Wimmer et al. (2008). The geometry of the model is controlled by a set of action units and animation units. Any shape  $s$  can be written as a sum of mean shape  $\bar{s}$  and a set of action units and shape units.

$$s(\alpha, \sigma) = \bar{s} + \phi_a \alpha + \phi_s \sigma \quad (3)$$

Where  $\phi_a$  is the matrix of action unit vectors and  $\phi_s$  is the matrix of shape vectors. Whereas  $\alpha$  denotes action units parameters and  $\sigma$  denotes shape parameters Li & Jain (2005). Model deformation governs under facial action coding systems (FACS) principles Ekman & Friesen (1978). The scaling, rotation and translation of the model is described by

$$s(\alpha, \sigma, \pi) = mRs(\alpha, \sigma) + t \quad (4)$$

Where  $R$  and  $t$  are rotation and translation matrices respectively,  $m$  is the scaling factor and  $\pi$  contains six pose parameters plus a scaling factor. By changing the model parameters, it is possible to generate some global rotations and translations. We extract 85 parameters to control the structural deformation.

## 6. Textural representation and parameterization

For texture extraction from the face images after model fitting, two different approaches are studied in this chapter. In section 6.1 texture extraction is performed using conventional AAM method, whereas in section 6.2 texture mapping approach is studied. Texture map is formed by storing each triangular patch to a block in memory. This block represents surface texture extracted from 3D surface of the face. Once texture is extracted, it is parametrized by using mean texture  $g_m$  and matrix of eigenvectors  $P_g$  to obtain the parameter vector  $b_g$  Li & Jain (2005).

$$g = g_m + P_g b_g \quad (5)$$

### 6.1 Image warping

Once structural information of the image is obtained from model fitting, we extract texture from the face region by mapping it to a reference shape. A reference shape is extracted by finding the mean shape over the dataset. Image texture is extracted using planar subdivisions of the reference and the example shapes. We use delaunay triangulations for the convex hull of the facial landmarks. Texture warping between the triangles is performed using affine transformation. This texture warping is used for CKFE database Kanade et al. (2000) and MMI database Maat et al. (2009). By warping texture to a reference shape, facial expressions are neutralized and hence useful for face recognition.

### 6.2 Optimal texture representation

Each triangular patch represents meaningful texture which is stored in a square block of the texture map. A single unit of the texture map represents a triangular patch. We experiment with three different sizes of the texture blocks and choose an optimal size for our experimentation. These three block sizes include  $2^3 \times 2^3$ ,  $2^4 \times 2^4$  and  $2^5 \times 2^5$ . We calculate energy function from these texture maps of individual persons and observe the energy spectrum of the images in our database for each triangular patch. If  $N$  is the total number of images, and  $p_i$  be a texel value (which is equal to a single pixel value) in texture map, then we define energy function as:

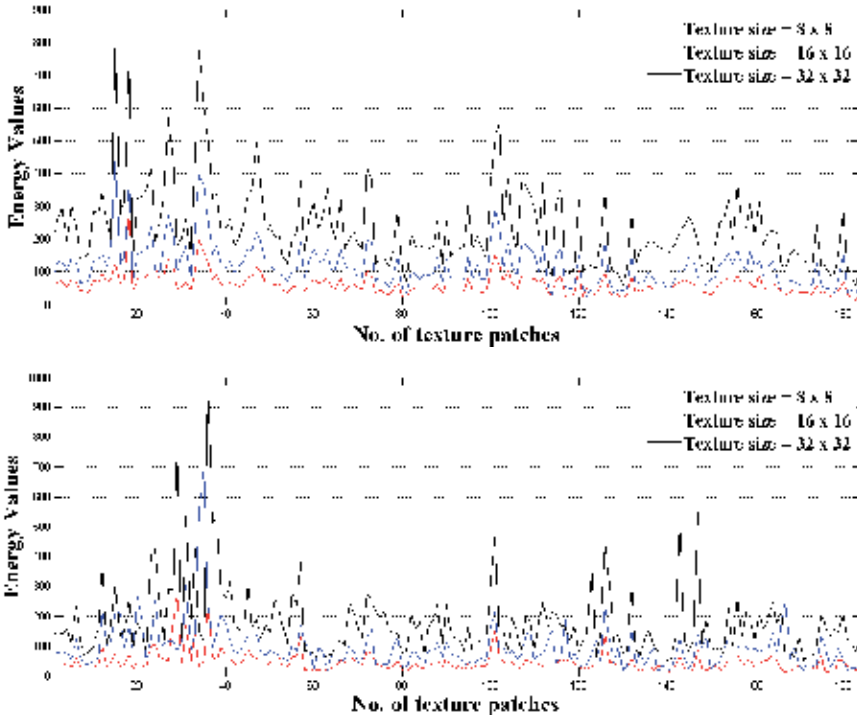


Fig. 3. Energy spectrum of two randomly selected subjects from PIE database. Energy values for each patch is comparatively calculated and observed for three different texture sizes.

$$E_j = \frac{1}{N} \sum_{i=1}^N (p_i - \bar{p}_j)^2 \quad (6)$$

Where  $\bar{p}_j$  is the mean value of the pixels in  $j^{th}$  block,  $j = 1 \dots M$  and  $M = 184$  is the number of blocks in a texture map. In addition to Equation 6, we find variance energy by using PCA for each block and observe the energy spectrum. The variation within the given block has similar behavior for two kinds of energy functions except a slight variation in the energy values. Figure 3 shows the energy values for two different subjects randomly chosen from our experiments. It can be seen from Figure 3 that behavior of the textural components is similar between different texture sizes. The size of the raw feature vector extracted directly from texture map increases exponentially with the increase of texture block size. If  $d \times d$  is the size of the block, then the length of the raw feature vector is  $\frac{d(d+1)}{2}$ . This vector length calculation depends upon how texture is stored in the texture map. This can be seen in Figure 4. We store each triangular patch from the face surface to upper triangle of the texture block. The size of raw feature vector extracted for  $d = 2^3$ ,  $d = 2^4$  and  $d = 2^5$  is 6624, 25024 and 97152 respectively. Any higher value will exponentially increase the raw vector without any improvement in the texture energy. We do not consider higher values due to increase in vector length. The overall recognition rate produced by different texture sizes from eight randomly selected subjects with 2145 images from PIE database is shown in Figure 5. The results are obtained using decision trees and Bayesian networks for classification. The classification

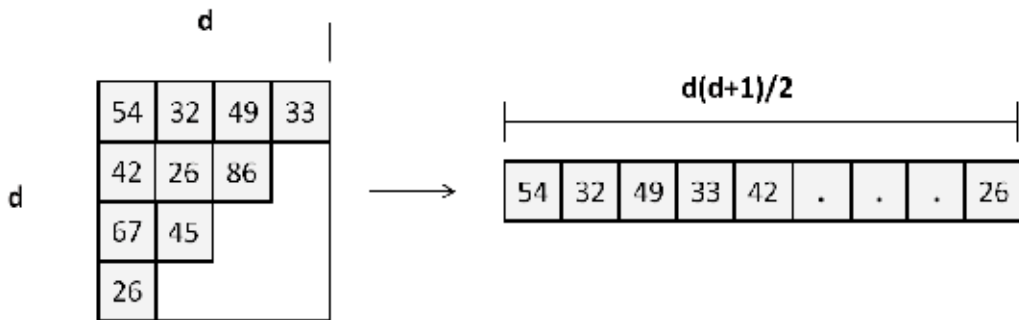


Fig. 4. Texture from each triangular patch is stored as upper triangle of the texture block in texture map. A raw feature vector is obtained by concatenating the pixel values from each block.

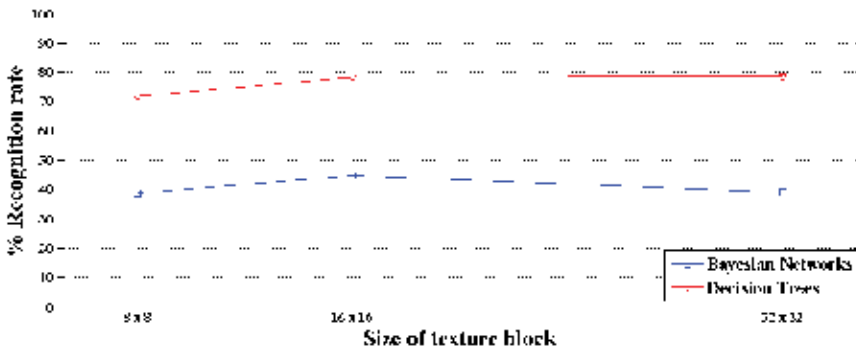


Fig. 5. Comparison over eight random subjects from the database with three different sizes of texture blocks. Recognition rate slightly improved as texture size is increased however causes a high increase on the length of raw feature vector. We compromise on texture block of size  $16 \times 16$ .

procedure is given in detail in next section 10. By trading off between the performance and size of the feature vectors, we choose texture block size to  $16 \times 16$  during our experiments.

### 6.3 Affine vs. perspective transformation

We consider perspective transformation because affine warping of the rendered triangle is not invariant to 3D rigid transformations Riaz et al. (2010). In general, texture warping is performed using affine transformation from a given image to a reference shape Cootes et al. (1998)Riaz, Mayer, Wimmer, Beetz & Radig (2009)Riaz, Mayer, Beetz & Radig (2009b)Wimmer et al. (2009). This preserves affinity after the transformation. However, for faces with different views triangular patches on the edges are not well defined and these triangles are tilted such that they contain very less information about the texture as compared to those triangles which are frontal. In order to equally weight all triangles and we use a homogeneous transformation.



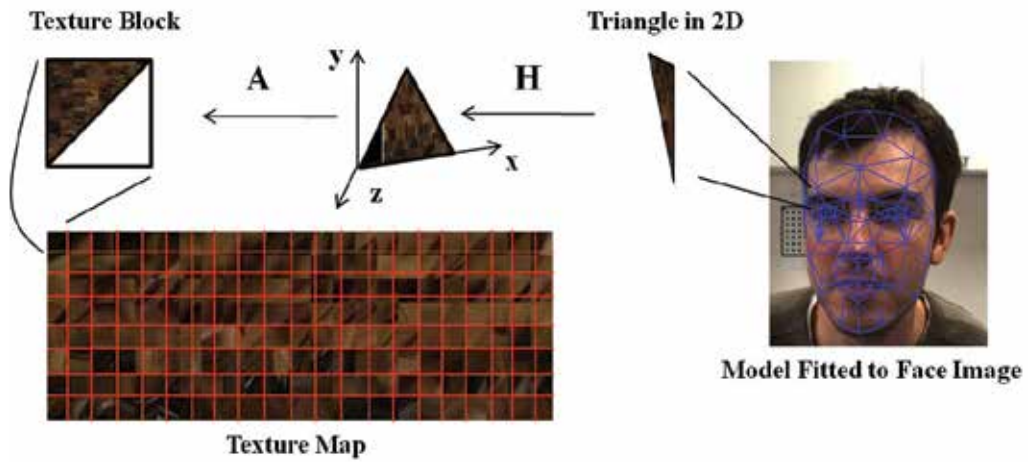


Fig. 6. Detailed texture extraction approach. Each triangular patch from the face surface is stored as a block in a texture map. This texture map is further used for feature extraction.

The homogeneous transformation  $\mathbf{M}$  is the given by

$$M = AK [R \ -Rt] = AH \quad (7)$$

Where  $\mathbf{K}$  is the camera matrix. An undistorted texture map from the face region is calculated in two steps. Firstly, we find the homography  $\mathbf{H}$ , by obtaining the rotation and translation of the triangle, by supposing that the initial triangle lies on the texture plane, the first vertex lies on the origin and the first edge lies on the x-axis. Secondly, affine transformation  $\mathbf{A}$  is calculated, so that the mapped triangle on the texture plane fits the upper triangle of the rectangular texture block. The lower triangular area is not considered in this regard.

Where  $\mathbf{R}$  and  $\mathbf{t}$  are the unknown to be calculated. In order to fit any arbitrary triangle to this upper triangle, we use an affine transformation  $\mathbf{A}$ . This process is shown in Figure 6. For detail, refer to Riaz et al. (2010).

## 7. Temporal features

Further, temporal features of the facial changes are also calculated that take movement over time into consideration. Local motion of feature points is observed using optical flow. We do not specify the location of these feature points manually but distribute equally in the whole face region. The number of feature points is chosen in a way that the system is still capable of performing in real time and therefore inherits a tradeoff between accuracy and runtime performance. Since the motion of the feature points are relative so we choose 140 points in total to observe the optical flow. We again use PCA over the motion vectors to reduce the descriptors.

If  $\mathbf{t}$  is the velocity vector,

$$\mathbf{t} = \mathbf{t}_m + \mathbf{P}_t \mathbf{b}_t \quad (8)$$

Where temporal parameters  $\mathbf{b}_t$  are computed using matrix of eigenvectors  $\mathbf{P}_t$  and mean velocity vectors  $\mathbf{t}_m$ .

| Feature Type     | LBP Features | DCT Features | PCA Features |
|------------------|--------------|--------------|--------------|
| Recognition Rate | 78.56%       | 80.60%       | 83.06%       |

Table 2. Comparison of three different feature types for face recognition. The overall recognition rate shows the number of images correctly classified. PCA outperforms the other two feature types and is used for further experimentations.

## 8. Feature fusion

We combine all extracted features into a single feature vector. Single image information is considered by the structural and textural features whereas image sequence information is considered by the temporal features. The overall feature vector becomes:

$$u = (b_{s,1}, \dots, b_{s,m}, b_{g,1}, \dots, b_{g,n}, b_{t,1}, \dots, b_{t,p}) \quad (9)$$

Where  $b_s$ ,  $b_g$  and  $b_t$  are shape, textural and temporal parameters respectively with  $m$ ,  $n$  and  $p$  being the number of parameters retained from subspace in each case. Equation 9 is called multi-feature. We extract 85 structural features, 74 textural features and 12 temporal features textural parameters to form a combined feature vector for each image. These features are then used for decision tree (DT) and bayesian network (BN) for different classifications. The face feature vector consists of the shape, texture and temporal variations, which sufficiently defines global and local variations of the face. All the subjects in the database are labeled for classification. Since features arise from different sources, it is not quite obvious to fuse them together to get a feature set. This can cause the dominance of the features with higher values and ones with low values are ignored. We use simple scaling of the features in  $[0, 1]$ . However, any suitable method for feature fusion can be applied here.

## 9. Comparative texture descriptors

From above texture representations, features are extracted using three different approaches, a) PCA, b) discrete cosine transform (DCT) and c) local binary pattern (LBP). Each texture map consists of 184 texture blocks where each texture block corresponds to texture in a triangular surface. The size of each block is  $16 \times 16$  pixels. This size is chosen by trading off between accuracy and efficiency. DCT coefficients are extracted in a zig-zag pattern from top-left corner of each block. We extract five coefficients per block and obtain a feature set of length  $5 \times 184 = 920$ . The advantage of using DCT over conventional approach is two fold, 1) it reduces the dimensions to a great extent, 2) DCT coefficients contain low frequency information which are robust to distortions and noise. For LBP descriptor, we consider those pixels for coding which lie inside face area. An LBP histogram of 255 gray levels and color histogram are calculated and used as texture features. The results of three different features types on all subjects of PIE database session from November 2000 to December 2000 Terence et al. (2002) is shown in Table 2. A J48 decision tree from Weka Witten & Frank (2005) is used as classifier. The detail about classifier specification is given in section 10. It can be seen that PCA outperforms the other two features types. For further experimentation, we use PCA for feature extraction.

### 9.1 Local energy based descriptors

From section 6.2, we have texture extracted from the face images and stored in a texture map in triangular patches of same sizes. Each patch represents a specific area of the face. Instead

of using conventional subspace learning which do not preserve localization information, we find heuristic features representative of each patch. If training data consists of  $M$  images, then we can calculate the variance energy of each patch using:

$$e_j = \frac{1}{M} \sum_{i=1}^N (p_i - \bar{p}_{ji})^2 \quad (10)$$

Where  $i = 1, \dots, N$  and  $N = 136$ , which is the number of pixel in triangle from texture map,  $\bar{p}_j$  is the average value of  $j$ th patch, with  $j = 1, 2, \dots, 184$ . Finally a feature vector  $E$  is formed by finding energy descriptor for each triangular patch and is written as:

$$E = \{e_1, e_2, \dots, e_j\} \quad (11)$$

There are three major benefits of using patch based representation.

- The extracted feature vector although is compact and sufficient to perform well in view invariant face recognition.
- It avoids subspace learning and new faces can be added easily in the database.
- Since such representation preserves localization of the facial features, it outperforms the conventional AAM and holistic approaches.

## 10. Experimental evaluation

In order to validate the extracted feature, we have used different subjects from three different databases called, CMU-PIE database Terence et al. (2002), MMI database Maat et al. (2009) and Cohn-Kanade facial expressions database (CKFED) Kanade et al. (2000). These databases consist of face images with different variations, like varying poses, facial expressions, gender information and talking faces. MMI and CKFED contain image sequences with temporal information. CKFED consists of 97 subjects range in age from 18 to 30 years. Sixty-five percent are female, 15 percent are African-American and three percent Asian or Latino. The MMI facial expression database holds over 2000 videos and over 500 images of about 50 subjects displaying various facial expressions on command. In case of CKFE and MMI databases we compute spatio-temporal feature from the image sequences. CMU-PIE database is collected between October and December 2000 consisting of 41,368 images of 68 people. Each person is captured with 13 different poses, 43 different illumination conditions, and with 4 different expressions. We take spatial feature and test them against frontal and half-profile poses. The texture extracted in this case is stored as a texture map after removing perspective distortion. During all experiments, we use two-third of the feature set for building the classifier model with 10-fold cross validation to avoid overfitting. The remaining feature set is used for testing purpose. We use all subjects from MMI and CKFED and partially use PIE database to perform face recognition and person dependent facial expressions and gender classification.

Since STMF set arises from different sources, so decision tree (DT) is applied for classification. However, other classifiers can also be applied here depending upon the application. We choose J48 decision tree algorithm for experimentation which uses tree pruning called subtree raising and recursively classifies until the last leaf is pure. We use same configuration for all classifiers trained during the experiments. The parameters used in decision tree are: confidence factor  $C = 0.25$ , with two minimum number of instances per leaf and C4.5 approach for reduced error-pruning Witten & Frank (2005). For further validation, we use random

| <b>PIE - Session (Oct-Nov 2000)</b> |                       |                           |                       |
|-------------------------------------|-----------------------|---------------------------|-----------------------|
| Vector Size                         | 181                   | 257                       | <b>184</b>            |
| Random Forest                       | 64.45%                | 70.12%                    | <b>89.12%</b>         |
| Decision Tree                       | 61.29%                | 69.65%                    | <b>75.13%</b>         |
| <b>PIE - Session Nov-Dec 2000</b>   |                       |                           |                       |
| Vector Size                         | 337                   | 295                       | <b>184</b>            |
| Random Forest                       | 90.15%                | 90.60%                    | <b>98.50%</b>         |
| Decision Tree                       | 88.42%                | 89.53%                    | <b>94.89%</b>         |
| <b>Feature Type</b>                 | <b>Shape free AAM</b> | <b>3D Surface Texture</b> | <b>Energy feature</b> |

Table 3. Comparison of three different approaches used for face recognition. The overall recognition rate shows the number of images correctly classified. PCA representation of the surface texture from a face outperforms conventional AAM approach, however energy based descriptors are not only compact but perform even better than two other approaches.

|       | <b>BN</b> | <b>DT</b> |
|-------|-----------|-----------|
| CKFED | 90.66%    | 98.50%    |
| MMI   | 90.32%    | 99.29%    |

Table 4. Face recognition across facial expressions on two different databases using Bayesian networks (BN) and decision tree (DT).

forests for classification with default Weka Witten & Frank (2005) parameters and 10-fold cross validation. The results coincide with those from decision trees. During all experimentation, we use same training and testing approach. For subspace learning, one-third of the database is used while the remaining part is projected to this space. We retain 97% of the eigenvalues during the subspace learning.

### 10.1 Expression invariant face recognition

MMI and CKFE databases contain six basic facial expressions in the form of image sequences. Although neutral expression is present as a seventh expression but we exclude it during experiments and solve the problem as six class problem. All images are frontal and hence we use spatio-temporal features with texture warped to reference shape. Most of the face recognition information is available in textural components and hence we obtain stability of our feature in face recognition results in the presence of facial expressions. Texture warping neutralizes the effect of facial expressions. The recognition results using decision tree and Bayesian networks are shown in Table 4. However, structural and temporal part of the same feature set contain sufficient facial expression information (refer to Table 1). In this way, a single STMF is representative of facial expression, face recognition and gender information.

### 10.2 Pose invariant face recognition

In section 6.2, we explained the detailed process for texture extraction. Since the model is defined over a coarse mesh of vertices, so it is useful to consider texture map as an image with undistorted texture patches. In the presence of different facial poses, triangles at the face edges are tilted such that texture information is extremely distorted. In order to solve this problem, we project each triangle on a block of  $16 \times 16$  pixel size. The block size is chosen by trading off between efficiency and accuracy. In this procedure, each triangle is weighted

| Approach                             | % Accuracy   |
|--------------------------------------|--------------|
| model based Mayer et al. (2009)      | 87.1%        |
| TAN Cohen et al. (2003)              | 83.3%        |
| LBP + SVM Shan et al. (2009)         | 92.6%        |
| IEBM Asthana et al. (2009)           | 92.9%        |
| Fixed Jacobian Asthana et al. (2009) | 89.6%        |
| <b>STMF + DT</b>                     | <b>93.2%</b> |

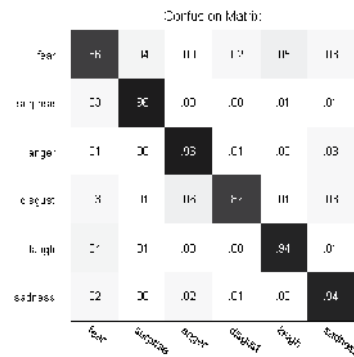


Table 5. Facial expressions recognition in comparison to different approaches Asthana et al. (2009) and Confusion matrix from our results

| Approach                 | Classification rate |
|--------------------------|---------------------|
| Pixels + SVM + Fusion    | 88.5%               |
| LBP + SVM + Fusion       | 92.1%               |
| VLBP + SVM               | 84.5%               |
| EVLBP + AdaBoost         | 84.6%               |
| <b>our approach+ BDT</b> | <b>94.8%</b>        |

Table 6. Comparison of gender classification in comparison to different approaches in Hadid & Pietikainen (2009)

equally toward the feature calculation and face edges are not destroyed but rather provide the detailed texture information that might be lost during conventional image warping approach. The recognition results with different approaches are given in Table 3.

### 10.3 Facial expressions and gender classification

Facial expressions recognition is performed on CKFED with six universal facial expressions: anger, disgust, fear, laugh, sadness and surprise. Each video sequence starts from a neutral face and reaches up to the peak of the particular expression. We exclude neutral expression during the experiments because it is included in all image sequences and cause more confusion. However, a neutral expression can also be considered as a seventh expression during classification. It can be automatically segmented using velocity vectors magnitudes. STMF features with their three structural, textural and temporal constituents are used for experiments. Finally, the results are compared in Table 5 with the state-of-the-art approaches which uses comparable system along with the confusion matrix from our experiments.

We further estimated age using FGNet database *FG-NET AGING DATABASE* (n.d.). This database contains 1002 images of 62 subjects with images of different ages ranging from 0 to 69 years. We divide the whole dataset in seven classes. Since the database consists of static images hence we experiment only with shape and textural component of the feature set. A classification rate of 49.70% is achieved with texture whereas the classification rate improved to 57.29% using support vector machine based classification. The mean absolute error (MAE) was 0.769.

## 11. Conclusions and future work

This chapter explained the STM for unconstrained face recognition. The spatial part of this feature set consists of structural (section 5) and textural (section 6) information. Two different types of texture extraction approaches are discussed in detail in section 6.1 and section 6.2. Further a comparative study of three different textural features has been studied which shows that PCA outperforms LBP and DCT. Since PCA is global representation of a face, hence it does not contain local information. In section 9.1 a local representation for each triangular patch is calculated. Since this local representation is added to the extracted features, hence it further improves the results and outperforms holistic PCA. Since feature set given in equation 9 is consistent with Table 1 hence it can be used for facial expression recognition, gender classification and age estimation. The results are shown in section 10.1. This chapter provides a comprehensive overview and a compact description of 3D face modeling, face recognition, classifying soft-biometric traits including facial expressions, gender and age. However such systems require more memory and are relatively slower as compared to conventional image based approaches. Future goal of this work is to enhance its efficiency to apply it in real time for interactive systems. Further more diverse conditions with large variabilities are required to be tested.

## 12. References

- Ahlberg, J. (2001). An experiment on 3d face model adaptation using the active appearance algorithm, *Image Coding Group, Dept of Electric Engineering, Linköping University*.
- Asthana, A., Saragih, J., Wagner, M. & Goecke, R. (2009). Evaluating aam fitting methods for facial expression recognition, *Affective Computing and Intelligent Interaction* 1(1): 598–605.
- Belhumeur, P., Hespanha, J. & Kriegman, D. (1997). Eigenfaces vs fisherfaces: Recognition using class specific linear projection, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 19(7).
- Blanz, V. & Vetter, T. (2003). Face recognition based on fitting a 3d morphable model, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(9): 1063–1074.
- Bronstein, A., Bronstein, M., Kimmel, R. & Spira, A. (2004). 3d face recognition without facial surface reconstruction, *Proceedings of European Conference of Computer Vision*.
- Cohen, I., Sebe, N., Garg, A., Lawrence, C. & Huanga, T. (2003). Facial expression recognition from video sequences: temporal and static modeling, *Computer Vision and Image Understanding*, Elsevier Inc., pp. 160–187.
- Cootes, T., Edwards, G. & Taylor, C. (1998). Active appearance models, *Proceedings of European Conference on Computer Vision* 2: 484–498.
- Edwards, G., Cootes, T. & Taylor, C. (1998). Face recognition using active appearance models, *European Conference on Computer Vision*, Springer, pp. 581–695.
- Edwards, G. J., Taylor, C. J. & Cootes, T. F. (1998). Interpreting face images using active appearance models, *Proceedings of the 3rd. International Conference on Face & Gesture Recognition*, FG '98, IEEE Computer Society, Washington, DC, USA, pp. 300–. URL: <http://portal.acm.org/citation.cfm?id=520809.796067>
- Edwards, G., Lanitis, C., Taylor, C. & Cootes, T. (1996). Statistical models of face images: Improving specificity., *British Machine Vision Conference*.

- Ekman, P. & Friesen, W. (1978). The facial action coding system: A technique for the measurement of facial movement, *Consulting Psychologists Press*.
- Fasel, B. & Luetttin, J. (2003). Automatic facial expression analysis: A survey, *PATTERN RECOGNITION* 36(1): 259–275.
- FG-NET AGING DATABASE (n.d.). <http://www.fgnet.rsunit.com/>.
- Hadid, A. & Pietikainen, M. (2009). Manifold learning for gender classification from face sequences, *Advances in Biometrics*, Springer Verlag., pp. 82–91.
- Kanade, T., Cohn, J. & Tian, Y. (2000). Comprehensive database for facial expression analysis, *Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition* pp. 46–53.
- Li, S. Z. & Jain, A. K. (2005). *Handbook of Face Recognition*, Springer.
- Maat, L., Sondak, R., Valstar, M., Pantic, M. & Gaia, P. (2009). Man machine interaction (mmi) database.
- Mayer, C., Wimmer, M., Eggers, M. & Radig, B. (2009). Facial expressions recognition with 3d deformable models, *2009 Second International Conferences on Advances in Computer-Human Interactions*, IEEE Computer Society, pp. 26–31.
- O’Toole, A. J. (2009). Cognitive and computational approaches to face recognition, *The University of Texas at Dallas*.
- Park, U. & Jain, A. K. (2007). 3d model-based face recognition in video, *2nd International Conference on Biometrics*.
- Riaz, Z., Gedikli, S., Beetz, M. & Radig, B. (2010). 3d face modeling for multi-feature extraction for intelligent systems, *Computer Vision for Multimedia Applications: Methods and Solutions 1*: 73–89.
- Riaz, Z., Mayer, C., Beetz, M. & Radig, B. (2009a). Face recognition using wireframe model across facial expressions, *Proceedings of the 2009 joint COST 2101 and 2102 international conference on Biometric ID management and multimodal communication, BioID MultiComm’09*, pp. 122–129.
- Riaz, Z., Mayer, C., Beetz, M. & Radig, B. (2009b). Model based analysis of face images for facial feature extraction, *Computer Analysis of Images and Pattern* pp. 99–106.
- Riaz, Z., Mayer, C., Wimmer, M., Beetz, M. & Radig, B. (2009). A model based approach for expression invariant face recognition, *International Conference on Biometrics*, Springer.
- Shan, C., Gong, S. & McOwan, P. (2009). Facial expression recognition based on local binary patterns: A comprehensive study, *Computer Vision and Image Understanding*, Elsevier Inc., pp. 803–816.
- Sinha, P., Balas, B., Ostrovsky, Y. & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about, *Proceedings of IEEE* 94(II): 1948–1962.
- Sirovich, L. & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces, *J. Opt. Soc.* 4(3): 519–524.
- Terence, S., Baker, S. & Bsat, M. (2002). The cmu pose, illumination, and expression (pie) database, *FGR ’02: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, IEEE Computer Society, Washington, DC, USA, p. 53.
- Turk, M. & Pentland, A. (1991). Face recognition using eigenfaces, *IEEE Computer Society Conference on In Computer Vision and Pattern Recognition, 1991. Proceedings CVPR ’91* pp. 586–591.
- Viola, P. & Jones, M. J. (2004). Robust real-time face detection, *International Journal of Computer Vision* 57(2): 137–154.

- Wimmer, M., Riaz, Z., Mayer, C. & Radig, B. (2009). Recognizing facial expressions using model-based image interpretation, *Advances in Human-Computer Interaction*, I-Tech.
- Wimmer, M., Stulp, F., Pietzsch, S. & Radig, B. (2008). Learning local objective functions for robust face model fitting, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 30(8): 1357–1370.
- Wimmer, M., Stulp, F., Tschechne, S. & Radig, B. (2006). Learning robust objective functions for model fitting in image understanding applications, *Proceedings of the 17th British Machine Vision Conference* pp. 1159–1168.
- Witten, I. H. & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann, San Francisco.
- Zhao, W. & Chellappa, R. (2005). *Face Processing: Advanced Modeling and Methods*, Elsevier.
- Zhao, W., Chellappa, R., Phillips, P. J. & Rosenfeld, A. (2003). Face recognition: A literature survey.



# Digital Signature: A Novel Adaptative Image Segmentation Approach

David Freire-Obregón<sup>1</sup>, Modesto Castrillón-Santana<sup>1</sup>  
and Oscar Déniz-Suárez<sup>2</sup>

<sup>1</sup>SIANI, Universidad de Las Palmas de Gran Canaria

<sup>2</sup>VISILAB, Universidad de Castilla-La Mancha  
Spain

## 1. Introduction

Partitioning or segmenting an entire image into distinct recognizable regions is a central challenge in computer vision which has received increasing attention in recent years.

It is remarkable how a simple task to humans is not so simple in computer vision, but it can be explained. To humans, an image is not just a random collection of pixels; it is a meaningful arrangement of regions and objects (Figure 1 shows a variety of images).

Despite the variations of these images, humans have no problem interpreting them. We can agree about the different regions in the images and recognize the different objects. Human visual grouping was studied extensively by the Gestalt psychologists. Thus, Wertheimer pointed out the importance of perceptual grouping and organization in vision and listed several key factors Wertheimer (1938), that lead to human perceptual grouping: similarity,



Fig. 1. Some challenging images for a segmentation algorithm. Our goal is to develop a single grouping procedure which can deal with all these types of images. Image Source: Sample images in MIT/CMU test set for frontal face detection (Sung & Poggio (1999)).

proximity, continuity, symmetry, parallelism, closure and familiarity. In computer vision, these factors have been used as guidelines for many grouping algorithms. Thus, the most studied version of grouping in computer vision is image segmentation. Image segmentation techniques can be classified into two families: (1) region-based, and (2) contour-based approaches.

- Region-based approaches try to find partitions of the image pixels into sets corresponding to coherent image properties such as brightness, color and texture.
- Contour-based approaches usually start with a first stage of edge detection, followed by a linking process that seeks to exploit curvilinear continuity.

On the other hand, in order to distinguish good segmentations from bad segmentations the, already mentioned, classical Gestalt theory has developed various principles of grouping (Palmer (1999); Wertheimer (1938)) such as proximity, similarity and good continuation. As Ren & Malik (2003) pointed out, the principle of good continuation states that a good segmentation should have smooth

1. Intra-region similarity: the elements in a region are similar. This includes similar brightness, similar texture, and low contour energy inside the region;
2. inter-region (dis)similarity: the elements in different regions are dissimilar. This in turn includes dissimilar brightness, dissimilar texture, and high contour energy on region boundaries.

These classical principles of grouping have inspired many previous approaches to segmentation. However, the Gestalt principles distinguish competing segmentations only when everything else is equal. Many of the previous works have made ad-hoc decisions for using and combining these cues. However, even to this day, many of the computational issues of perceptual grouping have remained unresolved.

In this work, we present an image descriptor based on self-similarities which is able to capture the general structure of an image. Computed descriptors are similar for images with the same layout, even if textures and colors are different. Similarly to Shechtman & Irani (2007), images are partitioned into smaller cells which, conveniently compared with a patch located at the image center, yield a vector of values that describes local aspect correspondences. In this Chapter, we demonstrate the effectiveness of our approach on two main topics in computer vision's research: facial expression recognition and hair/skin segmentation.

The outline of the Chapter is as follows. Section 2 reviews the previous segmentation approaches. The theoretical concept of digital signature and its main features will be discussed in Section 3 at a high level. In Section 4, some experimental results for classification and segmentation process in order to illustrate the effectiveness and usefulness of the proposed approach. The conclusions of this work in Section 5 will conclude this Chapter.

## 2. Related work

Many image segmentation methods have been proposed over the last several decades. As new segmentation methods have been proposed, a variety of evaluation methods have been used to compare new segmentation methods to prior methods. These methods are fundamentally very different, and can be partitioned based on the diversity in segment types:

- Algorithms for extracting uniformly colored regions (Comaniciu & Meer (2002); Shi & Malik (2000)).

- Algorithms for extracting textured regions (Galun et al. (2003); Malik et al. (1999)).
- Algorithms for extracting regions with a distinct empirical color distribution (Kadir & Brady (2003); Li et al. (2004); Rother et al. (2004)).

Thus, most of the methods require using image features that characterize the regions to be segmented. Particularly, texture and color have been independently and extensively used in the area. On the other hand, some algorithms can be also categorized on unsupervised (Shi & Malik (2000)), while others require user interaction (Rother et al. (2004)). Some algorithms employ symmetry cues for image segmentation (Riklin-Raviv et al. (2006)), while others use high-level semantic cues provided by object classes (i.e., class-based segmentation, see Borenstein & Ullman (2002); Leibe & Schiele (2003); Levin & Weiss (2006)).

There are also variants in the segmentation tasks, ranging from segmentation of a single input image, through simultaneous segmentation of a pair of images (Rother et al. (2006)) or multiple images.

Bagon et al. (2008) proposed a single unified approach to define and extract visually meaningful image segments, without any explicit modelling. Their approach defines a 'good image segment' as one which is 'easy to compose' (like a puzzle) using its own parts. It captures a wide range of segment types: uniformly colored segments, through textured segments, and even complex objects.

Shechtman & Irani (2007) proposed a segmentation approach based on 'local self-similarity descriptor'. It captures self-similarity of color, edges, repetitive patterns and complex textures in a single unified way. These self-similarity descriptors are estimated on a dense grid of points in image/video data, at multiple scales. Moreover, our Digital Signature is based in this work.

### 3. Digital signature

We present an image descriptor based on self-similarities which is able to capture the general structure of an image. Computed descriptors are similar for images with the same layout, even if textures and colors are different, similarly to Shechtman & Irani (2007). Images are partitioned into smaller cells which, conveniently compared with a main patch located in the image, yield a vector of comparison results that describes local aspect correspondences.

The Digital Signature (DS) descriptor is computed from a square image subdivided into  $n \times n$  cells, where each cell corresponds to an  $m \times m$  pixels image patch. The number of cells and their pixel size have effect on how much an image structure is generalized. A low number of cells will not capture many structural details, while too many small cells will produce a too detailed descriptor. The present approach will consider overlapping cells, which may be required to capture subtle structural details.

Once an image is partitioned, an  $m \times m$  main patch located in the image (which does not have to correspond to a cell in the image partition) is compared with all partition cells. In order to achieve greater generalization, image patches are compared computing the Sum of Squared Differences (SSD) between pixel values (or the Sum of Absolute Differences (SAD), which is computationally less expensive). Each cell-center comparison is consecutively stored in a  $m \times m$  dimensions DS descriptor vector.

In a more general mathematical sense, a Digital Signature  $z$  for a proposed cell  $p$ , is a function that makes a cell comparative that fall into each of the disjoint categories (similar to histogram's bins), whereas the graph of a digital signature is merely one way to represent a digital signature descriptor. Thus, if we let  $ncells$  be the total number of cells,  $Wcsize$  be the

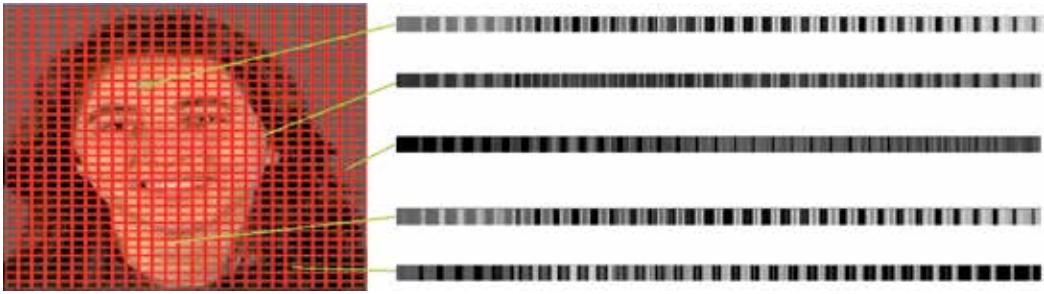


Fig. 2. Digital Signature Descriptor example using a 11x11 partition. The barcode-like vector represents all comparisons between each cell and the other cells. Image source: DaFEx Database (Battocchi & Pianesi (2004)).

width for each cell and  $Hcsize$  be the height for each cell, the descriptor meets the following conditions for each channel:

$$Digital\_Signature_z = \sum_p^{ncells} \left| \int_i^{Wcsize} \int_j^{Hcsize} cell_z(i,j) - \int_u^{Wcsize} \int_v^{Hcsize} cell_p(u,v) \right| \quad (1)$$

Such description overcomes color, contrast and textures. Images are described in terms of their general structure, similarly to Shechtman & Irani (2007). An image showing a white upper half and a black lower half will produce exactly the same descriptor as an image showing a black upper half and a white lower half. Local aspect correspondences are exactly the same: the upper half is different from the lower half. Rotations, however, are not considered.

Digital Signature descriptors are specially useful to describe points defined by a scale salient point detector (like DoG or SURF (Bay & Tuytelaars (2006))). The DS descriptor is shown as a barcode for representation purposes.

Thus, given an input image (i.e. a scale salient point or a known region like detected mouths), a number of cells  $n$  and their pixel size  $m$ , DS is computed as follows:

1. The image is divided into different cells inside a template sized  $(n \times m) \times (n \times m)$  pixels.
2. The template is partitioned into  $n \times n$  cells, each of them sized  $m \times m$  pixels.
3. Each cell is compared with any other template cell, and each result is consecutively stored in the  $n \times n$  DS descriptor vector. Thus, each cell provides its own Digital Signature.

In order to illustrate the effectiveness and usefulness of the proposed approach we present two different tests. For the first test, similar Digital Signatures from different images are applied to classify mouths into smiling or non-smiling gestures, for the second test, different Digital Signatures from the same image are compared in order to obtain a good image segmentation (See Figure 2).

## 4. Experimental results

### 4.1 Facial expression recognition

After extensive research in Psychology, it is now known that emotions play a significant role in human decision making processes (Damasio (1994); Picard (1997)). The ability to show and interpret them is therefore also important for human-machine interaction. Perceptual User Interfaces (PUIs) use multiple input modalities to capitalize on all the communication cues, thus maximizing the bandwidth of communication between a user and a computer.

Examples of PUIs have included: assistants for the disabled, augmented reality, interactive entertainment, virtual environments, intelligent kiosks, etc.

Some facial expressions can be very subtle and difficult to recognize even between humans. Besides, in human-computer interaction the range of expressions displayed is typically reduced. In front of a computer, for example, the subjects rarely display accentuated surprise or anger expressions as he/she could display when interacting with another human subject.

The human smile is a distinct facial configuration that could be recognized by a computer with greater precision and robustness. Besides, it is a significantly useful facial expression, as it allows to sense happiness or enjoyment and even approval (and also the lack of them) (Ekman & Friesen (1982)). As opposed to facial expression recognition, smile detection research has produced less literature. Lip edge features and a perceptron were used in Ito et al. (2005). The lip zone is obviously the most important, since human smiles involve mainly the Zygomatic muscle pair, which raises the mouth ends. Edge features alone, however, may be insufficient. Smile detection was also tackled in the BROAFERENCE system to assess TV or multimedia content (Kowalik et al. (2005)). This system was based on tracking the positions of a number of mouth points and using them as features feeding a neural network classifier. The work Shinohara & Otsu (2004), in turn, used Higher-order Local Autocorrelation, achieving near 98% recognition rates. More recently, the comprehensive work Whitehill et al. (2008) contends that there is a large performance gap between typical tests made in the literature and results obtained in real-life conditions. The authors conclude that the training set may be all-important, specially in terms of variability and size, which should be on the order of thousands of images.

The Sony Cybershot DSC T-200 digital camera has an ingenious "smile shutter" mode. Using proprietary algorithms, the camera automatically detects the smiling face and closes the shutter. To detect the different degrees of smiles by the subject, smile detection sensitivity can be set to high, medium or low. Some reviews argue that: *"the technology is not still so much sensitive that it can capture minor facial changes. Your facial expression has to change considerably for the camera to realize that"* (Entertainment.millionface.com: *Smile detection technology in camera* (2008)), or *"The camera's smile detection - which is one of its more novel features - is reported to be inaccurate and touchy"* (Swik.net: *Sonys Cyber-shot T200 gets its first review* (2008)). Whatever the case, detection rates or details of the algorithm are not available, and so it is difficult to compare results with this system. Canon also has a similar smile detection system. This section describes different techniques that are applied to the smile detection problem in video streams and shows the benefits of our new approach.

#### 4.1.1 Representation

In order to show the performance of our new approach, it will be compared with other image descriptors such as Local Binary Patterns and Principal Components Analysis. The Local Binary Pattern (LBP) is an image descriptor commonly used for classification and retrieval. Introduced by Ojala et al. (2002) for texture classification, they are characterized by invariance to monotonic changes in illumination and low processing cost.

Given a pixel, the LBP operator thresholds the circular neighborhood within a distance by the pixel gray value, and labels the center pixel considering the result as a binary pattern. The basic version considers the pixel as the center of a  $3 \times 3$  window and builds the binary pattern based on the eight neighbors of the center pixel, as shown in Figure 3. However, the LBP definition can be easily extended to any radius,  $R$ , considering  $P$  neighbors Ojala et al. (2002):

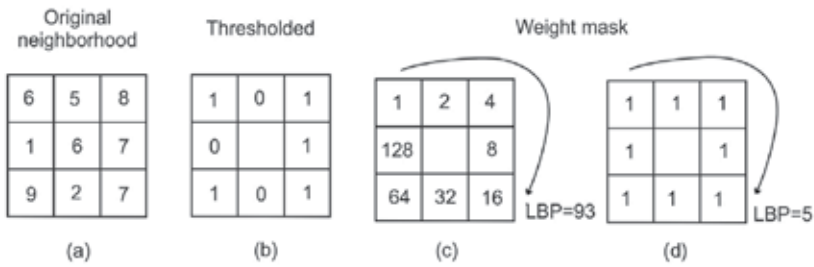


Fig. 3. The basic version of the Local Binary Pattern computation (c) and the Simplified LBP codification (d).

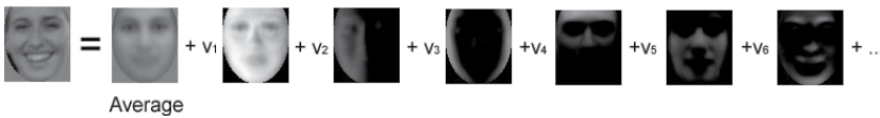


Fig. 4. The Principal Components Analysis (PCA) decomposition.

$$LBP_{P,R}(x_c, y_c) = \sum_{k=0}^{P-1} s(g_p - g_c) 2^k, \quad s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

Rotation invariance is achieved in the LBP based representation considering the local binary pattern as circular. The experience achieved by Ojala et al. (2002) suggested that just a particular subset of local binary patterns are typically present in most of the pixels contained in real images. They refer to these patterns as uniform. Uniform patterns are characterized by the fact that they contain, at most, two bitwise transitions from 0 to 1 or viceversa. For example, 00000000, 00011100 and 10000011 are uniform patterns. In the experiments carried out by Ojala et al. (2002) with texture images, uniform patterns account for a bit less than 90% of all patterns when using the 3x3 neighborhood.

More recently LBPs have been used to describe facial appearance. Once the LBP image is obtained, most authors apply a histogram based representation approach (Sébastien Marcel & Heusch (2007)). However, as pointed out by some recent works, the histogram based representation loses relative location information (Sébastien Marcel & Heusch (2007); Tao & Veldhuis (2007)), thus LBP can also be used as a preprocessing method. Using LBP as preprocessing method, having the effect of emphasizing edges and noise. To reduce the noise influence, Tao & Veldhuis (2007) proposed recently a modification in Equation 2. Instead of weighting the neighbors differently, their weights are all the same, obtaining the so called Simplified LBPs, see Figure 3-d. Their approach has shown some benefits applied to facial verification, due to the fact that by simplifying the weights, the image becomes more robust to illumination changes, having a maximum of nine different values per pixel. The total number of local patterns are largely reduced so the image has a more constrained value domain.

$$LBP_{P,R}(x_c, y_c) = \sum_{k=0}^{P-1} s(g_p - g_c), \quad s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3)$$

In the experiments described below, both approaches will be investigated, i.e. using the histogram based approach, but also using Uniform LBP and Simplified LBP as a preprocessing step.

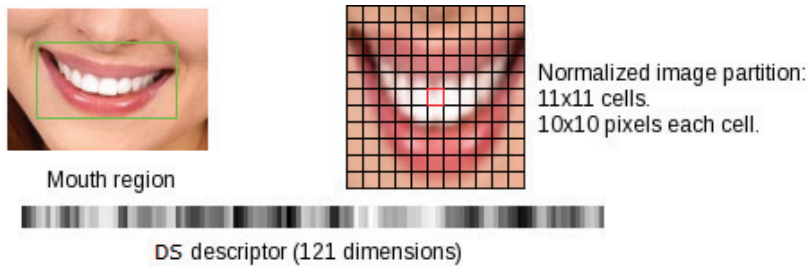


Fig. 5. DS Descriptor example using a 11x11 partition. The barcode-like vector represents all comparisons between each cell and the central patch. In this case, only the center patch will be considered to obtain our descriptor.

Raw face images are highly dimensional. A classical technique applied for face representation to avoid the consequent processing overload problem is Principal Components Analysis (PCA) decomposition (Kirby & Sirovich (1990)). PCA decomposition is a method that reduces data dimensionality, without a significant loss of information, by performing a covariance analysis between factors. As such, it is suitable for highly dimensional data sets, such as face images. A normalized image of the target object, i.e. a face, is projected in the PCA space, see Figure 4. The appearance of the different individuals is then represented in a space of lower dimensionality by means of a number of those resulting coefficients,  $v_i$  (Turk & Pentland (1991)).

We now discuss the central contribution of this section, the incorporation of the Digital Signature descriptor. However, images containing smiling mouths require local brightness to be preserved: teeth are always brighter than surrounding skin and that must be captured by the descriptor. Thus, instead of using SSD, patches are compared using Sum of Differences. Otherwise, a closed mouth would produce the same descriptor as a smiling mouth: lips are surrounded by differently colored skin, exactly as teeth are surrounded by differently colored lips. Figure 5 shows an example with an  $11 \times 11$  cell partition, each cell sized  $10 \times 10$  pixels. For this experiment, the DS is applied to classify mouths found by a face detector (Castrillón Santana et al. (2007)) into smiling or non-smiling gestures. Smiling mouths look similar no matter the skin color or the presence of facial hair. This generality can be registered by a self-similarity descriptor like DS.

Thus, for this test, our Digital Signature descriptor is computed as follows:

1. The image is resized to a template sized  $(n \times m) \times (n \times m)$  pixels.
2. The template is partitioned into  $n \times n$  cells, each of them sized  $m \times m$  pixels.
3. A main cell sized  $m \times m$  pixels is captured from the template image.
4. The main cell is compared with each template cell, and each result is consecutively stored in the  $n \times n$  DS descriptor vector.

#### 4.1.2 Classification and face detection

Two more points are needed to be justified. The first one is the classification method. In order to tell whether two images (smiling or not smiling) have a similar structure, their corresponding DS descriptors can be compared computing SAD between both vectors. However, given that the present work aims at classifying mouth images in two categories, a Support Vector Machine approach (Burges (1998)) is used. The SVM is a set of related supervised learning

methods used for classification and regression. They belong to a family of generalized linear classifiers. A property of SVMs is that they simultaneously minimize the empirical classification error and maximize the geometric margin; hence they are also known as maximum margin classifiers.

The second point is to explain how are faces going to be extracted from videos. Several approaches have recently appeared for real-time face detection (Schneiderman & Kanade (2000); Viola & Jones (2004)), all of them aiming at making the problem less environment dependent. Focusing on live video stream processing, a face detector based on cue combination (Castrillón Santana et al. (2007)) outperforms well known single-cue based detectors such as Viola & Jones (2004). This approach provides a more reliable tool for real time interaction in PUIs. The face detection system used to extract faces from video streams in this work (see Castrillón Santana et al. (2007) for details) integrates, among other cues, different classifiers based on the general object detection framework by Viola and Jones Viola & Jones (2004), skin color, multilevel tracking, etc. In order to further minimize the influence of false alarms, the facial feature detector capabilities were extended, locating not only faces but also eyes, nose and mouth. This reduces the number of false alarms, for it is less probable that multiple detectors, i.e. face and its inner features, are activated simultaneously with a false alarm. Its important to point that each detected face is normalized to  $59 \times 65$  pixels using the position of the eyes.



Fig. 6. Facial element detection results for a video stream extracted from the DaFEx Database.

The facial element detection procedure is only applied in those areas which bear evidence of containing a face. This is true for regions in the current frame, where a face has been detected, or in areas with a detected face in the previous frame. Figure 6 shows the possibilities of the face detector.

#### 4.1.3 Performance and evaluation

The DaFEx (Battocchi & Pianesi (2004)) database was used in this experiment. In this database 8 professional actors showed 7 expressions (6 basic facial expressions + 1 neutral) on 3 intensity levels (low, medium, high) twice. The frames of the 48 'happy' videos were extracted of the database sequences (see Figure 7) for test. The total number of images contained in the 'happy' videos of dataset is 12,988, from which, 6,783 are smiling images and 6,205 are no smiling images. These images have been also normalized according to eye positions obtaining  $59 \times 65$  samples and have been annotated manually by a human.

As briefly mentioned above, the experimental setup considers one possibility as input: the mouth. The input image is a grayscale image, and for representation purposes we have used the following approaches for the tests:

- PCA. A PCA space obtained from the original gray images of  $59 \times 65$  pixels.
- ULBP Hist. A concatenation of histograms based on the gray image or the resulting ULBP image.





Fig. 7. Low, medium and high intensity happy expressions from DaFEx

| DaFEx             | Approach          | False Negative | False Positive | Error Rate |
|-------------------|-------------------|----------------|----------------|------------|
| Complete Database | PCA               | 11.1%          | 16.3%          | 13.5%      |
|                   | ULBP Hist.        | 19.5%          | 23.6%          | 21.4%      |
|                   | SLBP Im. Val.     | 20.2%          | 21.3%          | 20.7%      |
|                   | Digital Signature | 10.2%          | 13.0%          | 11.5%      |
| LOW Intensity     | PCA               | 13.6%          | 19.1%          | 16.2%      |
|                   | ULBP Hist.        | 24.9%          | 24.0%          | 24.4%      |
|                   | SLBP Im. Val.     | 22.1%          | 23.8%          | 22.9%      |
|                   | Digital Signature | 12.3%          | 16.6%          | 14.3%      |
| MED Intensity     | PCA               | 13.6%          | 14.4%          | 13.9%      |
|                   | ULBP Hist.        | 19.9%          | 22.5%          | 21.1%      |
|                   | SLBP Im. Val.     | 20.9%          | 20.4%          | 20.6%      |
|                   | Digital Signature | 11.6%          | 13.7%          | 12.6%      |
| HIGH Intensity    | PCA               | 6.3%           | 15.5%          | 10.6%      |
|                   | ULBP Hist.        | 13.7%          | 24.5%          | 18.8%      |
|                   | SLBP Im. Val.     | 17.6%          | 19.7%          | 18.6%      |
|                   | Digital Signature | 6.8%           | 8.7%           | 7.7%       |

Table 1. Error rates achieved using each approach over the DaFEx database, considering the complete database or dividing it into different intensities (see text). Four different approaches were applied: PCA on grayscale images considering 130 coefficients, Normalized Histogram method on the Uniform LBP representation, Normalized Image Values method on the Simplified LBP representation and the Digital Signature on grayscale images.

- SLBP Im. Val. A concatenation of the image values based on the gray images or the resulting SLBP image.
- Digital Signature. DS descriptor obtained from the original gray images of  $59 \times 65$  pixels.

Similar experimental conditions have been used for every approach considered in this setup. The test sets are built randomly, having an identical number of images for both classes. Results presented correspond to the percentage of wrong classified samples of all test samples.

As it can be appreciated in Table 1, best results in almost every situation are achieved with our new approach. None of the LBP based representations outperforms that approach. However, even if the Uniform LBP approach evidences a larger improvement when normalized histograms are used, the Simplified LBP approach reported better results than Uniform LBP. As already stated in Tao & Veldhuis (2007) this preprocessing provides benefits in the context of facial analysis.

The PCA based representation achieves a better error rate than the LBP's approaches. However, it doesn't keep a good balance between false positive and false negative. PCA

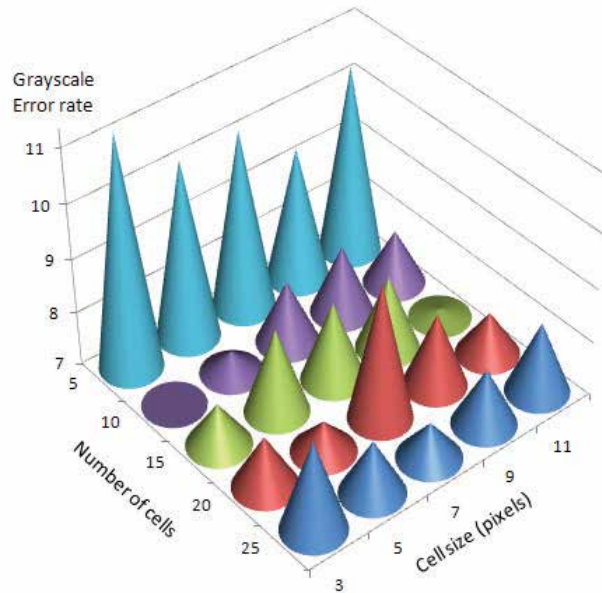


Fig. 8. Results achieved by the Digital Signature approach for the high intensity test. As it can be appreciated, best results are achieved for  $n = 10$  and  $m = 3$ .

deserves an additional observation, not always the increasing of the space dimension for PCA reports better results. This is the main reason for keeping only 130 coefficients for the image descriptor. When the DS descriptor is used, the test achieved the lowest error rate.

For the Digital Signature approach, error-rate behavior is also quite similar to behaviour obtained previously with PCA and Image Value tests rates. Again, the lowest error rates were achieved by the grayscale image test. For DS, it is important to mention that overlap is not considered between cells. Firstly, several tests without overlapping were made in order to find optimum DS parameters (number of cells and cells' size yielding the lowest error rate). For smile detection it was found that  $10 \times 10$  cells of  $3 \times 3$  pixels performed best (See Figure 8) thanks to the closeness between the size of the extracted DS main cell ( $30 \times 30$  pixels) and the original size of the mouth capture ( $20 \times 12$  pixels). It is also shown that worst results are achieved for configurations less than  $10 \times 10$  cells because of the loss of information due to resizing in the Normalization step. Beyond that number of cells and for bigger sizes, behaviour is irregular due to the fact that information extracted is not reliable because of the false information introduced when the mouth is resized to fit the DS cell. When images are upsampled, redundant and useless information is created. Unfortunately, when overlap was introduced, the achieved error rates were higher than without overlapping. Used images were too small for overlapping regions to be significant. Unlikely what happens with the PCA approach in this case there is a good balance between false positive and false negative. Thus, the new approach has got better results than other approaches that have been successfully used before for smile detection (Freire et al. (2009a;b)).

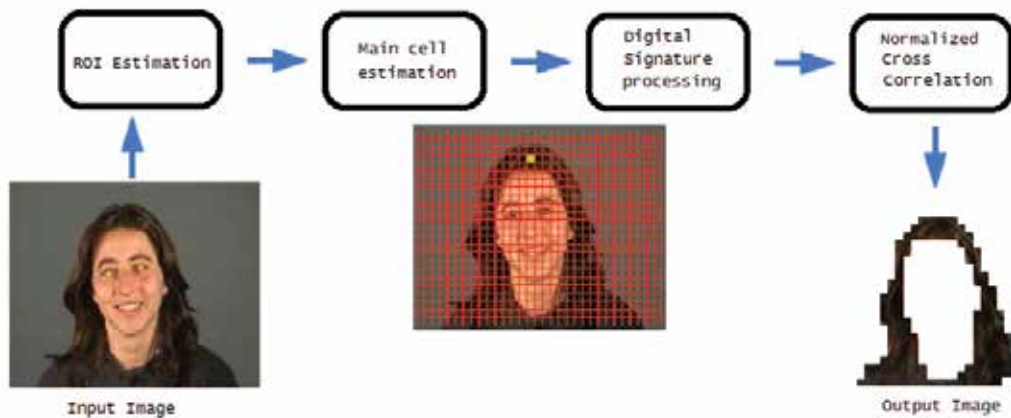


Fig. 9. Block diagram of our segmentation method: (a) input image, (b) Region of interest estimation, (c) main cell estimation (highlighted in green), (d) Digital Signature processing for each cell, (e) Normalized Cross Correlation between each cell's DS and the main cell's DS, and (f) the hair extracted by the proposed method.

#### 4.2 Image segmentation

Face recognition techniques have attracted much attention over the years and many algorithms have been developed. Since it has many potential applications in computer vision and automatic access control system, its research has rapidly expanded by not only engineers but also neuroscientists. Especially, face segmentation is an essential step of face recognition system since most face classification techniques tend to only work with face images. Therefore face segmentation has to correctly extract only face part of given large image. However, apart from facial features, hair style or clothing also reflects important personal traits.

Clothing segmentation is widely used in many computer vision tasks, such as dressed people detection (Ioffe & Forsyth (2001); Ronfard et al. (2002)), identification, image editing, human sketches and portraits for graphics rendering (Chen et al. (2004; 2006)). However, existing segmentation methods suffer from variations in colors and styles, different lighting conditions, cluttering backgrounds and occlusions generated by poses or other objects. Most existing methods use clothing models (Chen et al. (2006); Lee & Cohen (2006); Sprague & Luo (2002)). Typically, clothing models are trained with tagged samples, and then the clothing is extracted from images by comparing to the models. For example, in Sprague & Luo (2002), a dress model and a shirt model were trained, respectively, and then clothing detection was performed on the pre-segmented image by selecting a better match against the trained models. Such methods can only handle a few clothing styles and rely on the pre-segmentation accuracy.

On the other hand, human hair related applications have attracted increasing interest in recent years, since hair plays a significant role in the overall appearance of an individual. To achieve these tasks, hair segmentation is generally the first prerequisite step. However, to our knowledge, in most previous studies, hair is assumed segmented already or manually labeled. Furthermore, besides the above hair-related applications, many computer vision tasks can also benefit from segmented hair. For instance, it provides an important clue for gender classification, since hair styles of male and female are generally different. Hair can also facilitate age estimation since hair distribution and color gradually changes with the increase

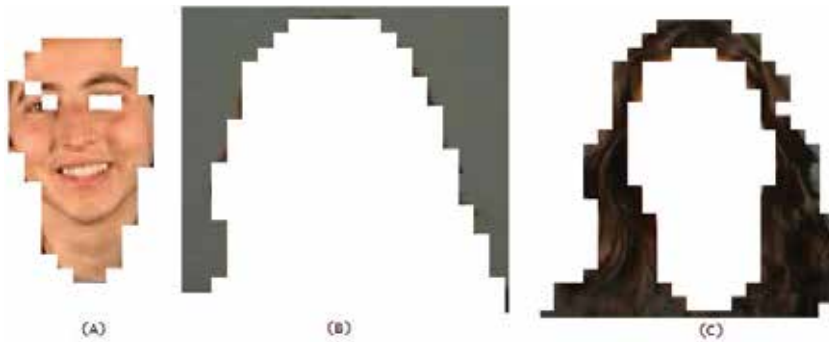


Fig. 10. Segmentation results on this frame image sequence are shown in subimages (a) to (c). Segment in (a) correspond to the person's skin, segment in (c) correspond to the person's hair and (b) correspond to the background. This results are achieved for  $n = 35$  and  $m = 11$ .

of age, especially for old men/women. Hair information can even contribute a lot to face recognition, considering that the hair style of one normal subject does not abruptly change frequently. To sum up, more attention should be paid to automatic hair segmentation. The work of Yacoob and Davis Yacoob & Davis (2006) is the only prior work we can find on hair detection. Their approach constructs a simple color model and uses it to recognize the hair pixel. However, their detection can only work under controlled background environment and very less hair color variation. On the other hand, hair modelling, synthesis, and animation have already become active research topics in computer graphics (Kajiya & Kay (1989); Marschner et al. (2003); Moon & Marschner (2006); Paris et al. (2004); Wei et al. (2005)). In this section, without any predefined model, we propose a new method to segment parts of an image (e.g. clothing, hair style or skin) by using the Digital Signature technique.

#### 4.2.1 The proposed method

Thus, given an input image, a number of cells  $n$  and their pixel size  $m$ , the proposed method is computed as follows (See Figure 9):

1. The facial detector described in Section 4.1.3 provides the eyes position. Thus, a region of interest can be estimated.
2. The image is divided into different cells inside a template sized  $(n \times m) \times (n \times m)$  pixels.
3. The template is partitioned into  $n \times n$  cells, each of them sized  $m \times m$  pixels.
4. Taking the eye positions into account, it is possible to estimate a main cell position.
5. Each cell is compared with any other template cell, and each result is consecutively stored in the  $n \times n$  DS descriptor vector. Thus, each cell provides its own Digital Signature (See Figure 2).
6. Each cell's DS will be compared to the main cell's DS. This will help to decide whether or not the proposed cell is similar to the main cell.

#### 4.2.2 Results

Figure 10 displays a segmentation results with a simple color frame. The Digital Signature succeeded to generate accurate skin and background segmentation. It fails to generate accurate hair segmentation because the hair and cloth have similar color in the example. It can be also appreciated an spatial aliasing effect due to the fact that the cell's shape is square.

One way to reduce the spatial aliasing is to decrease the cell's size but, this fact could be harmful for the method because it will cause an information loss for each cell.

RGB color images were considered here for test. In each segmentation, the value of the one free parameter, cell size in Equation 1, was kept constant:  $11 \times 11$ , despite the different characteristics of the images. Figure 11 shows the results of applying the segmentation algorithm to two more images under unconstrained lighting environment. The overall result of this study was that the segmentation results were generally stable to perturbations of the cell size as far as the distance between eyes remains similar between images.



Fig. 11. Some components of the partition applying a Digital Signature cell size of  $11 \times 11$  ( $n = 11$ ). Images source: askmen.com(Askmen.com (2011)).

## 5. Conclusions

In this Chapter, we have introduced a novel descriptor for segmenting an image into a regular grid of cells. We have argued that the regular grid confers a number of useful properties, such as a powerful representation approach for classification purposes. We have also demonstrated that despite this topological constraint, we can achieve segmentation performance comparable with current algorithms.

For the first test, this Chapter described a smile detection using different LBP approaches, as well as PCA image representation, combined with SVM. The potentiality of the Digital Signature based representation for smile verification has been shown. The DS based representation presented in this paper outperforms other approaches with an improvement over a 5%. Uniform LBP does not respond to a statistical spatial patterns locality. This means, that there is no gradual change between adjacent blocks preprocessed with Uniform LBP. Depending on the value of a pixel inside one of the blocks, codification between two adjacent pixels can be, for example, from pattern 2 to pattern 9. Translated to the space domain of SVM, this means that dimensions can be too far away. Simplified LBP keeps the statistical spatial patterns locality. There is a gradual change between adjacent preprocessed pixels. Translated to the SVM's space domain, this gradual change means that similar points are closer in this space. Our future line is focus on the potentiality of the DS descriptor for generic applications such as image retrieval. In this paper we have developed a static smile classifier achieving, in

some cases, a 93% of success rate. Due to this success rate, smile detection in video streams, where temporal coherence is implicit, will be studied in short term, as a cue to get the ability to recognize the dynamics of the smile expression.

For the second test, we have analyzed how the Digital Signature cell's descriptors can be used to segment an input image without any previous training. Once the eyes position is obtained, the template of the main cell can be found inside the image by comparing their digital signatures. The model works very well under certain circumstances:

1. There is a variety of patterns between the main cell and the rest of cells.
2. Under constraint and unconstraint lighting environment.
3. The face/hair/clothes features are relatively small to fill holes (See false negatives for the first dress in Figure 11)

The above insight opens the door for further research into the balance between complexity of the model and the way in which it is used for inference. In order to reduce the aliasing effect other geometrical shape for the cells can be also tested. On the other hand the cell's size depends on image cues such as the image's size or other heuristic method depending on the segmentation process (i.e. for face segmentation, the distance between the eyes could be an important cue for the cell size).

## 6. Acknowledgments

We would like to thank Dr. Thomas B. Moeslund and Dr. Nasrollahi for their advice, criticism and encouragement during the development of this algorithm. This work has been partially supported by the Research Project TIN2008-06068, funded by the Ministerio de Ciencia y Educación (Government of Spain).

## 7. References

- Askmen.com* (2011).
- Bagon, S., Boiman, O. & Irani, M. (2008). What is a good image segment?, *ECCV*.
- Battocchi, A. & Pianesi, F. (2004). Dafex: Un database di espressioni facciali dinamiche, *SLI-GSCP Workshop Comunicazione Parlata e Manifestazione delle Emozioni*.
- Bay, H. & Tuytelaars, T. (2006). Surf: Speeded up robust features, *Proceedings of the Ninth European Conference on Computer Vision*.
- Borenstein, E. & Ullman, S. (2002). Class-specific, top-down segmentation, *ECCV*.
- Burges, C. (1998). A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery* 2(2): 121–167.
- Castrillón Santana, M., Déniz Suárez, O., Hernández Tejera, M. & Guerra Artal, C. (2007). ENCARA2: Real-time detection of multiple faces at different resolutions in video streams, *Journal of Visual Communication and Image Representation* pp. 130–140.
- Chen, H., Liu, Z., Rose, C. & Xu, Y. (2004). Example-based composite sketching of human portraits, *Proceedings of Third International Symposium on Non Photorealistic Animation and Rendering (NPAR)*, pp. 95–102.
- Chen, H., Liu, Z., Zhu, S. & Xu, Y. (2006). Composite templates for cloth modeling and sketching, *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 943–950.
- Comaniciu, D. & Meer, P. (2002). Mean shift: a robust approach toward feature space analysis, *PAMI*.

- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*, Picador.
- Ekman, P. & Friesen, W. (1982). Felt, false, and miserable smiles, *Journal of Nonverbal Behavior* 6(4): 238–252.
- Entertainment.millionface.com: *Smile detection technology in camera* (2008).
- Freire, D., Castrillón Santana, M. & Déniz Suárez, O. (2009a). A novel approach for smile detection combining ulbp and pca, *Proceedings of the EUROCAST*.
- Freire, D., Castrillón Santana, M. & Déniz Suárez, O. (2009b). Smile detection using local binary patterns and support vector machines, *Proceedings of the VISIGRAPP*.
- Galun, M., Sharon, E., Basri, R. & Brand, A. (2003). Texture segmentation by multiscale aggregation of filter responses and shape elements, *ICCV*.
- Ioffe, S. & Forsyth, D. (2001). Probabilistic methods for finding people, *International Journal of Computer Vision* 43(1): 45–68.
- Ito, A., Wang, X., Suzuki, M. & Makino, S. (2005). Smile and laughter recognition using speech processing and face recognition from conversation video, *Proc. of the 2005 IEEE Int. Conf. on Cyberworlds (CW'05)*.
- Kadir, T. & Brady, M. (2003). Unsupervised non-parametric region segmentation using level sets, *ICCV*.
- Kajiya, J. & Kay, T. (1989). Rendering fur with three dimensional textures, *Computer Graphics* 23(3): 271–280.
- Kirby, Y. & Sirovich, L. (1990). Application of the Karhunen-Loève procedure for the characterization of human faces, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 12(1).
- Kowalik, U., Aoki, T. & Yasuda, H. (2005). Broaference - a next generation multimedia terminal providing direct feedback on audience's satisfaction level, *INTERACT*, pp. 974–977.
- Lee, M. & Cohen, I. (2006). A model-based approach for estimating human 3d poses in static images, *IEEE Trans. Pattern Analysis and Machine Intelligence* 28(6): 905–916.
- Leibe, B. & Schiele, B. (2003). Interleaved object categorization and segmentation, *BMVC*.
- Levin, A. & Weiss, Y. (2006). Learning to combine bottom-up and top-down segmentation, *ECCV*.
- Li, Y., Sun, J., Tang, C. & Shum, H. (2004). Lazy snapping, *ACM TOG*.
- Malik, J., Shi, J., Leung, T. & Belongie, S. (1999). Textons, contours and regions: Cue integration in image segmentation, *ICCV*.
- Marschner, S., H.W.Jensen, M.Cammarano, Worley, S. & Hanrahan, P. (2003). Light scattering from human hair fibers, *Proceedings of the SIGGRAPH*, pp. 780–791.
- Moon, J. T. & Marschner, S. R. (2006). Simulating multiple scattering in hair using a photon mapping approach, *Proceedings of the SIGGRAPH*.
- Ojala, T., Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 24(7): 971–987.
- Palmer, S. (1999). *Vision science: Photons to phenomenology*, MIT Press.
- Paris, S., Briceno, H. M. & Sillion, F. X. (2004). Capture of hair geometry from multiple images, *Proceedings of the SIGGRAPH*.
- Picard, R. W. (1997). *Affective Computing*, MIT Press.
- Ren, X. & Malik, J. (2003). Learning a classification model for segmentation, *9th Int. Conf. Computer Vision*, Vol. 1, pp. 10–17.
- Riklin-Raviv, T., Kiryati, N. & Sochen, N. (2006). Segmentation by level sets and symmetry, *CVPR*.

- Ronfard, R., Schmid, C. & Triggs, B. (2002). Learning to parse pictures of people, *Proceedings of ECCV*, pp. 700–714.
- Rother, C., Kolmogorov, V. & Blake, A. (2004). Grabcut: Interactive foreground extraction using iterated graph cuts, *SIGGRAPH*.
- Rother, C., Minka, T., Blake, A. & Kolmogorov, V. (2006). Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs, *CVPR*.
- Schneiderman, H. & Kanade, T. (2000). A statistical method for 3d object detection applied to faces and cars, *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1746–1759.
- Sébastien Marcel, Y. R. & Heusch, G. (2007). On the recent use of local binary patterns for face authentication, *International Journal of Image and Video Preprocessing, Special Issue on Facial Image Processing*.
- Shechtman, E. & Irani, M. (2007). Matching local self-similarities across images and videos, *IEEE Conference on Computer Vision and Pattern Recognition 2007 (CVPR'07)*.
- Shi, J. & Malik, J. (2000). Normalized cuts and image segmentation, *PAMI*.
- Shinohara, Y. & Otsu, N. (2004). Facial expression recognition using Fisher weight maps, *Procs. of the IEEE Int. Conf. on Automatic Face and Gesture Recognition*.
- Sprague, N. & Luo, J. (2002). Clothed people detection in still images, *Proceedings of the IEEE International Conference on Pattern Recognition (ICPR)*, pp. 585–689.
- Sung, K. & Poggio, T. (1999). Example-based learning for view-based human face detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(1): 39–51.
- Swik.net: Sonys Cyber-shot T200 gets its first review* (2008).
- Tao, Q. & Veldhuis, R. (2007). Illumination normalization based on simplified local binary patterns for a face verification system, *Proc. of the Biometrics Symposium*, pp. 1–6.
- Turk, M. & Pentland, A. (1991). Eigenfaces for recognition, *J. Cognitive Neuroscience* 3(1): 71–86.
- Viola, P. & Jones, M. J. (2004). Robust real-time face detection, *International Journal of Computer Vision* 57(2): 151–173.
- Wei, Y., Ofek, E., Quan, L. & Shum, H. (2005). Modelling hair from multiple views, *Proceedings of the VISSAP*.
- Wertheimer, M. (1938). Laws of organization in perceptual forms (partial translation), *A Sourcebook of Gestalt Psychology* pp. 71–88.
- Whitehill, J., Littlewort, G., Fasel, I., Bartlett, M. & Movellan, J. (2008). Developing a practical smile detector. Submitted to *Transactions on Pattern Analysis and Machine Intelligence*. Last accessed: March 2008.
- Yacoob, Y. & Davis, L. S. (2006). Detection and analysis of hair, *IEEE Trans. Pattern Analysis and Machine Intelligence* 28(7).



## **Part 3**

# **Iris Segmentation and Identification**



## Solutions for Iris Segmentation

Milena Bueno Pereira Carneiro, Antônio Cláudio P. Veiga,  
Edna Lúcia Flôres and Gilberto A. Carrijo  
*Federal University of Uberlândia - Department of Electrical Engineering,  
Brazil*

### 1. Introduction

The growing concern with security and access control to places and sensitive information has contributed for the increased utilization of biometric systems. Biometry is the name given to the techniques used to recognize people automatically through physical and behavioural characteristics of the human body such as those found on the face, fingerprint, hand geometry, iris, signature or voice. From all the biometric options, iris recognition deserves special attention as the iris contains a huge and unique richness of characteristics, which do not change over time and enables the construction of extremely reliable and accurate systems.

The iris recognition process is relatively complex and involves several stages of processing as illustrated in Figure 1. The first stage corresponds to the localization of the region of the iris on the image of the eye, which also involves the extraction of the regions corrupted by the superior and inferior eyelids and eyelashes.

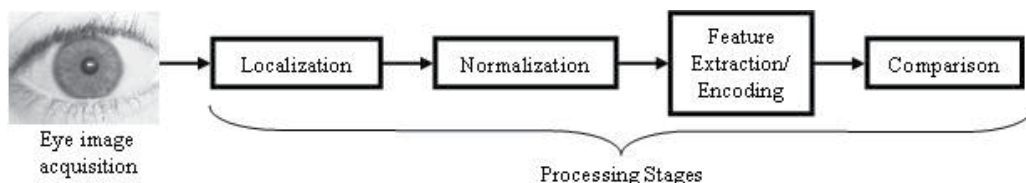


Fig. 1. Processing stages of an iris recognition system.

Once the region of the iris has been located, it must go through the normalization process. This stage is responsible for solving the problem of dimensional inconsistencies that are generally provoked by the variation of the distance between the individual and the image capture device and also by the variation of the size of the pupil due to varying levels of luminosity of the environment. The most utilized normalization method is called the “rubber sheet model” which was first suggested by Daugman (Daugman, 1993). In this model, all the annular region of the iris is uniformly sampled, in the radial direction as well as the angular direction, and is represented through the polar coordinate system, consequently generating a rectangular image.

After the normalization, the information from the iris is extracted and encoded so that the comparison of the images can finally be possible.

This chapter will focus on the first processing stage that is the segmentation (or localization) of the iris region in an eye image. The efficiency of the localization stage is essential to the success of the iris recognition system given the fact that error rates will increase if regions that do not belong to the iris are coded and processed.

In the next section, two traditional iris segmentation algorithms and a method to detect eyelashes and eyelids are described, implemented and some experimental results are collected. In section 3, we will propose a new method for iris segmentation. The proposal is based on the application of Memetic Algorithms to detect the circles which define the iris region. The implementation details of the new method are described and the experimental results are presented.

Section 4 presents an analysis of the effects of using severe compressed images. The performance of all the segmentation algorithms described previously will be measured for images with different levels of compression. Two different compression techniques will be employed.

Images from the public iris image database UBIRIS (Proença & Alexandre, 2005) were used to execute the experimental tests.

## 2. Traditional methods for iris segmentation

The pupil border and the external border of the iris, that together define the iris region, present a near elliptical contour. Nevertheless, the most traditional researches on iris recognition approximate them to circles with minimal loss in performance. Consequently, circle detection algorithms are usually applied to perform the iris localization.

Investigating the iris localization methodologies in the literature, it is possible to identify two major strategies on which those methodologies are based. Some methods are template-based and usually involve the maximization of some equations. Otherwise, they are boundary-based and demand the construction of an edge map for a later application of some geometric form-fitting algorithm.

Two methods were evaluated, each of them representing one of the possible strategies cited. The most traditional and widely used template-based method is the Integro-Differential Operator proposed by John Daugman (Daugman, 1993).

The most important boundary-based methods apply the Circular Hough Transform to find circles in an edge map, which was first suggested by Wildes (Wildes, 1997). This localization methodology was used, with several small variations, in many works such as (Huang et al., 2004; Ma et al., 2004; Masek, 2003) and many others.

The segmentation of the iris region also includes the detection and the exclusion of the eyelids and eyelashes interferences. The method proposed by Libor Masek (Masek, 2003) to exclude those interferences was also evaluated.

The main features of the two iris segmentation methods and the eyelids and eyelashes detection method applied are briefly explained in the following subsections.

### 2.1 Circular Hough transform

The Circular Hough Transform (CHT) (Gonzalez & Woods, 2002) was proposed by Paul Hough in 1962. This technique is able to recognize the circles present in an image and can be used to obtain the parameters that define the circle that represents the pupil border and the circle that represents the external iris border. These parameters are the radius and the coordinates of the circle center.

The first step is to convert the gray-scaled eye image into a binary edge map. The construction of the edge map is accomplished by the Canny edge detection method (Canny, 1986) with the incorporation of gradient information. The details of the Canny operator will not be included to this paper. It is important to know that the application of the Canny operator demands the adjustment of some parameters. After the adjustment, the same parameter values were used in all experimentations.

The Hough procedure requires the generation of a vote accumulation matrix with the number of dimensions equal to the number of parameters necessary to define the geometric form. For a circle, the accumulator will have 3 dimensions.

In the CHT, each edge pixel with coordinates  $(x,y)$  in the image space is mapped for the parameters space determining two of the parameters (for example,  $x_c$  and  $y_c$ ) and finding the third one (for example,  $r$ ) which resolves the circumference equation given in (1).

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \quad (1)$$

As a result, the point with coordinates  $(x_c, y_c, r)$  is obtained in the parameters space which represents a possible circle in the image.

At each set of parameters obtained  $(x_c, y_c, r)$ , the value of the accumulator in this position  $A(x_c, y_c, r)$  is incremented, i.e., a vote is attributed to that position. When all the pixels have been processed, the highest values of the accumulator  $A$  (i.e. the position that has received the highest number of votes) will indicate the parameters of probable circles in the image.

## 2.2 Integro-differential operator

John Daugman (Daugman, 1993) proposed an integro-differential operator to locate the circular regions of the iris and the pupil. This operator is defined by the equation (2).

$$\max(r, x_0, y_0) \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (2)$$

Where  $I(x,y)$  is the image of the eye,  $r$  is the search radius,  $G_\sigma(r)$  represents a Gaussian smoothing function and  $s$  is the contour of the  $r$  radius circle with the center in  $(x_0, y_0)$ .

The operator searches for the circular path where there is the greatest change in pixel values when there is a variation of the radius and the  $x$  and  $y$  coordinates from the center of the circle. The operator is applied iteratively with the amount of smoothing being progressively reduced in order to attain precise localization.

Although the iris search results depend on the pupil search, it is not possible to assume that the iris external circle has the same center of the pupil. Therefore, the three parameters that define the circle of a pupil must be estimated separately from the ones that define the iris.

## 2.3 Detection of eyelids and eyelashes

A procedure suggested by Libor Masek (Masek, 2003) was used to perform the detection of the region of the iris covered by the eyelids and eyelashes. To isolate the eyelids, it was assumed that their edges could be approximated by a line segment. The first step is to find a line that corresponds to the edge of the superior eyelid and one that corresponds to the inferior eyelid. To achieve this, the Linear Hough Transform was used. A second line is then drawn horizontally intercepting the first at the edge point of the iris closest to the pupil. This procedure is done for both the superior and the inferior eyelids. The region above the superior eyelid line and the region below the inferior eyelid line are excluded.

It may happen that, in some images, there is no occlusion of the iris by the eyelids. Therefore, if the maximum value in the Hough space is smaller than a predetermined threshold, no line is identified, which represents a non occlusion. Moreover, a line is only considered when it is found out of the pupil region and in the iris region.

To isolate the eyelashes a threshold determination technique is used, considering that in the group of images used the eyelashes are in general a little darker when compared to the rest of the image. Consequently, all of the pixels in the image darker than the threshold established are considered to be pixels which belong to the eyelashes and are consequently excluded.

## 2.4 Experimental results

There are presently several public and freely available iris image databases for biometric purposes. The majority of them incorporates few types of noise, almost exclusively related with eyelid and eyelash obstruction. The UBIRIS database (Proença & Alexandre, 2005) tries to simulate noncooperative imaging conditions. Its capture conditions produced images with very heterogeneous characteristics regarding focus, motion blur, contrast, brightness, as well as iris occlusions by eyelids or eyelashes and specular and lighting reflections. Mainly because of that, the UBIRIS database was chosen to be used in this work for the experimental tests, aiming to approach to a more realistic situation which makes the iris localization challenging. It was used a total of 1201 gray scale eye images with resolution 200x150.

The two iris segmentation methods and the eyelashes and eyelids detection method were applied to the images from the UBIRIS database. When the circular Hough transform was applied to the 1201 images, it obtained success in localizing the iris region in 1119 images, which means an efficiency of 93.17% and an error rate of 6.83%.

The integro-differential operator was able to correctly localize the iris region in 1135 images, which means an efficiency of 94.50% and an error rate of 5.5%.

The algorithm used to detect the eyelashes and eyelids interferences presented an efficiency of 95.25% (1144 images) and an error rate of 4.8%.

The classification of the results as correct or not correct was manually done by a visual inspection of the result of the application of each method in each image.

## 3. Proposed method for iris segmentation

We propose a method based on memetic algorithm to localize the iris region. In this section, the fundamentals of the proposal are described and the utilization of evolutionary algorithms is justified.

As well as the CHT, the proposed method demands that the original image passes through an edge detection algorithm. There are several edge detection techniques in the literature and, as well as in the implementation of the CHT, the Canny method (Canny, 1986) was employed.

It is necessary to obtain a binary edge image as that shown in Figure 2. In this figure, the white pixels represent edge pixels.

As suggested by Wildes (Wildes, 1997), only vertical edges were detected for external iris border localization, whereas vertical and horizontal gradients were equally weighted for pupil localization.

From Figure 2, one notices that the pupil and the iris contours are apparent. Suppose that, using a geometric figure of a circle (imaginary circle), one wishes to overlap, on the edge

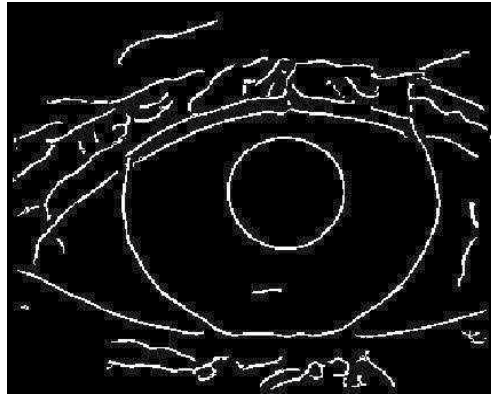


Fig. 2. Eye image after the application of the Canny edge detection algorithm.

image, the highest quantity of white pixels (edge pixels) possible. It is intuitive that, to reach this objective, the imaginary circle should be plotted onto the pupil or on the external border of the iris.

From this principle, it is possible to find the two desired circles through a scanning on the edge image. This scanning consists of varying the parameters of the imaginary circle into the image limits and, for each set of parameters ( $x_c$ ,  $y_c$ , radius), to plot the corresponding circle on the edge image and to count the quantity of edge pixels that are overlapped. Thus, in the end of the scanning, the imaginary circle that has overlapped more edge pixels, is considered to be the contour of the pupil or the iris.

The search for each one of the desired circles must be carried out separately. The images of the database used (UBIRIS (Proença & Alexandre, 2005)) have a resolution of 200 pixels along the horizontal direction (coordinate  $x$ ) and 150 pixels along the vertical direction (coordinate  $y$ ). For those images, it can be considered that the radius of the pupil ranges from 5 to 36 pixels while the radius of the iris ranges from 40 to 71 pixels. Hence, to implement a complete scanning, the coordinate  $x_c$  must vary from 1 to 200, and for each assumed  $x_c$ , the coordinate  $y_c$  must vary from 1 to 150. Moreover, for each pair of center coordinates ( $x_c$ ,  $y_c$ ) the radius must vary from 5 to 36 for a pupil search or from 40 to 71 for an iris search.

Analyzing this procedure, one observes that, to find the external border of the iris, it is necessary to consider  $200 * 150 * (5 - 36) = 930.000$  imaginary circles, and to find the pupil, it is necessary to consider  $200 * 150 * (40 - 71) = 930.000$  imaginary circles. Under these conditions, this procedure is impracticable, as it demands intense computational processing. This is a kind of problem for which there are many possible solutions and a vast search space. Some heuristic techniques are capable of approaching these problems. Here, Memetic Algorithms were employed to find the parameters of the iris and pupil circles in an adequate time, making the process less computationally demanding.

### 3.1 Memetic algorithms description

The called Memetic Algorithms (MAs) combine concepts of “local search” with the other operators typically used on Genetic Algorithms (GAs) that are inspired on Darwin's theory of evolution. Darwin established that, in a biological evolutionary process, after many generations, the population evolves according to the principles of natural selection and the

survival of the fittest. Memetic algorithms take into consideration not only the "genetic" evolution of individuals, but also a form of "cultural" evolution that is generally accomplished by a local search algorithms (Moscato, 2001).

The pseudo-code shown in Figure 3 contains the main steps of the computational implementation of the memetic algorithms. Firstly, in line 3, a population of individuals is generated randomly and stored in the variable *Pop*. The number of individuals in the population is *PopSize*. Each individual is represented by a chromosome, which is constituted by a code that represents a solution to the problem. All the generated individuals should be evaluated and a value is associated to each one of them in order to represent their fitness, which means, how well the solution fits the problem in question. The evaluation is carried out on line 5.

Lines 6 to 21 show the steps for creating the population of the next generation. The first step is the "Elitism", which is used to improve the convergence of the algorithm. It consists of holding back the *n* best individuals in each generation.

To complete the new population, some individuals are selected for reproduction. The selection mechanism takes into consideration their fitness, so that an individual has a selection probability proportional to its fitness. Once two individuals are selected (lines 9 and 10) their genes are recombined through the crossover operators with a probability of "CrossoverProbability". To achieve that, a "RandomNumber" between 0 and 1 is generated and if it is smaller than "CrossoverProbability" the crossover is executed and two new individuals are created (lines 13 and 14). The last genetic operator is the mutation, which alters the value of one or more genes of the chromosome chosen randomly with a small probability of "MutationProbability" (lines 15 to 17). Mutation increases the diversity of the population and assures that any point of the search space can be reached.

```

1 // MA Procedure
2
3 Pop = GenerateInitialPopulation(PopSize)
4 Repeat
5     CalculateFitness(Pop);
6     Elitism(n, Pop, NewPop);
7     i = n+1;
8     Repeat
9         Ind1 = Selection(Pop);
10        Ind2 = Selection(Pop);
11        NewPop(i) = Ind1;
12        NewPop(i+1) = Ind2;
13        If RandonNumber < CrossoverProbability
14            [NewPop(i), NewPop(i+1)] = Crossover(Ind1, Ind2);
15        If RandonNumber < MutationProbability
16            NewPop(i) = Mutation(NewPop(i));
17            NewPop(i+1) = Mutation(NewPop(i+1));
18        i = i+2;
19    Until NewPop is complete
20    LocalSearch(NewPop);
21    Pop = NewPop;
22 Until stop criterion is satisfied

```

Fig. 3. Pseudo-code of the memetic algorithm procedure.



The two new individuals are added to the new population. When the new population is complete, a local search procedure is executed (line 20) for each individual. The local search algorithm evaluates the individuals in a neighbourhood defined by a mechanism of neighbourhood generation and if a fitter individual is found, it substitutes the previous individual and is added to the new population. A pseudo-code of a genetic algorithm can be obtained from the pseudo-code of Figure 3 by simply excluding line 20.

The whole procedure is repeated until the stop criterion is satisfied. The stop criterion can be the number of generations, the execution time or an acceptable value of fitness.

The application of MAs to the iris localization is described in the next section.

### 3.2 Implementation of the proposed method using memetic algorithms

An implementation of a memetic algorithm has been utilized to solve the aforementioned problem.

First, it is necessary to define a structure for the chromosomes, which will represent possible solutions to the problem. A 17-bit binary representation is used, and each gene of the chromosome is assigned a value 0 or 1; the first 6 bits (or genes) refer to the coordinate  $x_c$  for the circle center, the next 6 bits refer to the coordinate  $y_c$  of the center and the last 5 bits represent the radius of the circle.

The 6-bit structure makes possible to represent 64 different values of  $x_c$  and  $y_c$ . As mentioned before, the size of the images is  $200 \times 150$ , but, in general, the iris is positioned in the central part of the eye image, thus, it is not necessary to look for the centers of the iris and the pupil in the extremities of the image. Therefore, the coordinate  $x_c$  is considered to be in the range of 69 to 132 and the coordinate  $y_c$  is considered as being in the range of 44 to 107. The radius is represented by a 5-bit structure and so, 32 different values can be obtained. For the pupil, the 32 different values are between 5 and 36, and for the iris, they are between 40 and 71.

Another extremely important step is the definition of a fitness function that assigns a fitness score for each individual. The number of edge pixels that match the circle represented by the chromosome defines the fitness score for that individual. Thus, the greater the number of matched pixels, the fitter an individual is.

The algorithm must generate an initial population representing the initial search space. The generation of the initial population is carried out by randomly assigning one binary value to each gene of a chromosome.

Two populations of individuals have been generated, one consisting of candidates to represent the external border of the iris, and the other to represent the pupil border. These populations are processed independently.

Two common genetic operators have been used: one-point crossover and flip mutation. In the first case, a pair of individuals is selected from the population and their chromosomes are trimmed at a random position; this yields two tails that are swapped between the two chromosomes so that two new different individuals are generated. Mutation can be applied to every descendant after the crossover. It alters some genes of some chromosomes at random and with a low probability.

There are different crossover and mutation operators for binary chromosomes. The one-point crossover and flip mutation are the simplest operators found in literature. As they reached good results they were chosen in order to minimize the computation effort and the execution time.

Before proceeding to the next generation, a local search is performed to add the “cultural evolution”. In the next section, the local search mechanism used to implement the memetic algorithm is described.

### 3.3 Implemented local search mechanism

Observation of the performance of the evolutionary part of the algorithm (with no local search) demonstrates that many times, solutions remain stuck at local minima for several generations. Graphically, it looks like the picture in Figure 4.

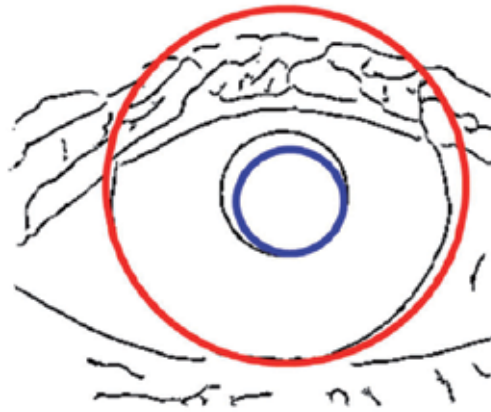


Fig. 4. The local minima issue that often occurs during the performance of a GA.

One notices that the imaginary circles, which represent the solution in question, are very close to an optimal solution, but the evolutionary algorithm could only generate an individual with these optimal characteristics after many generations. This is exactly where a local search becomes advantageous and accelerates the convergence.

The neighbourhood of the solution could be investigated by shifting the center of the circle to the right and to the left ( $x$  coordinate), upwards and downwards ( $y$  coordinate) and, for each of these variations, the radius should be ranged. However, this form of local search is computationally expensive, especially when the population contains a great number of individuals.

For that reason, in order to lessen the computational effort of the local search, a strategy for reducing the neighbourhood has been proposed and implemented, which guides the coordinates of the circle. Visually, one notices that the solution in Figure 4 is such that the circle touches the pupil border in a certain extension, i.e. some pixels of the imaginary circle are coincident with pixels of the pupil border. This information might be useful to guide the coordinates of the circle.

The local search mechanism is structured in the pseudo-code shown in Figure 5. It is applied to all individuals of the population (line 2), and each individual contains the parameters of a circle. The circle can be divided into quadrants and two diagonal lines can be defined as shown in Figure 6.

```

1 // Procedure LocalSearch(NewPop)
2 For each individual of the population
3     Repeat
4         Draw circle represented by the individual on edge image;
5         Discover each quadrant of the circle coincides with the highest...
6         ... amount of edge pixels;
7         If the highest amount of coincident edge pixels is in quadrant 1
8             SelectedDiagonal = diagonal A;
9         If the highest amount of coincident edge pixels is in quadrant 2
10            SelectedDiagonal = diagonal B;
11         If the highest amount of coincident edge pixels is in quadrant 3
12            SelectedDiagonal = diagonal A;
13         If the highest amount of coincident edge pixels is in quadrant 4
14            SelectedDiagonal = diagonal B;
15         i = 1;
16         While i <= centerRange Do
17             Shift the center along SelectedDiagonal;
18             j = 1;
19             While j <= radiusRange Do
20                 Vary radius;
21                 NewFitness = Calculate fitness of the new ...
22                 ...individual (new center and new radius)
23                 If NewFitness > Fitness of the original individual
24                     Original individual = New individual;
25                 j = j+1;
26             endwhile
27             i = i+1;
28         endwhile
29     Until no improvement is detected during a iteration.
30 endFor

```

Fig. 5. Pseudo-code of the local search procedure.

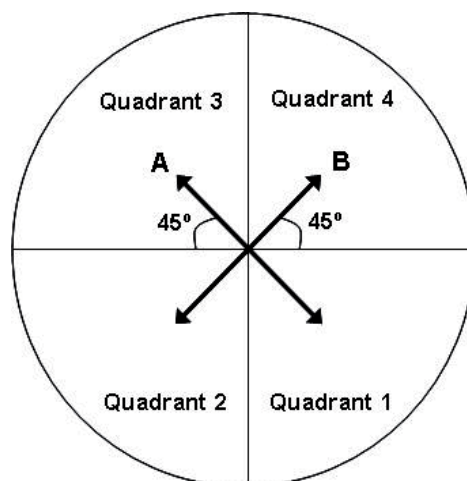


Fig. 6. Direction of the possible diagonal lines to be selected during the local search procedure.

To evaluate the fitness of an individual, the amount of edge pixels that coincide with the circle when it is plotted on the edge image is summed up. The local search algorithm verifies in which of the quadrants there have been the highest number of coincident pixels (line 5). If quadrants 1 or 3 have the highest amount of coincident pixels, the local search mechanism will vary the center coordinates  $x_c$  and  $y_c$  maintaining the center of the imaginary circle always along the diagonal line A (lines 7-8 and 11-12). If, on the other hand, the majority of the coincident pixels is in quadrants 2 or 4 the center of the imaginary circle must be shifted along the diagonal line B (lines 9-10 and 13-14). Therefore, to evaluate the neighbourhood, the center of the imaginary circle must be shifted a number of times equal to "centerRange" along the selected diagonal (lines 16 and 17), and, for each assumed position, the circle radius must also be ranged a number of times equal to "radiusRange" (line 19 and 20). For each situation, the new fitness must be calculated (line 21). If the new fitness is higher than the original fitness, the original individual is substituted by the new individual (line 23 and 24). This process is repeated for the same individual until improvements are no longer reached during an execution (line 29). Figure 7 uses the circle that represents the pupil border to illustrate this procedure.

Figure 7a shows a solution obtained by the genetic algorithm. Figure 7b shows the effect of shifting the center of the circle. After shifting the center, the radius is also ranged, as illustrated in Figure 7c. As there has been an enhancement in the fitness of that individual, the new individual substitutes the old one and another iteration of the local search algorithm is allowed. Once more, the amount of coincident pixels of each quadrant is evaluated, new coordinates for the center of the circle are generated (see Figure 7d) and new radiuses are investigated (see Figure 7e). The process is finished when no further enhancement in fitness is verified.

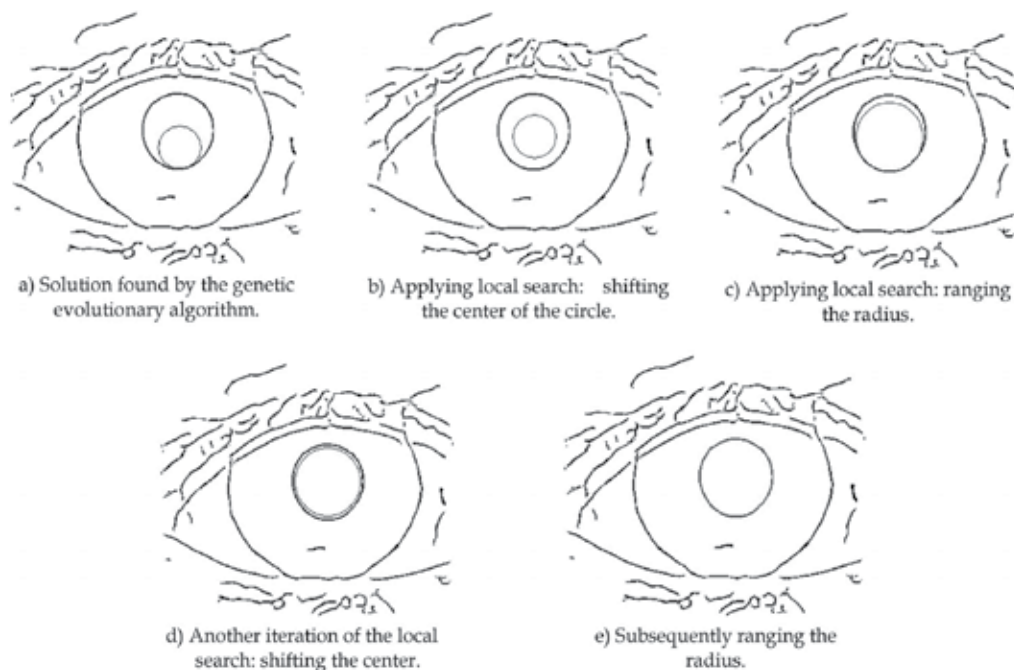


Fig. 7. Illustration of the local search procedure.

The result of this process is such that the local search alters the individual and improves its fitness score. This individual returns to the evolutionary algorithm with a better fitness compared to the one it originally had, due to the cultural evolution obtained by the local search, which exploited the available knowledge of the problem.

### 3.4 Parameters definition

The implementation of a MA requires the definition of some variable parameters that are population size, number of generations, crossover and mutation rates.

The bigger the population size, more possible solutions are represented and higher the chance of good individuals to be present in the population. On the other hand, the increase of the population size results in the increase of the processing time.

If an evolutionary algorithm is well projected, one expects that, in a given generation, the individuals are more adapted or, at least, equally adapted to the environment than in the previous generations. Therefore, the bigger the total number of generations, the greater the convergence to a good solution, but it is evident that the processing time becomes longer. The number of generations can be the stop criterion.

The crossover rate represents the probability of a recombination of two selected individuals. If the crossover rate is 0%, the entire new generation is composed of exact copies of the parents. If it is 100%, all offspring is created by crossover.

The mutation rate, in this work, represents the probability of altering one or more genes of a chromosome. If the mutation rate is 0% nothing is changed, otherwise, if it is 100%, all individuals will suffer some alteration and the procedure will be similar to a random search.

In order to define a suitable set of parameters, the algorithm was applied numerous times to all images of the database and its behaviour was analyzed while the parameters varied. During the simulation, the population size ranged uniformly from 10 to 200 individuals. For a greater number of individuals, the processing time becomes inconveniently long. The crossover and mutation rates ranged uniformly from 0% to 100%. The time of convergence of the algorithm was the metric used to evaluate the parameters.

Best results were achieved using a population of 50 individuals, a crossover rate of 90% and a mutation rate of 20%. It was verified that, for the majority of images, the iris region was correctly segmented until the fifth generation. So, the number of generations was considered equal to 8, which was used as the stop criterion of the MA.

### 3.5 Experimental results

Systems based on any heuristic techniques, such as Memetic Algorithms, are non deterministic. It means that, a different solution can be obtained each time the system is simulated. After defining the parameters of the algorithm (population size = 50 individuals; crossover rate = 90%; mutation rate = 20%), the system was simulated many times. In each simulation, the algorithm was applied to all 1201 images and the number of images that were correctly segmented was summed. After all simulations, the arithmetic mean and the standard deviation of the results were calculated.

Considering all the simulations, the average number of images that were correctly segmented was 1166 from 1201 images. In an average of 13 images only the pupil border was correctly segmented and in an average of 12 images only the external border of the iris

was correctly segmented. In an average of 10 images neither the pupil border nor the external border of the iris were correctly segmented. So the algorithm failed in segmenting a total of 35 images. It means that, when the algorithm is applied to images, an average 97.08% of the segmentations will be acceptable.

The standard deviation measured the spread of data about the mean and it was equal to 2.22. This is a small value and means that, for various simulations, the efficiency of the algorithm is always near the mean efficiency.

The results show that the method is robust enough to deal with not ideal conditions since a good efficiency rate could be achieved even with bad contrast and the presence of specular reflection in images.

#### **4. Analysing the performance of iris segmentation methods applied to severely compressed images**

An iris recognition system requires that the data related to the iris be stored for future comparison. The problem of storing the raw format of the image of the iris is that these images occupy a lot of space on disk and demand a high bandwidth for transmission. There is the alternative of storing only the code or template generated from the characteristics of the iris, which is considerably smaller. However, this code is generated by some proprietary algorithm, which would make the system dependent on some specific supplier. Presently, governmental organizations have demanded that biometric data be stored and recorded in the form of raw images, in order to have more negotiation freedom.

The fact that images are stored in raw format, that is, without any pre-processing, also has the advantage of allowing the data bank to take advantage of the inevitable evolutions of recognition algorithms in the future.

In this context, it is possible to see the necessity of image compression in biometry, and also the importance of knowing how much the captured image can be compressed without harming the performance of the iris recognition system.

Here, it will be verified what happens to the efficiency of the segmentation algorithms when the images undergo severe compression. For such, the fractal compression method that utilizes quadtree partitioning and the JPEG2000 compression are used with the objective of compressing the iris images at various compression rates, and analyzing up to what point compression would be viable.

Firstly, some features of the compression algorithms are presented and then, all the experimental results are shown.

##### **4.1 Compression algorithms**

Two different compression algorithms were applied to the iris images. One of them was the widely used JPEG2000 compression and the second one was chosen to be the Fractal compression that is based on different concepts.

The main characteristics of each of these techniques are presented in the following subsections which make clear that they have different principles. The main reason for the utilization of the second method was to guarantee that the conclusions about the behaviour of the segmentation algorithms are not dependent on the kind of compression technique that is applied.

#### 4.1.1 JPEG2000 compression

JPEG2000 is an image coding system that was created by the Joint Photographic Experts Group committee in 2000. It is a more powerful version of JPEG coding offering improved image quality at very high compression ratios (Information Technology - JPEG2000 Image Coding System, 2004).

JPEG2000 has a superior compression performance over JPEG and it is attributed to the use of Discrete Wavelet Transform (DWT) and a more sophisticated entropy encoding scheme instead of the Discrete Cosine Transform (DCT). As a result, artefacts are less visible and there is almost no blocking. Moreover, JPEG2000 decomposes the image into a multiple resolution representation in the course of its compression process allowing local areas within each image tile to be encoded using different sub bands of coefficients (Christopoulos et al., 2000).

The advantage of the JPEG2000 over JPEG in terms of image quality is more evident at very low bit rates (Christopoulos et al., 2000) that correspond to severe compression, as studied here.

The JPEG2000 compression and decompression at various quality factors was performed using the Linux tools *pantojpeg2k* and *jpeg2ktopam* from the JasPer JPEG2000 and Netpbm libraries.

#### 4.1.2 Fractal compression

Fractal compression is a technique based on the principle that real images have self similarity and that these similarities consist of redundant information, which can be eliminated. For this, several transformations are applied over all or part of an image which in turn makes each part of the image to have a convergence to a common point called the attractor (Fisher, 1995). This way, the images can be stored through groups of affine transformation coefficients, consequently of reducing the memory utilized.

Fractal compression differs from other lossy compression methods, such as JPEG, in a number of ways. JPEG achieves compression by discarding image data that is not required for the human eye to perceive the image. The resulting data is then further compressed using a lossless method of compression. To achieve greater compression ratios, more image data must be discarded, resulting in a poorer quality image with a pixelized (blocky) appearance (Murray & Van Ryper, 1996). Fractal images are not based on a map of pixels. Once an image has been converted into fractal code its relationship to a specific resolution has been lost; it becomes resolution independent. The image can be recreated to fill any screen size without the introduction of image artefacts or loss of sharpness that occurs in pixel-based compression schemes.

Here, a fractal compression method with quadtree partitioning (Fisher, 1995) was implemented in Java environment. The implementation assures great processing speed of the images and wide parameter manipulation with the purpose of balancing image quality and compression rate. The coefficients of the compression algorithm were adjusted so as to obtain images with various degrees of compression.

#### 4.2 Experimental results

The JPEG2000 compression and the fractal compression method were utilized to compress the images from the UBIRIS database (Proença & Alexandre, 2005) with the following

compression rates (CR): 0.7, 0.5, 0.3 and 0.15. In this way it was possible to evaluate the efficiency of the segmentation methods when the images suffer compression from moderate rates (0.7 rate) to extremely severe compression rates (0.15 rate).

Table 1 shows the effect of compression on the images. In average, the original images have a size of 22000 bytes, consequently compression at rates of 0.7, 0.5, 0.3 and 0.15 produce compressed images with an average size of 15400, 11000, 6600, and 3300 bytes respectively. In relation to the original image, compression at these rates represents an average reduction factor of 1.5:1, 2:1, 3.5:1 and 6.7:1, respectively.

In order to apply both the Circular Hough transform and the proposed method based on memetic algorithms to localize the iris region in an image, it is necessary to generate an edge map from this image first, remembering that the Canny edge detector operator were applied and the same parameter values were used throughout the experimentation.

|           | Size    | Reduction factor |
|-----------|---------|------------------|
| Original  | 22 KB   | 1 : 1            |
| CR = 0.7  | 15.4 KB | 1.5 : 1          |
| CR = 0.5  | 11 KB   | 2 : 1            |
| CR = 0.3  | 6.6 KB  | 3.5 : 1          |
| CR = 0.15 | 3.3 KB  | 6.7 : 1          |

Table 1. Effects of images compression - Average values.

Therefore before verifying the influence of compression in the iris segmentation, it is interesting to analyze the effect of compression on the edge map that is generated.

The edge map of each original image was compared to the edge map of the corresponding image compressed at the rates of: 0.7, 0.5, 0.3 and 0.15 using JPEG2000 compression and also Fractal compression. In average, 91.3% of the pixels considered to be edge pixels in the original image were also considered edge pixels in the edge map generated from the compressed image at a rate of 0.7 bpp when JPEG2000 compression was applied and 90.5% when Fractal compression was applied. For the edge maps generated from the images compressed at rates of 0.5, 0.3 and 0.15 bpp, this ratio was, in average of 81.1%, 78.5% and 49.8%, respectively, when JPEG2000 was used and 79.7%, 78.1% and 51.5%, respectively, when Fractal compression was used.

In the following subsections, the results obtained with each segmentation method are presented. It is important to emphasise that the segmentation rates that are presented in the following subsections represent the percentage of the images that were correctly segmented without compression which were also correctly segmented after compression. All results are summarized in Table 2.

#### 4.2.1 Results of the circular Hough transform application

We observed that with the compression rate of 0.7 bpp, there was little variation in the edge map and as a result, from the original images that had their iris region correctly found (or



segmented), the percentage of compressed images that had also the iris region correctly found was 99.8% for the JPEG2000 compression and 99.5% for the Fractal compression. This shows that moderate compression practically does not interfere in the efficiency of the Hough transform.

For the compression rates of 0.5 and 0.3, the edge map varied a little more, even so the algorithm was capable of correctly finding the iris region from, respectively, 99.4% and 98.4% of the images compressed by JPEG2000 and 99.3% and 98.5% of the images compressed by Fractal. This still represents a very acceptable efficiency.

With a compression rate of 0.15, the algorithm was capable of correctly finding the iris region from 93.5% of the images compressed by JPEG2000 and 94.7% of the images compressed by Fractal. The reduction of the efficiency of the algorithm was probably due to the more intense variation that occurred in the edge maps.

#### **4.2.2 Results of the integro-differential operator application**

When applying the integro-differential operator to original and compressed images it was observed that when the compression was performed with rates of 0.7 and 0.5 bpp, 100% of the original images that had their iris region correctly found were also properly segmented when compressed using both JPEG2000 and Fractal compression. Therefore, compression at these rates did not harm in any way the efficiency of the segmentation algorithm.

With a compression rate of 0.3 bpp, the algorithm was capable of correctly finding the iris region from 98.3% of the images compressed by JPEG2000 and 98.6% of the images compressed by Fractal, which represent a margin of error that can be ignored in face of the benefits of compression.

When a more severe compression rate was used (0.15), 88.2% of the images compressed by JPEG2000 and 89.7% of the images compressed by Fractal had their iris region correctly found. In this case, it is observed that the algorithm suffered greater interference from compression and for this technique this was especially due to the reduction in contrast between the white part of the eye and the iris and between the iris and the pupil in the compressed image.

#### **4.2.3 Results of the eyelids and eyelashes detection method**

The proposed method for the detection of eyelids and eyelashes was also applied to the original images and to compressed images with rates of 0.7, 0.5, 0.3 and 0.15 using JPEG2000 compression and Fractal compression. It was noticed that 100% of the original images that had the eyelids and eyelashes correctly detected also had them correctly detected using any of the compression rates and any of the compressing algorithms. This shows that compression, even when severe, does not harm the algorithm utilized to detect eyelids and eyelashes.

#### **4.2.4 Results of the proposed iris segmentation method**

Applying the proposed method based on MA, 100% of the original images that had their iris region correctly found were also properly segmented when compression was performed with rates of 0.7 and 0.5 bpp using both JPEG2000 and Fractal compression. This shows that despite the little variation of the edge map the method was robust enough to guarantee the same efficiency with or without compression.

For the compression rates of 0.3 and 0.15 bbb, the algorithm was capable of correctly finding the iris region from, respectively, 98.9% and 95.2% of the images compressed by JPEG2000 and 99.1% and 95.7% of the images compressed by Fractal.

For a clearer comparison, the results are summarized in Table 2.

|          | CR   | Circular Hough transform | Integro-differential operator | Eyelids and eyelashes detection | Proposed method |
|----------|------|--------------------------|-------------------------------|---------------------------------|-----------------|
| JPEG2000 | 0.7  | 99.8%                    | 100%                          | 100%                            | 100%            |
|          | 0.5  | 99.4%                    | 100%                          | 100%                            | 100%            |
|          | 0.3  | 98.4%                    | 98.3%                         | 100%                            | 98.9%           |
|          | 0.15 | 93.5%                    | 88.2%                         | 100%                            | 95.2%           |
| Fractal  | 0.7  | 93.5%                    | 100%                          | 100%                            | 100%            |
|          | 0.5  | 99.3%                    | 100%                          | 100%                            | 100%            |
|          | 0.3  | 98.5%                    | 98.6%                         | 100%                            | 99.1%           |
|          | 0.15 | 94.7%                    | 89.7%                         | 100%                            | 95.7%           |

Table 2. Effects of images compression in the segmentation stage. The values correspond to the percentage of the original images that were correctly segmented without compression which were also correctly segmented after compression.

## 5. Conclusion

The segmentation of the iris region is the first processing stage of an iris recognition system. The efficiency of this stage is essential to the success of the recognition task. In this work some traditional iris segmentation methods were firstly presented and evaluated by applying them to images from the UBIRIS database. Those images were chosen once they were captured under not ideal conditions and so, were challenging enough to the segmentation stage, what allows knowing how the methods would perform in real situations.

Then, a new segmentation method based on Memetic Algorithm that is an evolutionary algorithm was proposed and evaluated. Firstly, the basic concepts of Memetic Algorithms were described and then, the implementation of the proposed method was carefully detailed. The experimental results indicate that the proposed method is efficient at localizing the iris region in an image of an eye. The innovation related to the use of evolutionary algorithms to localize the iris region can be further explored in order to achieve even better results.

This work also examined the influence of severe image compression on the segmentation stage. Images of the UBIRIS database were compressed using two different compression algorithms: JPEG2000 and Fractal compression. It was possible to conclude that, in general, the most traditional boundary-based and template-based algorithms for iris segmentation continue presenting good performance even with compressed images.

Finally, it is concluded that in a complete iris recognition system, the segmentation stage will not represent a bottleneck, that is, it will not compromise the performance of the system in case compressed images are used.

The next work to be suggested is to perform a research of how compression interferes in the other processing stages, especially in the system's ability to recognize individuals precisely.

## 6. References

- Canny, J. (1986). A computational approach to edge detection, *IEEE Transactions on PAMI-8*, No.6.
- Christopoulos, C. ; Skodras, A. & Ebrahimi, T. (2000). The JPEG2000 still image coding system: An overview, *IEEE Trans. Consum. Electron*, Vol. 46, No.4, pp. 1103-1127.
- Daugman, J. D. (1993). High confidence visual recognition of person by a test of statistical independence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.11, pp. 1148-1161.
- Fisher, Y. (1995). *Fractal Image Compression - Theory and Application*, New York: Springer-Verlag, 1995. 341 p. ISBN: 0-387-94211-4 (New York) - 3-540-94211-4 (Berlin)
- Gonzalez, R. C. & Woods, R. E. (2002). *Digital image processing* (2<sup>nd</sup> edition), Prentice Hall, Upper Saddle River
- Huang, J.; Wang, Y.; Tan, T. & Cui, J. (2004). A new iris segmentation method for recognition, *In: Proceedings of the 17th International Conference on Pattern Recognition*, Vol.3, pp. 554- 557, DOI: 10.1109/ICPR.2004.1334589.
- Information Technology - JPEG2000 Image Coding System (2004). Int. Std. ISO/IEC 15444-1, 19 March 2011, Available from: <http://www.itu.int/rec/T-REC-T/en>.
- Ma, L.; Wang Y. & Zhang, D. (2004). Efficient iris recognition by characterizing key local variation, *In: IEEE Transactions on Image Processing*, Vol.13, No.6, pp 739-750.
- Masek L. (2003). *Recognition of human iris patterns for biometric identification*, Master Thesis. School of Computer Science and Software Engineering, University of Western Australia.
- Moscato P. (2001). *NP Optimization Problems, Approximability and Evolutionary Computation: From Practice to Theory*, Ph.D. dissertation, University of Campinas, Brazil.
- Murray, J. D. & Van Ryper, W. (1996). *Encyclopedia of Graphics File Formats*, (2<sup>nd</sup> edition), ISBN: 1-56592-161-5.
- Proença, H. & Alexandre, L. A. (2005) UBIRIS: a noisy iris image database, ICIAP 2005, *13th Int. Conf. on Image Analysis and Processing*, Cagliari, Italy, 6-8 September 2005, Lect. Notes Comput. Sci., 3617, pp. 970-977, ISBN 3-540-28869-4. <http://iris.di.ubi.pt>

Wildes, R. P. (1997). Iris recognition: An emerging biometric technology, *Proceedings of the IEEE*, Vol.85, No.9, pp. 1348-1363.

# Detecting Cholesterol Presence with Iris Recognition Algorithm

Ridza Azri Ramlee, Khairul Azha and Ranjit Singh Sarban Singh  
*Universiti Teknikal Malaysia Melaka (UTeM),  
Malaysia*

## 1. Introduction

Iris is a pigmented, round, contractile membrane of the eye, suspended between the cornea and lens and perforated by the pupil (Fig. 1). It regulates the amount of light entering the eye (Online Dictionary). According to (David J. Pesek, 2010), the eyes are connected and continuous with the brain's Dura mater through the fibrous sheath of the optic nerves, and they are connected directly with the sympathetic nervous system and spinal cord. The optic tract extends to the thalamus area of the brain. This creates a close association with the hypothalamus, pituitary and pineal glands. These endocrine glands, within the brain, are major control and processing centers for the entire body. Because of this anatomy and physiology, the eyes are in direct contact with the biochemical, hormonal, structural and metabolic processes of the body. This information is recorded in the various structures of the eye, i.e. iris, retina, sclera, cornea, pupil and conjunctiva. Thus, it can be said that the eyes are a reflex or window into the bioenergetics of the physical body and a person's feelings and thoughts (David J. Pesek., 2010). There are a lot of arguments between iridologists (iridology's practitioner) and the medical's practitioner. Due to this argument, numerous studies done by the medical's practitioner found that the diagnosis done by the iridologist upon the patient is not accurate (Allie Simon et al, 1979). However the study on relationship diseases to iris changes, still continuing for example the studied done on Ocular complication of adult rheumatoid arthritis don by S.C.Reddy and U.R.K.Rao in 1996 found that the mean duration of the arthritis and the mean duration of seropositivity were found to be significantly higher in patients with ocular (pigmented organ in eye) complication (S.C.Reddy et al, 1996). Another study done on bilateral retinal detachment in acute myeloid Leukemia by (K Pavithran et al., 2003), found that ocular manifestations are common in patient with acute Leukemia. This can result from direction infiltration by neoplastic cells of ocular tissues, including optic nerve, choroid, retina, iris and ciliary body, or secondary to hematology abnormalities such as anemia, thrombocytopenia, or hyperviscosity states or retinal destruction by opportunistic infection (K Pavithran et al., 2003). The history of Iridology study on iris was done by the physician Philippus Meyens in 1670 in a book explaining that the features of the irid called Chromatica Medica. In that book he wrote that the eye (iris) contains valuable information about the body. In 1881 a Hungarian physician, Dr. Ignatz Peczley who is claimed as the founder of modern Iridology wrote a book "Discoveries in the Field of Natural Science and Medicine, a guide to the study and

diagnosis from the eye." He introduced the first chart of the iris explaining zone in the iris. The idea of his study on iris, begun when he was a child, he was accidentally found the Owl with broken leg. He found a dark scar in the Owl's iris that scar turned white as the leg healed (Sandy Carter, 1999). The objective of this chapter is to explain how the presence of cholesterol in blood vessel can be detected by using iris recognition algorithm. This method used the John Daugman's and Libor masek's iris recognition methods and extends the study of eyes pattern to other application and in this case, the alternative medicine that is iridology. Based on the iris recognition methods and iridology chart, a MATLAB program has been created to detect the present of cholesterol in our body. However, further analysis must be done in order to know the exact range or level of cholesterol in blood vessel.

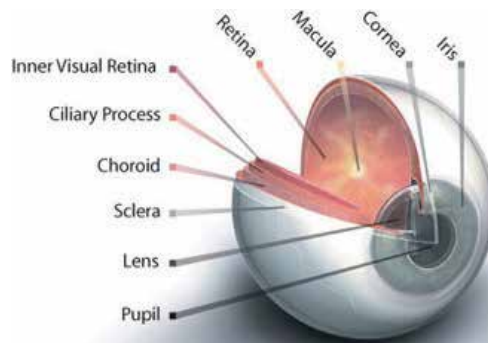


Fig. 1. Human Eye Anatomy.

### 1.1 Relationship cholesterol presence with arcus senilis

Hypercholesterolemia or a high level of cholesterol in the blood poses a significant threat to person's health. Even though it is not considered as a disease, it can be secondary to a disease and can help contribute to other many forms of diseases most notable are cardiovascular diseases. So, it is important to have our blood cholesterol levels checked. A contemporary technique to measure the cholesterol level is by doing blood test and the test is known as lipoprotein profile. The lipoprotein profile is done and preferably after a 9- to 12- hour fast and it measures the levels of total cholesterol.

The lipoprotein profile can be considered as intrusive means if it is just used merely for cholesterol screening. (N.Haq, M.D.Fox, 1991) introduced laser based technology as non-intrusive technique to measure blood cholesterol through skin. They proposed infrared (IR) absorption spectroscopic as the characterization of cholesterol in the skin. Based on U.S Food and Drug Administration (FDA), (FDA, 2004), skin contains approximately 11 percent by weight of all body cholesterol and when severe coronary artery disease is present, the numeric values obtained with the skin cholesterol test increases. Thus, the palm test for skin cholesterol is not to be useful in identifying people with less severe coronary artery disease and it is not intended to be used as a screening tool to determine the risk for coronary artery disease in general population.

In order to have a simple and non-intrusive means to be as a screening tool to detect cholesterol, we have considered alternative medicines. Iridology is one of the alternative medicines, which claims that iris pattern could reflect one's health and reveal the state of individual organs. According to iridology, cholesterol in body can be detected if there is a

“sodium ring” in the patient’s eyes. However, since there were statements that regard iridology as medical fraudulent (L.Berggren, 1985), we were looking at other medical statements that can relate cholesterol and other organs. We found out that high cholesterol can be detected from changes in iris pattern and they are called *Arcus Lipoides* (*Arcus Senilis* or *Arcus Juvenilis*). “*Arcus senilis* is a greyish or whitish arc or circle visible around the peripheral part of the corner in older adults. *Arcus senilis* is caused by lipid deposits in the deep layer of the peripheral cornea and not necessarily associated with high blood cholesterol. However, similar discoloration in the eyes of younger adults (*arcus juvenilis*) is often associated with high blood cholesterol (K.Hughes et. al., 1992).” This statement proves that iris pattern can be analyzed and used as another technique to detect cholesterol presence in body.

(Harold z. Pomerantz, 1962), conclude in his study the presence of *Arcus Senilis* before the age of 56 and large wrist size were found to appear with a frequency in coronary group which made their presence statistically significant at level 5%. Hypercholesterolemia was common finding in coronary patient who demonstrated *Arcus Senilis* and greying of hair. According to (Jae-Young Um et. al, 2005) although iridology has been criticized as an unfounded diagnostic tool, many iridologists are presently practicing in many areas. In Germany, 80% of Heilpraktiker (non-medically qualified health practitioners) practice iridology (Ernst, 2000). In this study, (Jae-Young Um et. al, 2005) investigated the ACE genotypes of hypertensive patients classified by their iris constitutions. As a result, 74.7% of hypertensive patients were neurogenic or cardio-renal connective tissue weakness type. Also, the frequencies of DD genotype were significantly higher in hypertensive patients than in controls. These results are consistent with the reports that DD genotype was associated with hypertension (Staessen et al., 2001). Therefore, (Jae-Young Um et. al, 2005) present the results support that D allele is a candidate gene for hypertension, and suggest an apparent relationship between ACE genotype and iris constitutions, as well as the novel possibility of molecular genetics understanding of iridology.

## 2. Eye image

The eye is the organ of sight, a nearly spherical hollow globe filled with fluids (humors). The outer layer or tunic (sclera, or white, and cornea) is fibrous and protective. The middle tunic layer (choroid, ciliary body and the iris) is vascular. The innermost layer (the retina) is nervous or sensory. The fluids in the eye are divided by the lens into the vitreous humor (behind the lens) and the aqueous humor (in front of the lens). The lens itself is flexible and suspended by ligaments which allow it to change shape to focus light on the retina, which is composed of sensory neurons (NLM, 2010). Fig. 1 shows the anatomy of human eye which contain the area of sclera and iris for references. The iris image needs to be extract from the original eye image. This solid iris image will be used in this system to verify the presence of cholesterol. Thus it is vital to isolate this part (iris) from the whole unwanted part in the eye (sample). This separation or segmentation is the process of remove the outer part of the eye (outside the iris circle), in order to get solid image of iris that useful for localisation the cholesterol lipid. Generally this eye breaks up into two parts, the first part is the inner region which is the iris and pupil boundary and the second part is the outer regions, the iris and sclera boundary. The quality of the images is very important to get the best result, thus the images should not have any impurities that can cause miss localization. These impurities include the flash reflection from camera and wrong angle of image capture.

In this project the sample of eye is very vital because analysis base on the data from human eyes. The easier way to have these samples is by using free database source. These database are freely available iris images database, they give permission to all researcher, student and et cetera for research or educational purpose. There are a few free database sources that can found in website, these databases such as CASIA, MMU, UPOL, UBIRIS en cetera. To use these database someone need to write email or send the e-form to the authors and state the particular information such as name, status, organization and purpose to use database. Some of these databases have the restriction such as username and password, and these data will be given base on request by the users when they email the author for use their database.

### 2.1 UBIRIS

UBIRIS database is comprised of 1877 images collected from 241 subjects within the University of Beira Interior 6 in two distinct sessions and constituted, at its release date, the world's largest public and free iris database for biometric purposes.

### 2.2 CASIA

(CASIA, 2003), iris image database (version 1.0, the only one that we had access to) includes 756 iris images from 108 eyes, hence 108 classes. For each eye, 7 images are captured in two sessions, where three samples are collected in the first and four in the second session. Similarly to the above described database, its images were captured within a highly constrained capturing environment, which conditioned the characteristics of the resultant images. They present very close and homogeneous characteristics and their noise factors are exclusively related with iris obstructions by eyelids and eyelashes (Fig. 2). Moreover, the post process of the images filled the pupil regions with black pixels, which some authors used to facilitate the segmentation task.

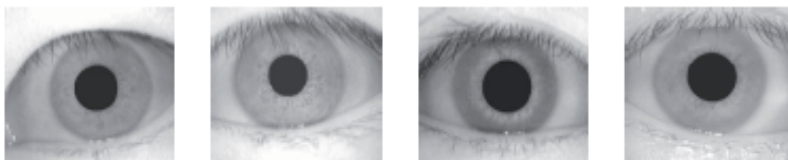


Fig. 2. Examples of iris images from the CASIA database.

### 2.3 MMU

The Multimedia University has developed a small data set of 450 iris images (MMU). They were captured through one of the most common iris recognition cameras presently functioning (LG IrisAccessR 2200). This is a semi-automated camera that operates at the range of 7-25 cm. Further, a new data set (MMU2) comprised of 995 iris images has been released and another common iris recognition camera (Panasonic BM-ET100US Authenticam) was used. The iris images are from 100 volunteers with different ages and nationalities. They come from Asia, Middle East, Africa and Europe and each of them contributed with five iris images from each eye. Obviously, the images are highly homogeneous and their noise factors are exclusively related with small iris obstructions by eyelids and eyelashes (Fig. 3).





Fig. 3. Examples of iris images from the MMU database.

## 2.4 Medical website

For this project the limitation to achieve the subject or patient to get real image of eye sample, the best place to get these real eye images is at ophthalmology department since this department deal with various case of eye problem that later can be refer to *Arcus Senilis* problem. For this reason, the samples only can be used from free source as from medical. The medical iris images can be obtain from TedMontgomery, National Library of Medicine, Mediscan clipart library licensed medical pictures and Mayo clinic and foundation for medical education and research medical and research training Fig. 4. show a few sample from above website.

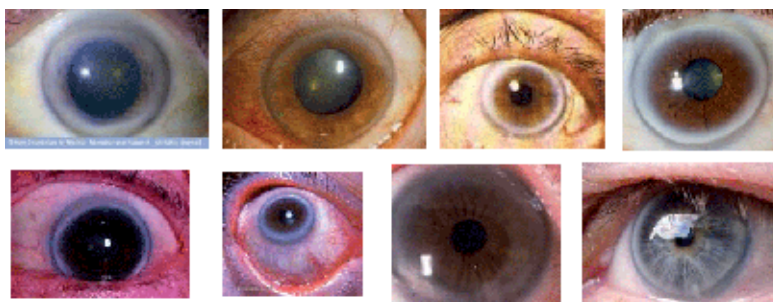


Fig. 4. Examples of iris images from medical website.

## 3. Iris recognition

Iris recognition is one of the most widely implemented biometric systems in use today. John Daugman is said to have developed the most widely used algorithms and most efficient methods of recognition, but there have been many new findings and algorithms (J. Daugman, 2004). (L. Masek, 2003) has verified the uniqueness of human iris patterns and developed “an open source” iris recognition system.

### 3.1 Hough transform

The Hough transform is a standard computer vision algorithm that can be used to determine the parameters of simple geometric objects, such as lines and circles, present in an image. The circular Hough transform can be employed to deduce the radius and centre coordinates of the pupil and iris regions. An automatic segmentation algorithm based on the circular Hough transform is employed (Theodore & Richard.W 2002). Firstly, an edge map is generated by calculating the first derivatives of intensity values in an eye image and then set the threshold base on the result. From the edge map, votes are cast in Hough space for the parameters of circles passing through each edge point. These parameters are the centre coordinates  $x_c$  and  $y_c$ , and the radius  $r$ , which are able to define any circle according to the equation

$$x_c^2 + y_c^2 + r^2 = 0 \quad (1)$$

A maximum point in the Hough space will correspond to the radius and centre coordinates of the circle best defined by the edge points. (Theodore & Richard.W 2002) and Kong and Zhang also make use of the parabolic Hough transform to detect the eyelids, approximating the upper and lower eyelids with parabolic arcs, which are represented as;

$$((-x-h_j)\sin\theta_j + (y-k_j)\cos\theta_j)^2 = a_j((x-h_j)\cos\theta_j + (y-k_j)\sin\theta_j) \quad (2)$$

$a_j$  controls the curvature,  $(h_j, k_j)$  is the peak of the parabola and  $\theta_j$  is the angle of rotation relative to the x-axis.

In performing the preceding edge detection step, (Theodore & Richard.W 2002) bias the derivatives in the horizontal direction for detecting the eyelids, and in the vertical direction for detecting the outer circular boundary of the iris. The motivation for this is that the eyelids are usually horizontally aligned, and also the eyelid edge map will corrupt the circular iris boundary edge map if using all gradient data. Taking only the vertical gradients for locating the iris boundary will reduce influence of the eyelids when performing circular Hough transform, and not all of the edge pixels defining the circle are required for successful localisation. Not only does this make circle localisation more accurate, it also makes it more efficient, since there are less edge points to cast votes in the Hough space.

### 3.2 Segmentation

This segmentation (localization) process is to search for the centre coordinates of the pupil and the iris along with their radius. These coordinates are marked as  $c_i, c_p$ , where  $c_i$  represented as the parameters of  $[x_c, y_c, r]$  of the limbic and iris boundary and  $c_p$  represented as the parameters of  $[x_c, y_c, r]$  of the pupil boundary. It makes use of (Theodore & RichardWildes 2002) method to select the possible centre coordinates first. The method consist of threshold followed by checking if the selected points (by threshold) correspond to a local minimum in their immediate neighborhood these points serve as the possible centre coordinates for the iris. Once the iris has been detected (using Daugman's method), the pupil's centre coordinates are found by searching a 10x10 neighbourhood around the iris centre and varying the radius until a maximum is found (J. Daugman, 2004). The input for this function is the image to be segmented and the input parameters in this function including  $rmin$  and  $rmax$  (the minimum and maximum values of the iris radius). The range of radius values to search for was set manually, depending on the database used. For the CASIA database (example, iris3.bmp),  $rmin$  is set to 55 pixels and  $rmax$  is set to 160 pixel.

The input for this function is the image to be segmented and the input parameters in this function including  $rmin$  and  $rmax$  (the minimum and maximum values of the iris radius). The range of radius values to search for was set manually, depending on the database used. For the CASIA database (example, iris3.bmp),  $rmin$  is set to 55 pixels and  $rmax$  is set to 160 pixels. The sample of *Arcus Senilis* eye (arcus7.bmp) is downloading from (Mediscan, 2000), for this image the  $rmin$  is set to 70 pixels and  $rmax$  is set to 260 pixels. Table 1 shows the others sample run for segmentation process.

| Image                    | rmin | rmax |
|--------------------------|------|------|
| Iris1.bmp (CASIA)        | 75   | 260  |
| Iris2.bmp (CASIA)        | 75   | 260  |
| Iris3.bmp (CASIA)        | 75   | 260  |
| Iris4.bmp (CASIA)        | 75   | 260  |
| Iris6.bmp (CASIA)        | 75   | 260  |
| c_mo1.bmp                | 50   | 83   |
| c_mo2.bmp                | 75   | 260  |
| normal1.bmp              | 75   | 260  |
| Arcus1.bmp (Medical Web) | 80   | 260  |
| Arcus2.bmp (Medical Web) | 80   | 260  |
| Arcus3.bmp (Medical Web) | 80   | 260  |
| Arcus4.bmp (Medical Web) | 80   | 260  |
| Arcus5.bmp (Medical Web) | 80   | 260  |
| Arcus6.bmp (Medical Web) | 80   | 260  |
| Arcus7.bmp (Medical Web) | 80   | 260  |
| Arcus2.bmp (Medical Web) | 80   | 260  |
| Arcus2.bmp (Medical Web) | 80   | 260  |
| Arcus2.bmp (Medical Web) | 80   | 260  |
| ubiris1.bmp (UBIRIS)     | 85   | 260  |
| ubiris2.bmp (UBIRIS)     | 85   | 260  |
| ubiris2.bmp (UBIRIS)     | 85   | 260  |

Table 1. Shows the example of the images and their rmin and rmax values.

The output of this function will be the value of  $c_p$  and  $c_i$ , which is the value of  $[x_c, y_c, r]$  for the pupillary boundary and the limbic / iris boundary and also the segmented image.

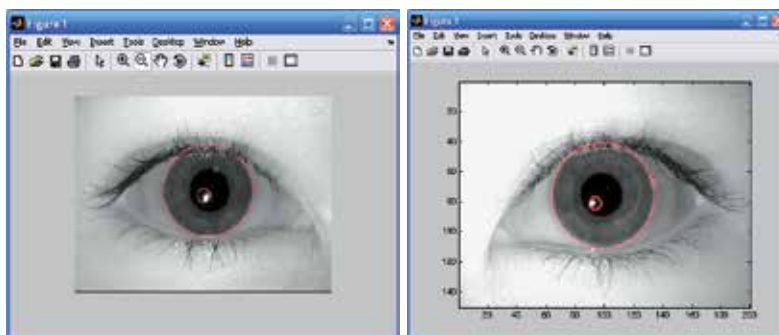


Fig. 5. An example where segmentation fails. The segmentation is failing to detection correctly the edges of the pupil border, but it segmented illumination light.

The program has be run on the image from CASIA as in the Fig. 5, this image give wrong detection on pupil boundary because segmentation on pupil is segmented on the illumination light rather than segmented the pupil boundary. Another example as shows in the Fig. 6, also fail to determine the edge of pupil but it detect the edge of the impurity illumination light. This will effect to the quality of the segmentation eyes image, cause to

imperfectly to detect iris and pupil boundary region of the eyes. But luckily the significant area of white ring (*Arcus Senilis*) lay at the boundary of sclera or iris up to the pupil, so as long as the segmentation done correctly on the iris it can considered succeed. This segmentation image will be crop base on the value of iris's radius.

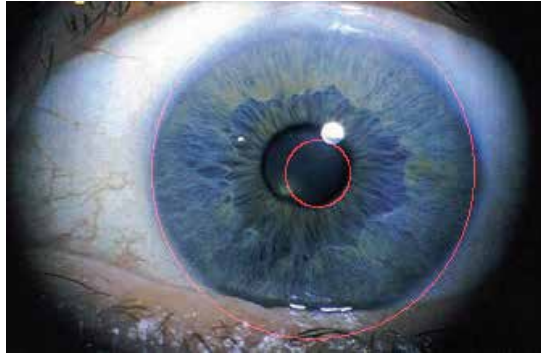


Fig. 6. Another example where segmentation fails. Canny edge detection fails to find the edges of the pupil border, but it takes edge of impurity of the illumination light.

The problem with this image is the existing of illumination from camera (on top of the pupil) will cause miss detection of iris and pupil boundary. This process consider failure if the purpose of this segmentation process is used for biometric iris recognition, but for this system since the image only need to be analyze in 30 percent from the limbic (border of iris) so this result can be used.

### 3.3 Normalization

After the iris is localized the next step is normalization (iris enrolment). It is a process after localization (segmentation) is to change the iris region to the fixed Fig. in order to make further analysis. From the process of normalization, the segmented image of the eye will give the value radius pupil and the iris. This image will be crop base on the value of iris radius, so that the unwanted area will be removing (eg sclera and limbic). Therefore only the intended area can be analyzed. According to (Frank L. Urbano, 2001), *Arcus Senilis*, or Corneal Arcus, is described as a yellowish-white ring around the cornea that is separated from the limbus by a clear zone 0.3 to 1 mm in width. It is caused by extracellular lipid deposition in the peripheral cornea, with the deposits consisting of cholesterol, cholesterol esters, phospholipids, and triglycerides. The fatty acids that make up many of the deposited lipid molecules include palmitic, stearic, oleic, and linoleic acids. Normally the area of white ring (*Arcus Senilis*), occurs from the sclera/iris up to 20 to 30 percents toward to pupil, so this is the only the main area that have to be analyzed. The other reason to normalize is to make the analysis process become easier rather than to examine the eye in circular shape. In rectangular shape analyze can be done either from top to bottom or from bottom to top.

(J. Daugman, 2004) describes details on algorithms used in iris recognition. He has introduced the Rubber Sheet Model that transforms the eye from circular shape into rectangular form and it is shown in Fig.7. This model remaps all point within the iris region to a pair of polar coordinates  $(r, \theta)$ , where  $\theta$  is the angle  $[0, 2\pi]$  and  $r$  is on the interval  $[0, 1]$ .

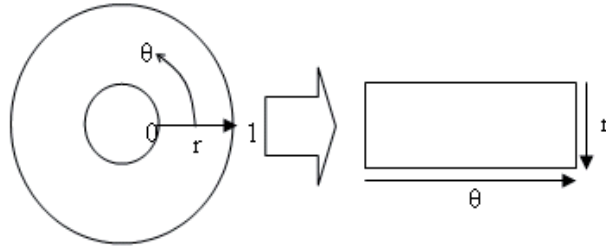


Fig. 7. Daugman's rubber sheet model.

The remapping of the iris image  $I(x, y)$  from Cartesian coordinates to the normalized non-concentric polar representation can be modelled as:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (3)$$

With

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_l(\theta) \quad (4)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_l(\theta) \quad (5)$$

Where  $I(x, y)$  is the iris region image,  $(x, y)$  are the original Cartesian coordinates,  $(r, \theta)$  are the corresponding normalised polar coordinates, and  $x_p, y_p$  and  $x_l, y_l$  are the coordinates of the pupil and iris boundaries along the  $\theta$  direction. The localization of the iris and the coordinate system is able to achieve invariance to 2D position and size of the iris, and to the dilation of the pupil within the iris.

The normalization process can be illustrated in Fig. 8. It is done by taking the reference point from the centre of the pupil and radial vectors pass via the iris area. There are two important data points along each radial line which are radial resolution for radial line in pupil and angular resolution for radial line around the iris region. Since the pupil can be non-matching with the iris therefore it need to remap to rescale the points depending to the angle around the iris and pupil. This formula given by:

$$r' = \sqrt{\alpha\beta} \pm \sqrt{\alpha\beta^2 - \alpha - r_l^2} \quad (6)$$

With

$$\alpha = o_x^2 + o_y^2 \quad (7)$$

$$\beta = \cos\left(\pi - \arctan\left(\frac{o_y}{o_x}\right) - \theta\right) \quad (8)$$

The shift of centre of the pupil relative to the iris centre, this given by  $O_x, O_y$ , and  $r'$  is the distance between edge of pupil and edge of the iris at the angle  $\theta$  around the region, and  $r_l$  is the radius of the iris. The remapping formula first gives the radius of the iris region 'doughnut' as a function of the angle  $\theta$ .

A constant number of points are chosen along each radial line, so that a constant number of radial data points are taken, irrespective of how narrow or wide the radius is at a particular angle. The normalised pattern was created by backtracking to find the Cartesian coordinates of data points from the radial and angular position in the normalised pattern. From the 'doughnut' iris region, normalisation produces a 2D array with horizontal dimensions of angular resolution and vertical dimensions of radial resolution.

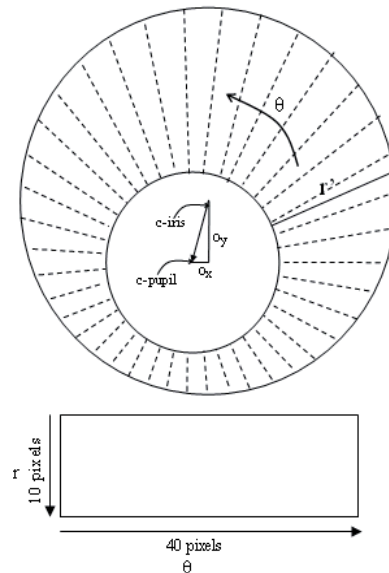


Fig. 8. Outline of the normalisation process with radial resolution of 10 pixels, and angular resolution of 40 pixels. Pupil displacement relative to the iris centre is exaggerated for illustration purposes.

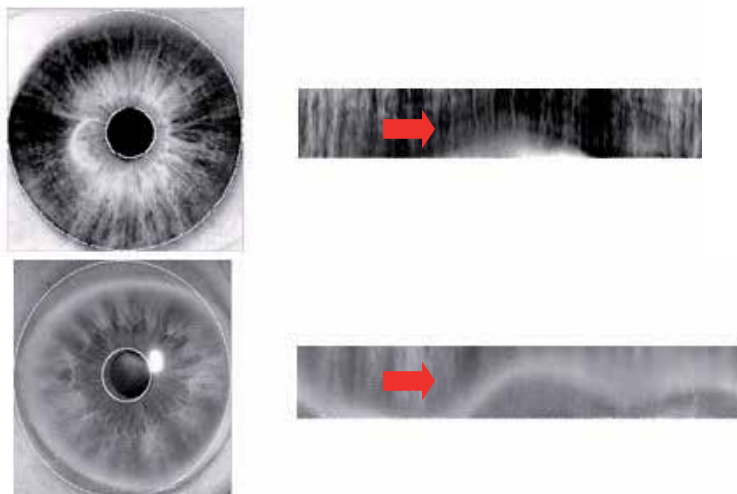


Fig. 9. Illustration of the normalization process for two images of the same iris taken under varying conditions. Top image normal eye, bottom image suspected illness eye.

The normalisation process proved to be successful and some results are shown in Fig 9. This normalization process will transform the segmented eye from the circular form to the rectangular shape. However normalisation output will be cropped until 30 percents, from the bottom eye (sclera and iris boundary) toward pupil.

Normalisation of two eye images of the same iris is shown in Fig 9. The pupil is smaller in the bottom image, however the normalisation process is able to rescale the iris region so that it has constant dimension. Note that rotational inconsistencies have not been accounted for by the normalisation process, and the two normalised patterns are slightly misaligned in the horizontal (angular) direction. Rotational inconsistencies will be accounted for in the matching stage.

It is difficult to do analysis if the image is in the original form therefore the image needs to be wrapped to transform the nature from circle to rectangular shape. This process only can be achieved by doing the conversion polar to rectangular.

The normalisation process proved to be successful and some results are shown in Fig. 9. However, the normalisation process was not able to perfectly reconstruct the same pattern from images with varying amounts of pupil dilation, since deformation of the iris results in small changes of its surface patterns.

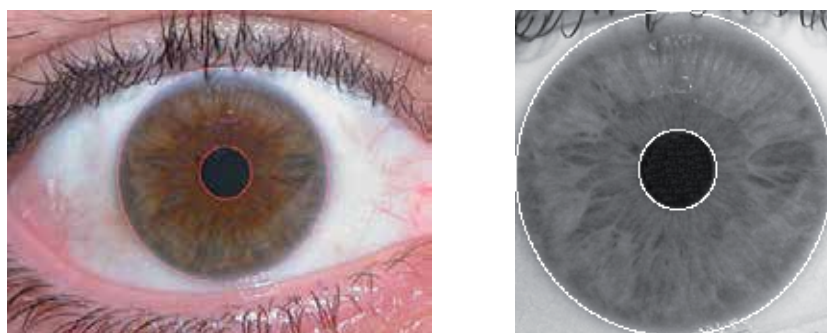


Fig. 10. Stages of localization with eye image 'ubiris1.bmp (340x260)' from the UBIRIS database left) original color eye image localization right) black and white (in gray) eye image localization.

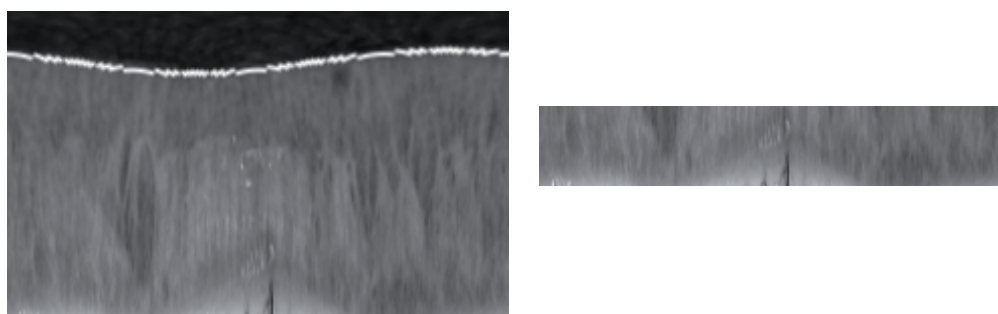


Fig. 11. Stages of normalization with eye image 'ubiris1.bmp (340x260)' from the UBIRIS database left) 100 percents display from polar to rectangular of localization eye. (Right) 30 percents display from polar to rectangular of localization eye, after crop process.

The normalization process is used for converting the circular iris into rectangular form with fixed dimension as shown in Fig 12. We can see clearly the circle shape (Localisation) turn to rectangular shape (normalisation) and also the signs in both pictures are labelled to illustrate the upper eyelid, pupil and white dot exit during segmentation and after normalisation.

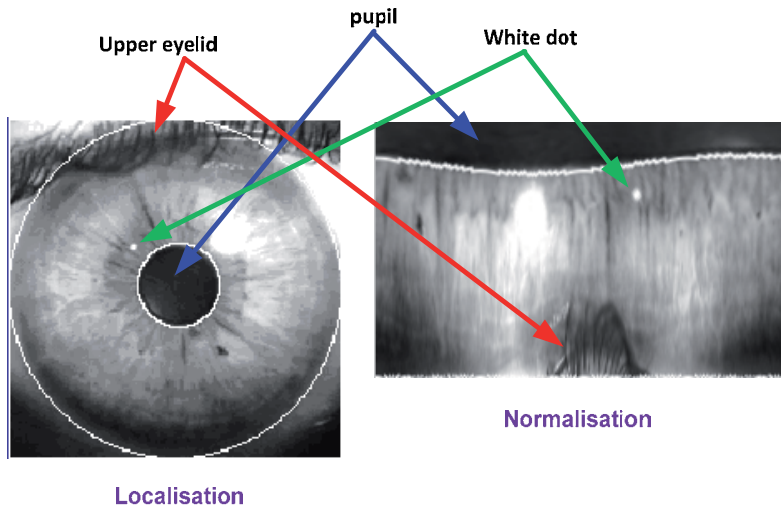


Fig. 12. Stages of normalization showing the same part in both shape circular and rectangular in the eye.

#### 4. Cholesterol detection system

The Daugman's Rubber Model is the main process in the cholesterol detection system proposed in this paper. The process starts with obtaining number of normal eyes images and images of eyes with Arcus Lipids. The normal eyes are available from UBIRIS databases while the Arcus Lipids eyes are taken from iridology clinic.

The next step is to isolate the actual iris region in digital eye image. The isolation process needs to be done to segment the outer boundary for the iris and the inner boundary for the pupil. This can only be done by searching the centre point of the pupil given by x and y axis. Hough transform is used to detect edge of the iris and pupil circle.

Next, the image has to be analyzed and this can only be done if it is transformed to normalized polar coordinates using Rubber Model. Since the "sodium ring", terminology given in iridology, or Arcus Lipids for the greyish or whitish arc in iris is only available at the bottom of this coordinate, thus only 30% of the iris part is considered in the normalization.

Lastly, to determine whether the eye has the ring, histogram of the image has to be plotted so that the decidability can be determined using OTSU's method. The algorithm assumes the image contains two classes of pixels (e.g. foreground and background) and finds the optimum threshold separating the two classes so that their combined spread (within-class variance) is minimal.



## 5. Results

Fig 13 shows the whole process of the cholesterol detection system using iris recognition and image processing algorithm comprises the following actions:

- Eye images acquire from database (CASIA, UBIRIS, MMU and medical web) or from digital camera.
- Process of pupil and iris localization and segmentation, to classify the required region.
- Attain normalization iris from circular shape to rectangular shape with full image (100%)
- Crop the normalization iris to 30% from full image.
- Run the normalization iris to get the histogram value.
- Using OTSU to calculate the optimum threshold to detect Arcus Lipids (Cholesterol presence).
- Results “Sodium ring detected” or “not detected” will be display in MATLAB window.

This experiment used the eye images from free available database sources that can be downloaded by permission of the author of the database. The eye images also can be taken from high quality digital camera where the subject should be categorised to two groups of people; the healthy people and the suspected people with cholesterol symptom. These databases are free and can be used for research and educational purpose. Amongst these databases such as Chinese Academy of science (CASIA), Multimedia University (MMU), UBIRIS, and UPOL. And for medical images it can be used from TedMontgomery, National Library of Medicine, Mediscan clipart library licensed medical pictures and Mayo clinic and foundation for medical education and research medical and research training. For these database the eye images can be categories to two groups, first group provide grey colour images such as from CASIA and MMU database, while the second group such UBIRIS and almost all the medical website provide colour images. The reason to use eye image from these databases because these images are easy to be achieved without need to find eye images from any subject (patients).

The next process is *localization and segmentation* of pupil and iris this is to classify the required region. In this process two important regions need to be segmented which are pupil (the inner circular, normally black circular) and iris (the outer part, pattern and colour area). The unused part such as sclera (the white part outside the iris) will be removed because this region is not aimed for this experiment.

*Eye image crop, base on iris radius*; after segmentation process (the image is converted to grey scale), the next process is to cropped the eye image base on the value of iris radius. This iris radius value is obtained from the process of segmentation. In the segmentation process three important parameters are achieved, such as circular pupil parameters ( $x$ ,  $y$ , and  $r$ ) for pupil and circular iris ( $x$ ,  $y$  and  $r$ ) for iris. To crop this grey scale eye image the  $x$ ,  $y$  and  $r$  value from the iris's radius are needed because this is the outer and the biggest circular in this eye image. The result from this process will be obviously eye image contain only iris and pupil image, that can be used for next process of normalization process.

*Normalization process*; this process involve transforming the shape of image achieve from segmentation process (cropped eye image), where the segmentation eye image in circular shape will be change to rectangular shape, as illustrate in Fig 8. The reason to change this shape is to make the analysis can be carried out easily compare to original shape (circular shape). If the analysis is performed in circular shape the analysis must be done around the

circle of the eye and this will be so difficult, while analysis upon rectangular shape is more practical and easy to be done. This method transforming to rectangular shape, is been introduced by Professor John Daughman in his paper under Rubber's shape model, which this method analysis is performed upon the rectangular eye image from the bottom image up to the top of the image or vice versa. This normalized process will be transformed all area of circular shape of the eye image to become rectangular shape and the process will be mark as 100 percent normalization process output. According to Frank L. Urbano, in his medical journal (Hospital Physician Past Article 2001), the sign of cholesterol deposits can be seen 30 percents from bottom (border of sclera) of the eye image. As referring to this statement the 100 percents normalization eye image will be crop to until 30 percents remaining. This process still normalization image, where the cropping taken the part from the pupil to a middle part of iris up to 70 percents area involve. The reason of this process because the white ring (cholesterol's deposit) usually affect in this area which is 30 percents from the bottom of sclera or iris region up to the top of the pupil and iris border.

The 30 percents normalization image furthermore is the main area where the analysis needs to be performed. For this thesis the analysis is carried out by using histogram function. The brightness of the image will be shown in bar chart of the histogram graph where the left side of x axis will represent the by dark area showing low brightness, and the right side of x axis will is represent by high level of brightness. The Histogram output than will be apply to OTSU threshold used to determine the threshold value of the image being analyzed. This OTSU threshold will give the average level of brightness in the eye image based on distribution of bar chart histogram from the output of histogram process. The OTSU graph will display and marked the threshold area for showing the threshold value.

Result from MATLAB later will display the string "Sodium ring is detected" or "no sodium ring is detected" depends on the eye image that be examined. For this experiment the boundary value of threshold in the examined eye image is set to 139 (this was decide after run 50 samples of the eye image), the average boundary of the illness or detected eye problem with sign of the presence cholesterol deposited. If the threshold value fallen below this value the eye image is considered as normal (no existing white ring or cholesterol), but if the threshold value rise up beyond this value (139), the subject or patient is detected a sign of the presence of cholesterol. In MATLAB window display will popup the message showing the result, base on value obtained in automated detecting cholesterol presence (ADCP), written in MATLAB m-file programming. Results indicate the range of normal and abnormal eye, based on the threshold value obtain from ADCP. This will determine either someone have the symptom of the cholesterol presence or not. The result however is display in command window, but yet this program can be run and display using Graphic User Interface (GUI), where MATLAB has tool to perform it.

Normalization process is shown in Fig. 14, where in this process the image has been localized and transformed from polar to rectangular using gray image of normal eye. The transformation has to be cropped up to 30 percents from sclera/iris toward pupil because this is the area the *Arcus Senilis* normally exists.

Fig 15 shows the histogram, threshold values and the statement about the condition of the normal and *Arcus Senilis* iris, respectively. From the histogram and OTSU's method, the decidability or threshold value to distinguish between normal eyes and eyes with *Arcus Lipids* is found to be 139. This value is determined after testing 30 images of normal eyes. If the cluster mean value is less than this threshold, this means than the eye is normal eye and if it is above 139, then the eye can be detected as eye with *Arcus Lipids*.

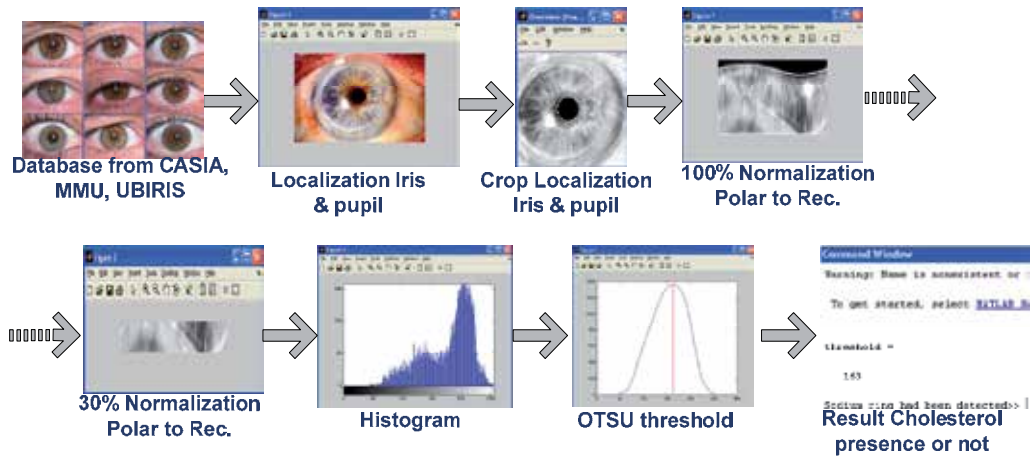


Fig. 13. Overall system for Cholesterol Detection.

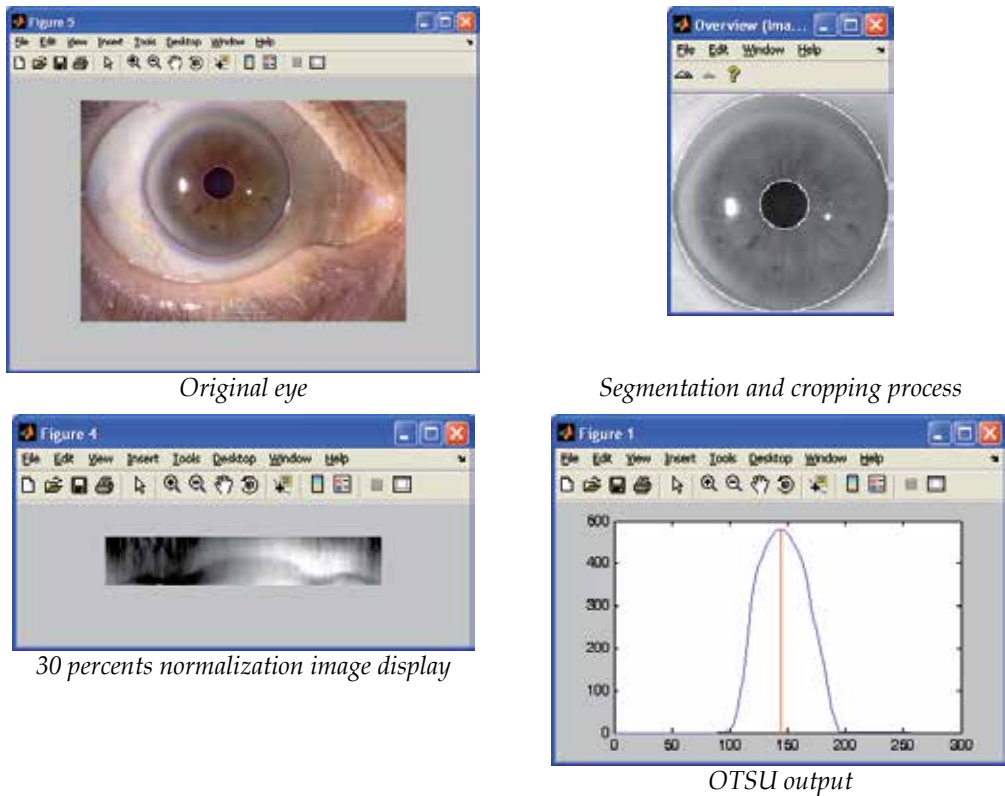
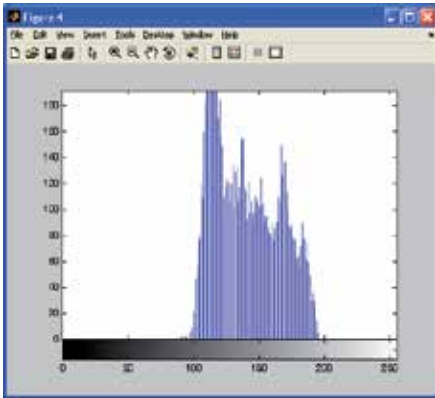


Fig. 14. Stages of localization with eye image 'Arcus1.bmp' from the medical web, (Clock wise from top left) original colour eye image localization, the iris and pupil detected correctly. (Top right) Gray eye image localization (Bottom left) 30 percents display from polar to rectangular of localization eye with enhancement. (Bottom right) OTSU threshold value for this eye is 144.



*Histogram output*



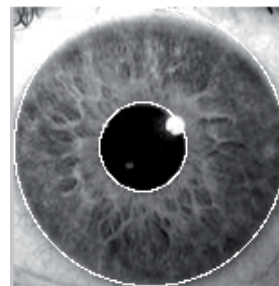
*Final result*

Fig. 15. Results from eye with “sodium ring’ or *Arcus senilis* i.e. ‘arcus1.bmp’: histogram, threshold value and statement of iris condition.

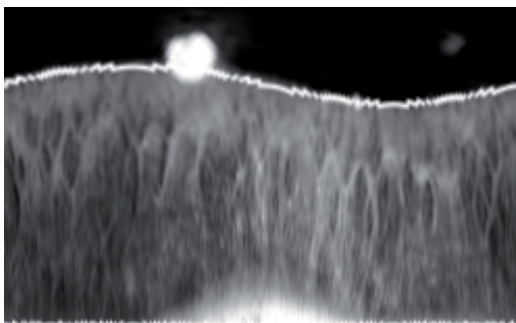
Another result for cholesterol presence detection in normal eye shown as in Fig. 16 below, where the eye images in this figure shown variety of process involve from original eye until analysis to determine cholesterol presence.



*Original eye*



*Segmentation and cropping process*



*100 percents normalization image display*



*30 percents normalization image display*

Fig. 16. Continued.

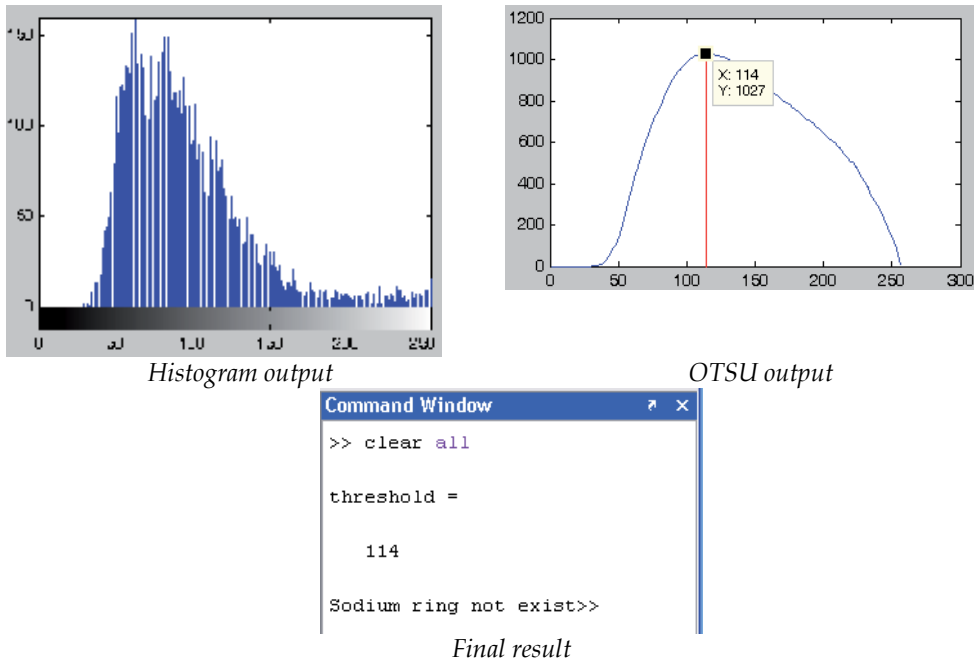


Fig. 16. Results from normal eye.

The output from this experiment shown in Fig. 16 indicate the eye image have no symptom of cholesterol presence this shown in command window result "Sodium ring not exist" and the threshold value give 114 which is below the set value (139). This meant the test eye image contain no symptom of cholesterol presence.

Another result from abnormal eye image is shown in Fig.17 below. The process follows the same method as describe in above procedure. For this image the threshold value determine is 148 which is higher than the set point 139 thus the result in command window display the message "Sodium ring had been detected", this indicate the eye encompass of cholesterol presence symptom.

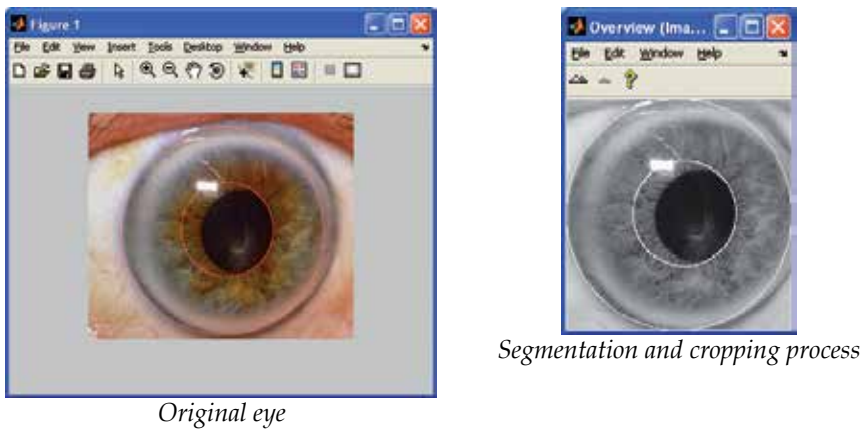
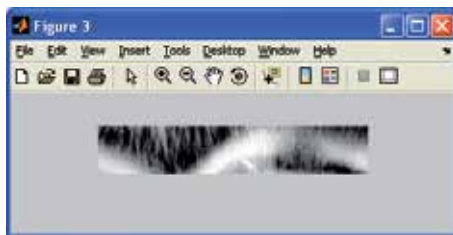
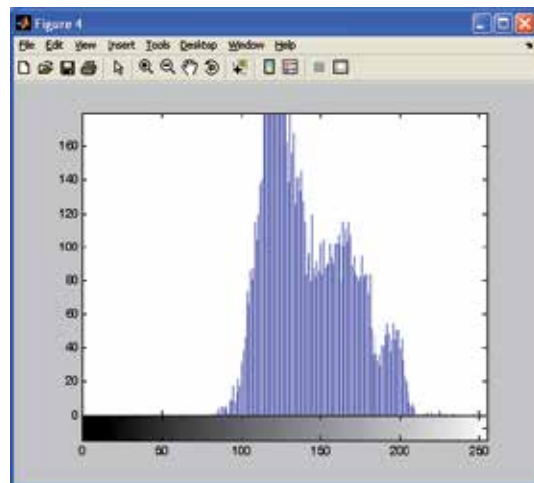


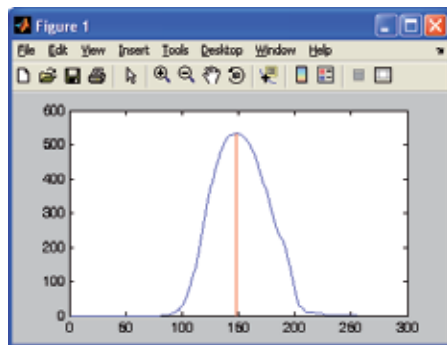
Fig. 17. Continued.



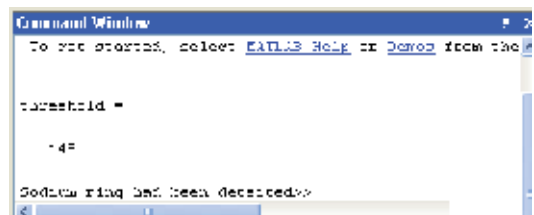
*30 percents normalization image display*



*Histogram output*



*OTSU output*



*Final result*

Fig. 17. Results from normal eye.

## 6. Conclusion

This work shows that there is a simple and non-intrusive method to detect cholesterol in body and iris recognition is not only mainly for biometric identification but it can also be used as a mean to detect cholesterol or maybe diagnose any diseases as iridology claimed it is supposed to be. However, this work is only preliminary work and experiments that are more extensive need to be run in future in order to know the real level of cholesterol in the body. This program had been executed on more than 50 samples of normal and abnormal eye images; it can be concluded that the threshold boundary of the normal and problem eye is about 139. This project had shown the entire process of detecting cholesterol presence using automated program (ADCP). However this programs written in m-file and the result display in command window. The improvement can be done such as the using GUI for execute and displaying the result and using other method for determines the threshold of the normal or problem eye. Others application that can use this program is to determine the eye problem due to other type of eye diseases such as cataract, glaucoma, diabetic, tumour etcetera.

## 7. References

- [1] (N.Haq, M.D.Fox, 1991), N.Haq, M.D.Fox, "Preliminary results of IR spectroscopic characterization of cholesterol," Bioengineering Conference, 1991., Proceeding of the 1991 IEEE Seventeenth annual Northeast, pp.175-176, 4-5 Apr 1991.
- [2] (FDA, 2004), U.S.Food and Drug Administration(FDA), "FDS Clears New Palm Test For Skin Cholesterol", FDA Talk Paper, June 24<sup>th</sup>, 2004, available at <http://www.fda.gov/bbs/topics/ANSWERS/2002/ANS01154.html>
- [3] (L.Berggren, 1985), L.Berggren, "Iridology: a critical review" Acta Ophthalmol, 63,pp.1-8,1985.
- [4] (K.Hughes et. al., 1992), K.Hughes, K.C.Lun, S.P.Sothy, A.C. thai, W.P. Leong, P.B. Yeo, "Corneal arcus and cardiovascular risk factors in Asians in Singapore,"Int. Journal Epidemiol,vol.21, pp.473-147, 1992.
- [5] (J. Daugman, 2004), J. Daugman. "How Iris Recognition Works," IEEE Transaction on Circuits and Systems for Video Technology, vol. 14, No.1, Jan 2004.
- [6] (L. Masek, 2003) L. Masek, "Recognition of Human Iris Patterns for Biometric Identification", Dissertation, University of Western Australia, 2003.
- [7] (CASIA, 2003), Chinese Academy ,Chinese Academy of Sciences - Institute of Automation. Database of 756 Greyscale Eye Images. <http://www.sinobiometrics.com> Version 1.0, 2003.
- [8] (MMU) Multimedia University -database contributes of 450 iris images <http://pesona.mmu.edu.my/~ccte/>
- [9] UBIRIS Iris Image Databse, available at <http://iris.di.ubi.pt/>
- [10] (UPOL) UPOL iris database, available at ["http://phoenix.inf.upol.cz/iris/download/"](http://phoenix.inf.upol.cz/iris/download/), update: 2003
- [11] (NLM, 2010), National Library of Medicine - National Institutes of Health- world's largest medical "<http://www.nlm.nih.gov/>" update: 2010
- [12] (Theodore, RichardWildes 2002)Theodore A. Camus, RichardWildes, "Reliable and Fast Eye Finding in Close-up Images", This work was sponsored by the Defense Advanced Projects Agency.The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the U.S. Defense Advanced Projects Agency, the U.S. Army Intelligence Center & Fort Huachuca, Directorate of Contracting Office, or the U.S. Government.
- [13] (Mediscan, 2000), Medical pictures and images -Mediscan clipart library Licensed medical pictures <http://www.mediscan.co.uk/>
- [14] N.OTSU, *A threshold selection method from gray-level histograms*. IEEE Trans. on System, Man and Cybernetics, 9(1):62--66, 1979
- [15] (Harold z. Pomerantz, 1962) Harold z. Pomerantz, M.D., Montreal, "The relationship Between Coronary Heart Disease and the Presence of Certain Physical Characteristics", Departments of Medicine and Cardiology, Reddy Memorial Hospital, Montreal, Canad. Med. Ass. J. Jan. 13, 1962, vol.86
- [16] (Jae-Young Um et. al, 2005) Jae-Young Um,\* Nyeon-Hyoung An,et. al, "Novel Approach of Molecular GeneticUnderstanding of Iridology: RelationshipBetween Iris Constitution and Angiotensin Converting Enzyme Gene Polymorphism", The American Journal of Chinese Medicine, Vol. 33, No. 3, 501-505, 2005 World

Scientific Publishing Company Institute for Advanced Research in Asian Science and Medicine, 2005.

- [17] (Frank L. Urbano, 2001) Frank L. Urbano, MD, "Ocular Signs of Hyperlipidemia", review of clinical signs, general internal medicine, Mount Laurel Primary Care Associates, Mount Laurel, NJ., Hospital Physician, November 2001
- [18] Staessen, J.A., J.G. Wang, E. Brand, C. Barlassina, W.H. Birkenhager, S.M. Herrmann, R. Fagard, L. Tizzoni and G. Bianchi. The deletion/insertion polymorphism of the angiotensin converting enzyme gene and cardiovascular-renal risk. *J. Hypertens.* 19: 1349-1358, 2001.
- [19] Ernst, E. Iridology: not useful and potentially harmful. *Arch. Ophthalmol.* 118: 120-121, 2000.



# Robust Feature Extraction and Iris Recognition for Biometric Personal Identification

Rahib Hidayat Abiyev and Kemal Ihsan Kilic  
*Department of Computer Engineering, Near East University, Nicosia,  
North Cyprus*

## 1. Introduction

Humans have distinctive and unique traits which can be used to distinguish them from other humans, acting as a form of identification. A number of traits characterising physiological or behavioral characteristics of human can be used for biometric identification. Basic physiological characteristics are face, facial thermograms, fingerprint, iris, retina, hand geometry, odour/scent. Voice, signature, typing rhythm, gait are related to behavioral characteristics. The critical attributes of these characteristics for reliably recognition are the variations of selected characteristic across the human population, uniqueness of these characteristics for each individual, their immutability over time (Jain et al.,1998). Human iris is the best characteristic when we consider these attributes.

The texture of iris is complex, unique, and very stable throughout life. Iris patterns have a high degree of randomness in their structure. This is what makes them unique. The iris is a protected internal organ and it can be used as an identity document or a password offering a very high degree of identity assurance. Also the human iris is immutable over time. From one year of age until death, the patterns of the iris are relatively constant (Jain et al.,1998, Adler,1965). Because of uniqueness and immutability, iris recognition is one of accurate and reliable human identification technique.

Nowadays biometrics technology plays important role in public security and information security domains. Iris recognition is one of the most reliable and accurate biometrics that plays an important role in identification of individuals. The iris recognition method deliver accurate results under varied environmental circumstances. Iris is the part between the pupil and the white sclera. The iris texture provides many minute characteristics such as freckles, coronas, stripes, furrows, crypts (Adler,1965). These visible characteristics are unique for each subject.

Iris recognition process can be separated into these basic stages: iris capturing, preprocessing and recognition of the iris region. Each of these steps uses different algorithms. Pre-processing includes iris localization, normalization, and enhancement. In iris localization step, the detection of the inner (pupillary) and outer (limbic) circles of the iris and the detection of the upper and lower bound of the eyelids are performed. The inner circle is located on the iris and pupil boundary, the outer circle is located on the sclera and iris boundary. Today researchers follow different methods in finding pupillary and limbic

boundaries. For these methods, the important problems are the accuracy of localization of iris boundaries, preprocessing speed, robustness (Daugman, 1994, Noh et al., 2003). For the iris localization step, several techniques have been developed. Daugman used integro-differential operator to detect inner and outer boundaries of the iris (Daugman, 2001; Daugman, 2003; Daugman, 2006). First derivative of image intensity and Hough transform methods are proposed for localization of iris boundaries in (Wiels, 1995). Using edge detection, the segmentation of iris is performed in (Boles & Boashash, 1998). Circular Hough transform is applied to find the iris boundaries in (Masek, 2003). In (Ma et al., 2002) grey-level information, Canny edge detection and Hough transform are applied for iris segmentation. In these researches iris locating algorithm based on the grey distribution of the features is presented. (Tisse et al., 2002) used a method that can locate a circle given three non-linear points. This is used for finding relatively accurate location of the circle, then the gradient decompose of Hough transform is applied for the accurate location of pupil. (Kanag & Xu, 2000) used Sobel operator and applied circle detection operator to locate the pupil boundary. In (Yuan et al., 2005) using threshold value, Canny operator and Hough transform the inner circle of iris is determined, then edge detection operator is used for outer circle. Pupil has dark colour, but in certain non ideal conditions because of specular reflections it can be illuminated unevenly. For such irregular iris images, researchers used intensity based techniques for iris localization (Daugman, 2007; Vatsa et al., 2008). (Vatsa et al., 2008) used two stage iris segmentation algorithm. In the first stage, iris boundaries are estimated by using elliptical model. In the second stage Mumford-Shah function is used to detect exact boundaries. Authors of (Zuo et al., 2008) explained robust automatic segmentation algorithm. One of the most important characteristics of iris localization system is its processing speed. Sometimes for the iris localization it may not be possible to utilize any method involving floating-point processing. For example if we have small embedded real-time system without any floating-point processing part, operations involving kernel convolutions will be unusable, even if they may offer more accurate results. Detailed discussions on the issues related to performance of segmentation methods can be found in (Liu et al., 2005, Cui et al., 2004, Abiyev & Altunkaya, 2008, 2009). Iris localization generally takes nearly more than half of the total processing time in the recognition systems, as pointed in (Cui et al., 2004). With this vision, we develop fast segmentation algorithms that require not many floating-point processing like convolution.

After the preprocessing stage feature vectors of iris images are extracted. These feature vectors are used for recognition of irises. Image recognition step basically deals with the "meaning" of the feature vectors, in which classification, labelling, matching occur. Various algorithms have been applied for feature extraction and pattern matching processes. These methods use local and global features of the iris.

Among many methods used for recognition today, these can be listed: Phase based approach (Daugman, 2001; Daugman, 2003; Daugman & Downing, 1994, Miyazawa et al., 2008), wavelet transform, zero crossing approach (Boles & Boashash, 1998; S. Avila & S. Reillo, 2002; Noh et al., 2002; Mallat, 1992), and texture analysis (Wiels, 1997; Boles & Boashash, 1998; Ma et al., 2003; Park et al., 2003; Ma et al., 2005). (Wang & Han, 2003, 2005) proposed independent component analysis is for iris recognition.

In phase based approach, the task of iris feature extraction is a process of phase demodulation. The phase is chosen as an iris feature. Daugman used multiscale quadrature wavelets to extract texture phase structure information of the iris to generate a 2,048-bit iris code and compared the difference between a pair of iris representations by

computing their Hamming distance. Boles and Boashash (1998) calculated a zero-crossing representation of 1D wavelet transform at various resolution levels of a concentric circle on an iris image to characterize the texture of the iris. Zero-crossings of wavelet transform provide meaningful information of images structure (Mallat,1992). S.Avila and S.Reillo (2002) further developed the method of Boles and Boashash by using different distance measures for matching. Many texture analysis methods can be adapted to recognize the iris images (Wildes,1997; Ma et al.,2003). Wildes represented the iris texture with a Laplacian pyramid constructed with four different resolution levels and used the normalized correlation to determine whether the input image and the model image are from the same class. (Ma et al., 2002) used multichannel Gabor filters to extract iris features. Future bank of circular symmetric filters was designed to capture the discriminating information along the angular direction of the iris image (Ma et al., 2003). (Tisse et al., 2002) uses a directional filter bank in order to decompose the iris images into eight directional subband outputs. In (Ma et al.,2005), iris recognition algorithm based on characterizing local variations of image structures was utilized. In (Lim et al.,2001) moment based iris blob matching algorithm is proposed and combination of this algorithm with local feature based classifiers is considered. (Abiyev & Altuknka,2008; Scotti, 2007) applied Neural Networks (NN) for iris pattern recognition.

Analysis of previous research demonstrated that the efficiency of personal identification system using iris recognition is determined by the fast and accurate recognition of the iris patterns. Great portion of the recognition time is spent on the localization of iris boundaries. Accuracy of recognition depends on the accuracy of iris localization and on the accuracy of classification. Some methods for iris localization are based on selecting threshold value for detecting pupil boundaries. Experiments on the CASIA Version 1 and Version 3 iris databases showed to us that segmentation algorithms sometimes may require intervention from researchers, for example fixing a certain threshold value for all the images. But this does not give good results, since many iris images captured under different illumination conditions and some of them contain noises. For example CASIA Version 1 and CASIA Version 3 databases have different illumination properties and CASIA Version 3 database iris images contain specular highlights in the pupil region. Existence of specular highlight poses difficulties for any algorithm that assumes pupil pixels have form of disc with the darkest pixels in the central region of the image. In this study we summarized our efforts on developing robust feature extraction methods. Two methods are proposed for feature extraction. The algorithms can find the image specific segmentation parameters from iris images. The first method is simple and fast method developed for high quality iris images (CASIA Version 1) where there is no specular highlight. Here black rectangle is used for pupillary boundary detection and gradient search method for limbic boundary detection. The second method is adaptive to degraded quality and can also perform iris segmentation in the presence of specular highlight (CASIA Version 3). Here Hough circle search for pupillary boundary detection and gradient search method for limbic boundary detection. After segmentation the features are extracted and used for recognition of irises. With the development of NN, iris recognition systems may gain speed, accuracy, learning ability. In this paper fast iris localization algorithm and NN based iris classification are proposed.

The rest of the paper is organized as follows: Section 2 includes the structure of iris recognition system, iris localization, normalization and enhancement algorithms. Section 3 is about iris classification using neural network. Section 4 presents experimental results of implemented algorithms. Section 5 includes conclusions.

## 2. Image processing

### 2.1 Structure of the iris recognition system

The block diagram of the iris recognition system is given in Fig.1. The image recognition system includes iris image acquisition and iris recognition. The iris image acquisition includes the lighting system, the positioning system, and the physical capture system (Wieldon,1997). The iris recognition includes preprocessing and neural networks. During iris acquisition, the iris image in the input sequence must be clear and sharp. Clarity of the iris's minute characteristics and sharpness of the boundary between the pupil and the iris, and the boundary between the iris and the sclera affects the quality of the iris image. A high quality image must be selected for iris recognition. In iris pre-processing, the iris is detected and extracted from an eye image and normalized. Normalized image after enhancement is represented by the feature vector that describes gray-scale values of the iris image. For classification neural network is used. Feature vector becomes the training data set for the neural network. The iris recognition system includes two operation modes: training mode and online mode. At first stage, the training of recognition system is carried out using greyscale values of iris images. After training, in online mode, neural network performs classification and recognizes the patterns that belong to a certain person's iris.

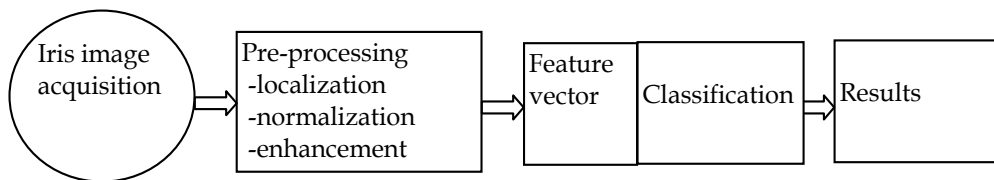


Fig. 1. A block diagram of the iris recognition system.

### 2.2 Iris localization

An eye image contains not only the iris region but also some parts that need to be separated from iris, such as the pupil, eyelids, sclera. For this reason, at the first step, segmentation should be done to localize and extract the iris region from the eye image. Iris localization is the detection of the iris area between pupil and sclera. So we need to detect the upper and lower boundaries of the iris and determine its inner and outer circles (Fig.2). Two different algorithms for localization of iris patterns are presented here. The first algorithm is called black-rectangle algorithm that uses grey values for segmentation, second algorithm uses Otsu thresholding and Hough circle search for pupillary boundary detection and gradient search method for limbic boundary detection.

A number of algorithms has been developed for iris localization. Most of them are based on the Hough transform (Wieldon,1997; Masek,2003). In these researches the canny edge detection algorithm with circular Hough transform is applied to detect the inner and outer boundaries of the iris. The circular Hough transform is employed to reveal the radius and centre coordinates of the pupil and iris regions. In this operation, starting from the upper left corner of iris, the circular Hough transform is applied. This algorithm is used for each inner and outer circle separately. The application of the Hough transform requires a long time to locate the boundaries of the iris and need fixing a certain threshold value for all the images. But this does not give good results, since many iris images captured under different

illumination conditions and some of them contain noise. In the paper two algorithms are proposed for iris localization. The first method is simple and fast method developed for high quality iris images (CASIA Version 1) where there is no specular highlight. Here black rectangle is used for pupillary boundary detection and gradient search method for limbic boundary detection. The second method is adaptive to degraded quality and can also perform iris segmentation in the presence of specular highlight (CASIA Version 3). Here Hough circle search is applied for pupillary boundary detection and gradient search method for limbic boundary detection. After segmentation the features are extracted and used for recognition of irises.

**Black-rectangle algorithm.** As mentioned iris localization includes detecting the boundaries between pupil and iris and also sclera and iris. To find the boundary between the pupil and iris, we must detect the location (centre coordinates and radius) of the pupil. The black rectangular technique that is applied in order to localize pupil and detect the inner circle of iris is given in algorithm 1. The pupil is a dark circular area in an eye image. Besides the pupil, eyelids and eyelashes are also characterized by black colour. In some cases, the pupil is not located in the middle of an eye image, and this causes difficulties in finding the exact location of the pupil using point-by-point comparison on the base of threshold technique. In this paper, we are looking for the black rectangular region in an iris image (Fig. 3). Choosing the size of the black rectangular area is important, and this affects the accurate determination of the pupil's position. If we choose a small size, then this area can be found in the eyelash region. In this paper a (10x10) rectangular area is used to accurately detect the location of the pupil (step 1 of algorithm 1). Searching starts from the vertical middle point of the iris image and continues to the right side of the image (see step 2 of algorithm 1). A threshold value is used to detect the black rectangular area. Starting from the middle vertical point of iris image, the greyscale value of each point is compared with the threshold value. As it is proven by many experiments, the greyscale values within the pupil are very small. So a threshold value can be easily chosen. If greyscale values in each point of the iris image are less than the threshold value, then the rectangular area will be detected. If this condition is not satisfactory for the selected position, then the search is continued from the next position. This process starts from the left side of the iris, and it continues until the end of the right side of the iris. In case the black rectangular area is not detected, the new position in the upper side of the vertical middle point of the image is selected and the search for the black rectangular area is resumed. If the black rectangular area is not found in the upper side of the eye image, then the search is continued in the down side of image. In Fig. 3(a), the searching points are shown by the lines. In Fig. 3(b), the black rectangular area is shown in white colour. After finding the black rectangular area, we start to detect the boundary of the pupil and iris. At first step, the points located in the boundary of pupil and iris, in horizontal direction, then the points in the vertical direction are detected (Fig. 4). In Fig. 4 the circle represents the pupil, and the rectangle that is inside the circle represents the rectangular black area. The border of the pupil and the iris has a much larger greyscale change value. Using a threshold value on the iris image, the algorithm detects the coordinates of the horizontal boundary points of  $(x_1, y_1)$  and  $(x_1, y_2)$ , as shown in Fig. 4. The same procedure is applied to find the coordinates of the vertical boundary points  $(x_3, y_3)$  and  $(x_4, y_3)$ . After finding the horizontal and vertical boundary points between the pupil and the iris, the following formula is used to find the centre coordinates  $(x_p, y_p)$  of the pupil.

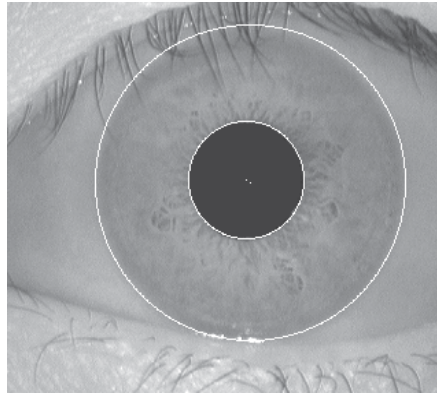


Fig. 2. A localised iris image.

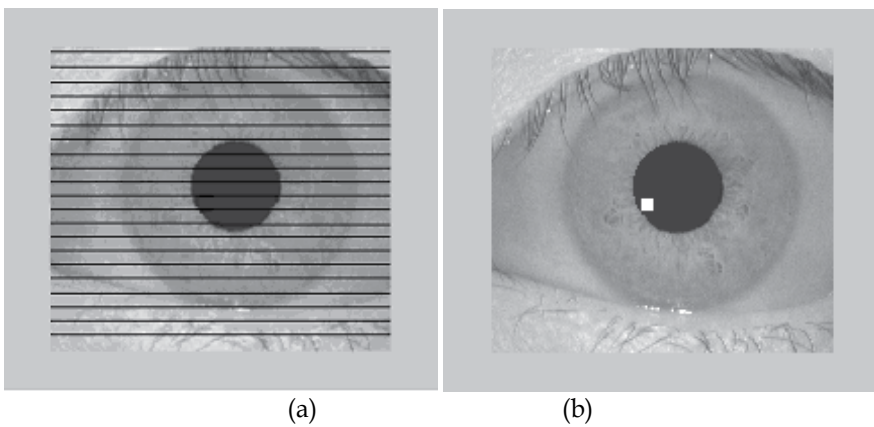


Fig. 3. Detecting the rectangular area: a) The lines that were drawn to detect rectangular areas, b) The result of detecting of rectangular area.

$$x_p = (x_3 + x_4) / 2, \quad y_p = (y_3 + y_4) / 2 \quad (1)$$

#### Algorithm 1

1. **Input** of Iris image and **Setting** Size of Black Rectangular area
2. **Determining** Horizontal middle point of iris image and **Searching** Black Rectangular area in horizontal direction
3. **If** Black Rectangular is found
4. **Then Detects** the coordinates of the horizontal boundary points of  $(x_1, y_1)$  and  $(x_1, y_2)$ , and vertical boundary points  $(x_3, y_3)$  and  $(x_4, y_3)$  between the pupil and the iris
5. **Determining** the centre coordinates  $(x_p, y_p)$  of the pupil using equation 1.
6. **Determining** the radius of the pupil using equation (2).
7. **Else Continue** Searching in the down side (or in upper side) of image's middle point
8. **Finding** Outer boundary of pupil and **Determining** the differences using (3) in horizontal direction
9. **Determining** Maximum values of Differences in left and right side of pupil and **Finding** boundary point between iris and sclera
10. **Determining** the centre coordinates  $(x_s, y_s)$  and radius of the iris using equation 5

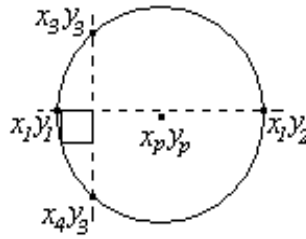


Fig. 4. Finding the centre of the pupil.

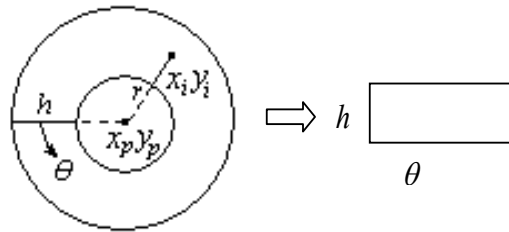


Fig. 5. Normalization of the iris.

The same procedure is applied for two different rectangular areas. In case of small differences between coordinates, the same procedure is applied for four and more different rectangular areas in order to detect a more accurate position of the pupil's centre. After determining the centre points, the radius of the pupil is computed using equation (2) (see steps 3-7 of algorithm 1).

$$r_p = \sqrt{(x_c - x_1)^2 + (y_c - y_1)^2}, \quad \text{or} \quad r_p = \sqrt{(x_c - x_3)^2 + (y_c - y_3)^2} \quad (2)$$

Because of the change of greyscale values in the outer boundaries of iris is very soft, the current edge detection methods are difficult to implement for detection of the outer boundaries. In this paper, another algorithm is applied in order to detect the outer boundaries of the iris. We start from the outer boundaries of the pupil and determine the difference of sum of greyscale values between the first ten elements and second ten elements in horizontal direction (see step 9 of algorithm 1). This process is continued in the left and right sectors of the iris. The difference corresponding to the maximum value is selected as boundary point. This procedure is implemented by the following formula.

$$DL_i = \sum_{i=10}^{y_p - (r_p + 10)} (S_{i+1} - S_i); \quad DR_j = \sum_{j=y_p + (r_p + 10)}^{right-10} (S_{j+1} - S_j) \quad (3)$$

Here  $DL$  and  $DR$  are the differences determined in the left and right sectors of the iris, correspondingly.  $x_p$  and  $y_p$  are centre coordinates of the pupil,  $r_p$  is radius of the pupil,  $right$  is the right most  $y$  coordinate of the iris image. In each point,  $S$  is calculated as

$$S_j = \sum_{k=j}^{k+10} I(i, k) \quad (4)$$

where  $i=x_p$ , for the left sector of iris  $j=10, \dots, y_p-(r_p+10)$ , and for the right sector of iris  $j=y_p+(r_p+10)$ .  $I_x(i,k)$  are greyscale values.

The centre and radius of the iris are determined using

$$y_s = (L + R) / 2, \quad r_s = (R - L) / 2 \quad (5)$$

$L=i$ , where  $i$  correspond to the value  $\max(|DL_i|)$ ,  $R=j$ , where  $j$  correspond to the value  $\max(|DR_j|)$  (see steps 10-11 of algorithm 1).

Using inner and outer circles, the normalization of iris is performed (Fig.5).

**Algorithm based on Otsu thresholding and Hough circle.** Black-rectangle searching algorithm could be efficiently used for fast localization iris images of CASIA Version 1 database. When iris images contain specular highlights then the detection of pupil region is subjected to some difficulties. Black-rectangle searching algorithm is difficult to apply such iris images. In CASIA Version 3 database, iris images have specular highlights in the pupil region. These specular highlights pose difficulties for any algorithm that assumes pupil pixels having form of disc with the darkest pixels in the central region of the image. In such cases one useful method is the utilization of Hough transformation that searches for circles in the image in order to detect pupillary boundary. Experiences on the CASIA Version 1 and Version 3 iris databases showed to us that segmentation algorithms sometimes may require intervention from researchers, for example fixing a certain threshold value. But sometimes it is necessary to use an automatic algorithm that will decide on such parameters depending on the certain image characteristics (i.e. size, illumination, average gray level, etc...). In this algorithm the combination of thresholding and Hough circle method is used for detecting pupil boundaries. The steps of the algorithm can be seen in algorithm 2. Below, explanations on the motivations and reasons for the steps involved in the algorithm is given. Generally iris databases created under different illumination settings. For each one selecting a good threshold value by programmer is possible only after some experimentation.

Furthermore different images from the same database can be captured under different illumination settings. So finding an automatic thresholding method is essential to overcome such difficulties when thresholding is necessary to separate object (iris or pupil) and background. In the algorithm proposed in this paper, Otsu thresholding (Otsu,1979) is used for this purpose. Experiments with Otsu Thresholding have demonstrated that a thresholding process for detecting iris boundaries can be automated. In that way, instead of using programmer given threshold value, the process of extracting it from the image can be automated. Several experiments showed that half of the Otsu threshold value, perfectly fits for this purpose (Step 1-2 in algorithm 2). Fig. 6(a) shows sample eye image before any processing. The process of selecting Otsu threshold value is given in Fig. 6(b). Here right vertical line demonstrates Otsu threshold value, and the left vertical line demonstrates half of Otsu threshold value. In some cases of thresholding, Otsu value by itself separates sclera from the rest of the eye in the image. So taking half of the Otsu value is a good approximation (also good heuristics) for separating pupil from the sclera region. Here multi-level Otsu methods can be utilized for better thresholding, which can be computationally more expensive.

Interested readers can find good reviews and surveys on thresholding and segmentation methods in (Sezgin & Sankur, 2004; Trier & Taxt,1995]. Separating pupil region from the image can let us to find certain characteristic of the pupil, at least approximate radius size when the "pupil pixels" are counted. Of course any thresholding that can separate pupil will



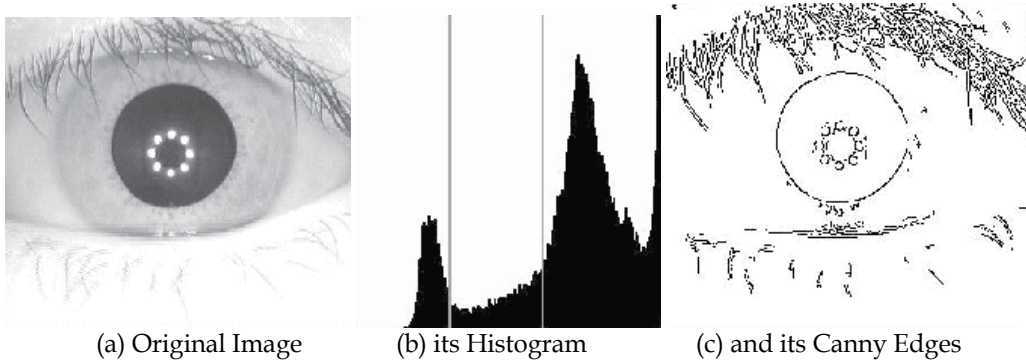


Fig. 6. (a) Original image (b) Half-Otsu (leftmost line) and Otsu values marked on the histogram of the Original Image.

also separate dark eyebrows too (Fig. 7(a)). But again here all is needed is an approximate radius plus or minus "delta". In other words small range for Hough circle search. That is  $r_{expected} \pm \Delta r$ . This information is valuable in searching circles in the Hough space having radius in the range of  $[r_{expected} \pm \Delta r]$ . Better and more efficient results from the Hough circle search can be obtained by reducing the radius range. It can be seen from the Fig. 7(a) and 7(b) that thresholding can give us good approximation for the pupil region.

#### Algorithm 2

1. **Finding** Otsu threshold value of the Original Image
2. **Thresholding** Original Image with the half of the Otsu value
3. **Dilating and Eroding (Opening)** Half-Otsu thresholded image Generate two instances of the resulting image: Inst1, Inst2
4. **Performing Canny-Edge Detection** on the thresholded image Inst1
5. **Counting** black pixels on the thresholded image Inst2 **Estimating** expected pupil radius ( $r_{expected}$ ) on the thresholded image Inst2
6. **Searching for Circles in Hough Space** on the thresholded image Inst1 Searching circles with radius =  $[r_{expected} \pm \Delta r]$
7. For Pupillary Boundary **Finding the best match** : The Hough circle that covers pupil edges best on the Inst1
8. **Median Filtering** Original Image **Performing Gradient Search** for limbic boundary on the Median Filtered Image
9. **Extracting** Normalized Image from the Original image
10. **Background Subtraction and Bicubic scaling** for the Normalized Image
11. **Return** Scaled and Enhanced Iris Band

The formula for the area of a disc is used for finding the approximate pupil radius. But before that one has to "clean" isolated black pixels and avoid counting them along with the black pixels that may remain on the borders. For the first problem, morphological operators with 3x3 rectangular structuring element, like in-place erode (where central pixel replaced by the minimum in the 3x3 neighbourhood) and in-place dilate (where central pixel replaced by the maximum in the 3x3 neighbourhood) (Step 3 in algorithm 2) on the thresholded binary image (white 0 and black 1) is used. Erosion is to clean isolated black pixels and dilation to "shrink" the regions that may remain in the pupil area due to specular highlight. Especially CASIA V3 database images acquired with camera having leds. The

specular highlight in the CASIA V3 images makes simple thresholding methods unusable for finding pupillary boundary. For the second problem it helps to focus on the smaller central region on the image, assuming that the pupil is in the central region of the image, which is the case for CASIA databases used in the experiments (Step 5 in algorithm 2). In the algorithm Hough circle searching function is used in finding limbic boundaries, which given radius range and several other parameters, returns the list of possible circles on the input image (Step 6 in algorithm 2). Thresholded (with half of the Otsu value) and morphologically processed image is given as an input to this function. In this case the function will be more focused on finding pupil boundary, which can be assumed as a circle (Fig. 7(c)). After Canny edge detection, function searches for the circles on the edge pixels. But one has to devise scoring method for finding the best matched circle for pupillary boundary. For this the circles returned by Hough circle search function are scored, based on the number of overlapped pixels with the Canny edge pixels. So the algorithm used separate copy of the Canny edge image from the same image that is given as an input to the Hough circle function (Step 4 in algorithm 2). For overlapping "relaxed" counting is used (Step 7 in our algorithm, see algorithm 2). Every pixels on the edge image that falls into the 3x3 neighbourhood of each circle pixels is counted (marked them to avoid double counting), instead of looking for "strict" overlapping. This has to be done since the pupillary boundary is not a perfect circle most of the time.

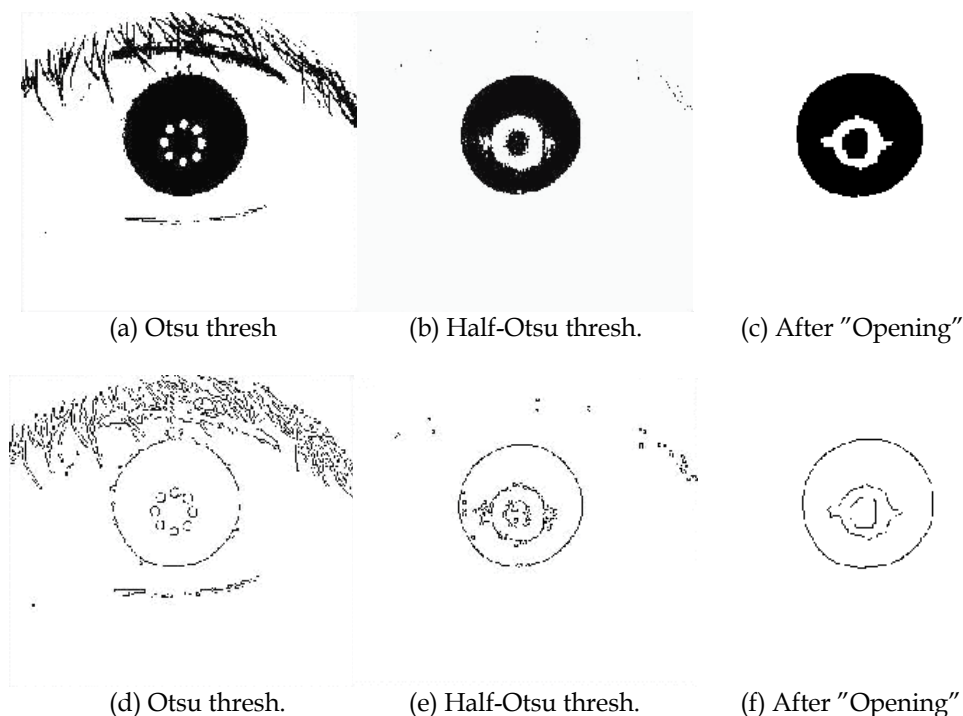


Fig. 7. 1st row: Otsu Thresholded image (a) and effect of "Opening" (c) (eroding and dilating black pixels) on Half-Otsu Thresholded image (b). 2nd row: Canny Edges of the; Otsu thresholded (d), Half-Otsu Thresholded (e), and after "Opening" (f).

After finding pupillary boundary, using above given formulas (3) simple gradient search is utilized in both (a) Original Image (b) its Histogram (c) and its Canny Edges left and right directions, where differences of average grey level of the two consecutive windows (size is adjusted according to the image size and average grey level) observed on the median filtered image. Here the algorithm needs to find the point where the grey level starts increasing, in other words where the sclera region (generally whiter than the iris region) starts (Step 8 in our algorithm, see algorithm 2). Doing such a search on the original image can not give good results, as the histogram of the central (passing from the centre of the pupil) line (band in fact, 3-pixel wide) of the original image may contain little "spikes" (see Fig. 8). These "spikes" may be problem in finding the "real" place where the grey level starts increasing (0 is black and 255 is white). To overcome this difficulty median filtering is used, since median filtering preserves edges and smoothes boundaries without being computationally very expensive. The sample result of this search can be seen in Fig. 8, where the centre of the best matched circle is marked with little star and it is drawn in grey colour over the black pixels of the edges.

In this type of search method, crucial parameters are the width of the windows and the threshold value that will determine the actual place where the grey level increases. The window size is adjusted according to the image width and also the threshold is adjusted according to the average grey level of the image. The idea here is that, in bright images sclera can be found with threshold value greater than the ones where you have darker image. So smaller threshold values are used for darker images. The algorithm infact does not separate images into dark or light. To do better parameter calculations, small fuzzy-logic front-end is added to the algorithm which it was not included on the flow-chart (see algorithm 2). Basically we utilized the information acquired in the experiments. In the experiments it has been seen that setting global (for all images in the database) values for certain parameters, like thresholds, does not give good results. So in the algorithm the idea of adjusting them according to certain image characteristics is practised, as it is said above. The algorithm groups images according to their grey level averages. For example images having average grey level less than 100 can be classified as very dark images. Based on this the algorithm formed several groups based on the several ranges of the average grey level: [0..99], [100..140], [141..170], [171..200], [201..255]. Five groups are formed, which can be named as very dark, dark, middle, bright, very bright. For each group the algorithm adjusted certain parameters accordingly, also considering image size and expected radius of the pupil (which is calculated in step 5 of the algorithm, see algorithm 2). Finally in steps 9 and 10 (see algorithm 2) using "homogeneous rubber sheet" (Daugman,2004) segmented (see Fig. 8 for the iris band marked on the image) iris images are normalized (Fig. 8). The formulas used for normalization are described in next section. After background subtraction is made followed by bicubic scaling and histogram equalization. The results of these operations (steps 9-11) are described in section 2.3. Every iris band that are segmented scaled to the uniform size of 360x64.

### 2.3 Iris normalization

The irises captured from the different people have different sizes. The size of the irises from the same eye may change due to illumination variations, distance from the camera, or other factors. At the same time, the iris and the pupil are non concentric. These factors may affect the result of iris matching. In order to avoid these factors and achieve more accurate

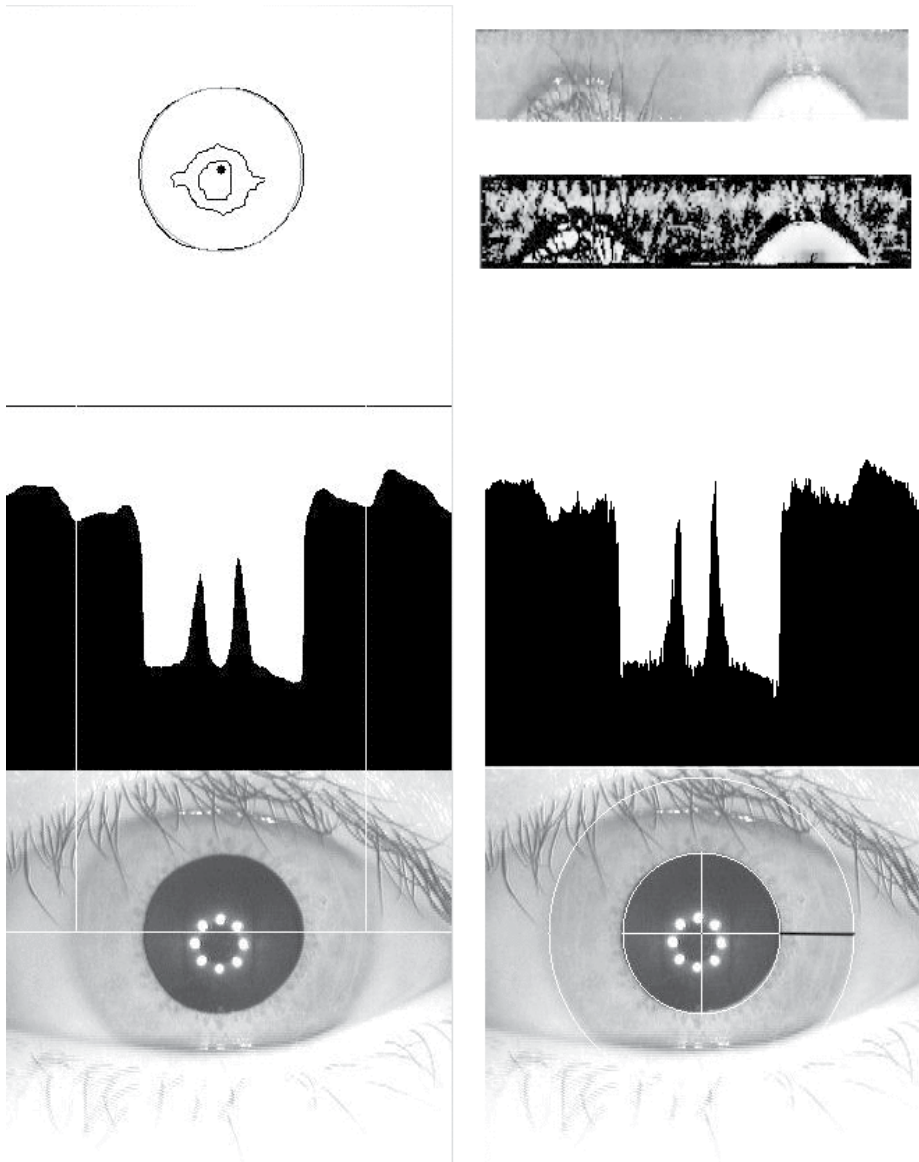


Fig. 8. Best-match Hough circle (left top) and detected limbic boundaries shown on the image (left bottom) and on the histogram of the central band (from the Median filtered image, left middle). Spikes on the histogram of the central (passing from the pupil centre) band from Original image (right middle). Detected iris band marked on the original image (right bottom). Normalized iris band (right uppermost). Scaled-Enhanced Normalized iris band (right second from the top).

recognition, the normalization of iris images is implemented. In normalization, the iris circular region is transformed to a rectangular region with a fixed size. With the boundaries detected, the iris region is normalized from Cartesian coordinates to polar representation. This operation is done using the following operation (Fig.9).

$$\theta \in [0, 2\pi], \quad r \in [R_p, R_L(\theta)] \quad (6)$$

$$x_i = x_p + r \cdot \cos(\theta); \quad y_i = y_p + r \cdot \sin(\theta)$$

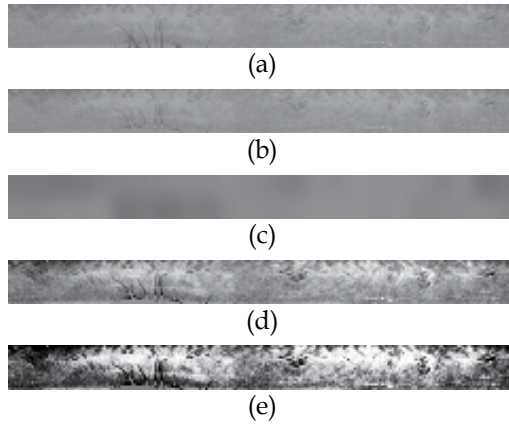


Fig. 9. a) Normalized image, b) Normalized image after removing eyelashes, c) Image of nonuniform background illumination, d) Image after subtracting background illumination, e) Enhanced image after histogram equalization.

Here  $(x_i, y_i)$  is the point located between the coordinates of the papillary and limbic boundaries in the direction  $\theta$ .  $(x_p, y_p)$  is the centre coordinate of the pupil,  $R_p$  is the radius of the pupil, and  $R_L(\theta)$  is the distance between centre of the pupil and the point of limbic boundary.

In the localization step, the eyelid detection is performed. The effect of eyelids is erased from the iris image using the linear Hough transform. After normalization (Fig. 9(a)), the effect of eyelashes is removed from the iris image (Fig. 9(b)). Analysis reveals that eyelashes are quite dark when compared with the rest of the eye image. For isolating eyelashes, a thresholding technique was used. To improve the contrast and brightness of image and obtain a well distributed texture image, an enhancement is applied. Received normalized image using averaging is resized. The mean of each 16x16 small block constitutes a coarse estimate of the background illumination. During enhancement, background illumination (Fig. 9(c)) is subtracted from the normalized image to compensate for a variety of lighting conditions. Then the lighting corrected image (Fig. 9(d)) is enhanced by histogram equalization. Fig. 9(e) demonstrates the preprocessing results of iris image. The texture characteristics of iris image are shown more clearly. Such preprocessing compensates for the nonuniform illumination and improves the contrast of the image.

Normalized iris provides important texture information. This spatial pattern of the iris is characterized by the frequency and orientation information that contains freckles, coronas, strips, furrows, crypts, and so on.

### 3. Neural network based iris pattern classification

#### 3.1 Neural Network based model

In this paper, a Neural Network (NN) is used to recognise the iris patterns. In this approach, the normalized and enhanced iris image is represented by a two-dimensional array. This array contains the greyscale values of the texture of the iris pattern. These values are input

signals for the neural network. Architecture of NN is given in Fig. 10. Two hidden layers are used in the NN. In this structure,  $x_1, x_2, \dots, x_m$  are greyscale values of input array that characterizes the iris texture information,  $P_1, P_2, \dots, P_n$  are output patterns that characterize the irises.

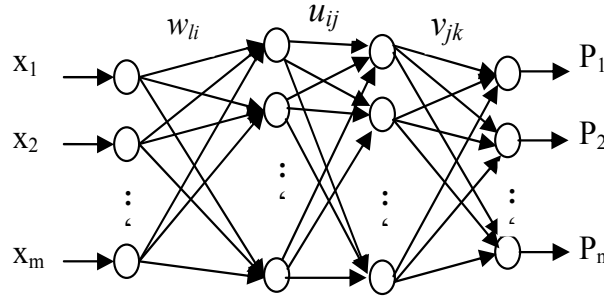


Fig. 10. Neural Network Architecture.

The  $k$ -th output of neural network is determined by the formula

$$P_k = f_k \left( \sum_{j=1}^{h2} v_{jk} \cdot f_j \left( \sum_{i=1}^{h1} u_{ij} \cdot f_i \left( \sum_{l=1}^m w_{li} x_l \right) \right) \right) \quad (7)$$

where  $v_{jk}$  are weights between the output and second hidden layers of network,  $u_{ij}$  are weights between the hidden layers,  $w_{li}$  are weights between the input and first hidden layers,  $f$  is the activation function that is used in neurons,  $x_l$  is input signal. Here  $k=1, \dots, n$ ,  $j=1, \dots, h2$ ,  $i=1, \dots, h1$ ,  $l=1, \dots, m$ ,  $m$  is number of neurons in input layer,  $n$  is number of neurons in output layer,  $h1$  and  $h2$  are number of neurons in first and second hidden layers, correspondingly.

In formula (7)  $P_k$  output signals of NN are determined as

$$P_k = \frac{1}{1 + e^{-\sum_{j=1}^{h2} v_{jk} y_j}} \quad (8)$$

where,

$$y_j = \frac{1}{1 + e^{-\sum_{i=1}^{h1} u_{ij} y_i}}; \quad y_i = \frac{1}{1 + e^{-\sum_{l=1}^m w_{li} x_l}} \quad (9)$$

Here  $y_i$  and  $y_j$  are output signals of first and second hidden layers, respectively.

After activation of neural network, the training of the parameters of NN start. The trained network is then used for the iris recognition in online regime.

### 3.2 Parameter learning

In this paper, a gradient based learning algorithm with adaptive learning rate is adopted. This allows to guarantee convergence and speed up learning processes. In addition a momentum is used to speed up learning processes.

At the beginning, the parameters of NN are generated randomly. The parameters  $v_{jk}$ ,  $u_{ij}$ , and  $w_{li}$  of NN are weight coefficients of second, third and last layers, respectively. Here  $k=1, \dots, n$ ,  $j=1, \dots, h_2$ ,  $i=1, \dots, h_1$ ,  $l=1, \dots, m$ . To generate NN recognition model, the training of the weight coefficients of  $v_{jk}$ ,  $u_{ij}$ , and  $w_{li}$  has been carried out.

During training the value of the following cost function is calculated.

$$E = \frac{1}{2} \sum_{k=1}^n (P_k^d - P_k)^2 \quad (10)$$

Here  $n$  is the number of output signals of the network and  $P_k^d$  and  $P_k$  are the desired and the current output values of the network, respectively. The parameters  $v_{jk}$ ,  $u_{ij}$ , and  $w_{li}$  of neural network are adjusted by using the following formulas.

$$\begin{aligned} w_{li}(t+1) &= w_{li}(t) - \gamma \frac{\partial E}{\partial w_{li}} + \lambda(w_{li}(t) - w_{li}(t-1)) \\ u_{ij}(t+1) &= u_{ij}(t) - \gamma \frac{\partial E}{\partial u_{ij}} + \lambda(u_{ij}(t) - u_{ij}(t-1)) \\ v_{jk}(t+1) &= v_{jk}(t) - \gamma \frac{\partial E}{\partial v_{jk}} + \lambda(v_{jk}(t) - v_{jk}(t-1)) \end{aligned} \quad (11)$$

Here  $\gamma$  is the learning rate, and  $\lambda$  is the momentum.

The derivatives in (11) are determined using the following formulas.

$$\begin{aligned} \frac{\partial E}{\partial v_{jk}} &= \frac{\partial E}{\partial P_k} \frac{\partial P_k}{\partial v_{jk}}; & \frac{\partial E}{\partial u_{ij}} &= \frac{\partial E}{\partial P_k} \frac{\partial P_k}{\partial y_j} \frac{\partial y_j}{\partial u_{ij}}; \\ \frac{\partial E}{\partial w_{li}} &= \frac{\partial E}{\partial P_k} \frac{\partial P_k}{\partial y_j} \frac{\partial y_j}{\partial y_i} \frac{\partial y_i}{\partial w_{li}}; \end{aligned} \quad (12)$$

The derivatives in (12) are determined as

$$\begin{aligned} \frac{\partial E}{\partial P_k} &= (P_k - P_k^d); & \frac{\partial P_k}{\partial v_{jk}} &= P_k(1 - P_k) \cdot y_j \\ \frac{\partial P_k}{\partial y_j} &= P_k(1 - P_k) \cdot v_{jk}; & \frac{\partial y_j}{\partial u_{ij}} &= y_j(1 - y_j) \cdot y_i \\ \frac{\partial y_j}{\partial y_i} &= y_j(1 - y_j) \cdot u_{ij}; & \frac{\partial y_i}{\partial w_{li}} &= y_i(1 - y_i) \cdot x_l \end{aligned} \quad (13)$$

Using equations (11-13), an update of the parameters of the neural network is carried out.

One important problem in learning algorithms is convergence. The convergence of the gradient descent method depends on the selection of the initial values of the learning rate. Usually, this value is selected in the interval [0-1]. A large value of the learning rate may lead to unstable learning, a small value of the learning rate results in a slow learning speed. In this paper, an adaptive approach is used for updating these parameters. That is, the learning of the NN parameters is started with a small value of the learning rate  $\gamma(t)$ . During

learning,  $\gamma(t)$  is increased if the value of change of error  $\Delta E(t)=E(t)-E(t+1)$  is positive, and decreased if negative. This strategy ensures a stable learning for the type-2 NN, guarantees the convergence and speeds up the learning.

#### 4. Experimental results

The CASIA iris image databases are used to evaluate the iris recognition algorithms. Currently this is one of largest iris database available in the public domain. The experiments are performed for two CASIA version 1 and CASIA version 3 iris databases. At first CASIA version 1 iris database is used. This image database contains 756 eye images from 108 different persons. Experiments are performed in two stages: iris segmentation and iris recognition. At first stage the above described black rectangle algorithm is applied for the localization of irises. The experiments were performed by using Matlab on Pentium IV PC. The average time for the detection of inner and outer circles of the iris images was 0.14s. The accuracy rate was 98.62%. Also using the same conditions, the computer modelling of the iris localization is carried out by means of Hough transform and Canny edge detection realized by Masek and integrodifferential operator realized by Daugman. The average time for iris localization using Hough transform is obtained 85 sec, and 90 sec using integrodifferential operator. Table 1 demonstrates the comparative results of different techniques used for iris localization. The results of Daugman method are difficult for comparison. If we use the algorithm which is given in (Daugman,2001; Daugman,2004) then the segmentation represents 57.7% of precision. If we take into account the improvements that were done by author then Daugman method presents 100% of precision. The experimental results have shown that the proposed iris localization rectangular area algorithm has better performance.

Second experiment was performed using algorithm based on Otsu thresholding and Hough circle. Experiments are performed using both CASIA1 version 1 and version 3 iris image databases. As mentioned CASIA version 1 image database contains 756 eye images from 108 different people, CASIA version 3 (Interval) database contains 2655 eye images from 249 different people. Experiments were performed on Intel® Quad Core 2.4GHz machine with openSUSE 11.0 Linux system and Intel C/C++ (V11.0.074) compilers with OpenCV (V1.0) and IPP (V6.0) libraries were used (all in 64 bit). IPP optimized mode is utilized for OpenCV, since IPP library can provide speedups up to 3.84x for multi-threading on a 4-core processor. The implemented algorithm can extract iris band in about 0.19 sec.

This roughly equals to 5 fps processing speed (for 320x280 resolution), with on line camera. The segmentation results for both databases are given in the Table 2. We believe the proposed algorithm can be used for on line iris extraction with an acceptable real-time performance. After segmentation we have checked extracted iris images visually to see if there is any extracted image that can not be suitable input for any classification method. Total 7 images rejected after inspection, from the CASIA V3. In the case of CASIA V1 images we did not see any unsuccessful segmentation.

In second stage the iris pattern classification using NN is performed. For classification 50 different persons were selected from iris database. From each person two irises are used for training and two irises for testing. The detected irises after normalization and enhancement are scaled by using averaging. This help to reduce the size of neural network. Then the images are represented by matrices. These matrices are the input signal for the neural



| Methodology | Accuracy rate | Average time |
|-------------|---------------|--------------|
| Daugman [4] | 57.7%         | 90 s         |
| Wildes [7]  | 86.49%        | 110 s        |
| Masek [9]   | 83.92%        | 85 s         |
| Proposed    | 98.62%        | 0.14 s       |

Table 1. Accuracy rate for iris segmentation.

| Version  | Images | Segmented | Percent | Avg (sec) |
|----------|--------|-----------|---------|-----------|
| CASIA V1 | 756    | 748       | 98.94   | 0.18      |
| CASIA V3 | 2655   | 2624      | 98.83   | 0.19      |

Table 2. Segmentation Results for CASIA V1 (all) and V3 (CASIA-IrisV3-Interval).

network. The outputs of the neural network are classes of iris patterns. Each class characterizes the certain person's iris. Two hidden layers are used in neural network. The numbers of neurons in first and second hidden layers are 120 and 81, correspondingly. In the hidden layers the number of neurons is selected after several experiments by trial. The use of less number of neurons in hidden layer does not provide enough accuracy of network. To improve computational power of neural network and to decrease training error we increment number of neurons in hidden layer. Neural learning algorithm is applied in order to solve iris classification. From each set of iris images, two patterns are used for training and two patterns for testing. After training the remaining images are used for testing. The recognition rate of NN system was 99.25%. The obtained recognition result is compared with the recognition results of other methods that utilize the same iris database. The results of this comparison are given in Table 3. As shown in the table, the identification result obtained using the neural network approach illustrates the success of its efficient use in iris recognition. But we need to note that future increasing the number of iris patterns decrease the recognition rate using NN.

| Methodology    | Accuracy rate |
|----------------|---------------|
| Daugman [4]    | 100%          |
| Boles [8]      | 92.64%        |
| Li Ma [11]     | 94.9%         |
| Avila [26]     | 97.89%        |
| Neural Network | 99.25%        |

Table 3. The recognition performance of comparing with existing methods.

## 5. Conclusion

Personal identification using iris recognition is presented. Iris segmentation algorithms are proposed for accurate and fast detection of irise patterns. First one is based on detection of black rectangular patterns on purples. Second algorithm is based Otsu thresholding and Hough circle of iris patterns that can be efficiently applied to images containing specular

highlights. The application of segmentation algorithms are demonstrated on CASIA version 1 and version 3 databases. Using black rectangle algorithm the average time for iris segmentation is obtained to be 0.14 sec on Pentium IV PC using Matlab. Accuracy rate of iris segmentation 98.62% is achieved. For second algorithm the accuracy of iris segmentation for CASIA version 1 is 98.84, for version 3 – 98.83%. The located iris after pre-processing is represented by a data set. Using this data set as input signal the neural network is used to recognize the iris patterns. The recognition accuracy was 99.25%. The obtained results demonstrate the efficiency of proposed algorithms.

## 6. References

- Jain A., Bolle R., and Kanti S. P. *Biometrics: Personal Identification in a Networked Society*. Kluwer, 1998.
- Adler A. *Physiology of Eye : Clinical Application*. London, The C.V. Mosby Company, fourth edition, 1965.
- Daugman J. *Biometric Personal Identification System Based on Iris Analysis*. US Patent no. 5291560, 1994.
- Daugman J. *Statistical richness of visual phase information: Update on recognizing persons by iris patterns*. *Int. Journal of Computer Vision*, 2001.
- Daugman J. *Demodulation by complex-valued wavelets for stochastic pattern recognition*. *Int. Journal of Wavelets, Multiresolution and Information Processing*, 2003.
- Daugman J. *How iris recognition works*. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21 – 30, January 2004.
- Wildes R. *Iris recognition: An emerging biometric technology*. *Proc. of the IEEE*, 85(9):1348 – 1363, September 1997.
- Boles W. and Boashash B. *A human identification technique using images of the iris and wavelet transform*. *IEEE Trans. on Signal Processing*, 46(4):1185 – 1188, 1998.
- Masek L. *Recognition of Human Iris Patterns for Biometric Identification*. BEng. Thesis. School of Computer Science and Software Engineering, The University of Western Australia, 2003.
- Ma L., Tan T., Wang Y., and Zhang D. *Personal identification based on iris texture analysis*. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 25(12):1519 – 1533, December 2003.
- Ma L., Wang Y. H., and Tan T. N. *Iris recognition based on multichannel gabor filtering*. *Proc. of the Fifth Asian Conference on Computer Vision, Australia*, pages 279 – 283, 2002.
- Tisse C., Martin L., Torres L. and Robert M. *Person identification technique using human iris recognition*. *Proc. of Vision Interface*, pages 294 – 299, 2002.
- Kanag H. and Xu G. *Iris recognition system*. *Journal of Circuit and Systems*, 15(1):11 – 15, 2000.
- Yuan W., Lin Z. and Xu L. *A rapid iris location method based on the structure of human eyes*. *Proc. Of 27th IEEE Annual Conference Engineering in Medicine and Biology, Shanghai, China, September 1-4 2005*.
- Daugman J. *New methods in iris recognition*. *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, 37(5):1168 – 1176, October 2007.
- Vatsa M., Singh R., and Noore A. *Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing*. *IEEE Trans.*

- on Systems, Man, and Cybernetics Part B: Cybernetics, 38(4):1021 – 1035, August 2008.
- Zuo J. and Schmid N. An Automatic Algorithm for Evaluating the Precision of Iris Segmentation, IEEE Second Int. Conf. on Biometrics Theory, Applications and Systems (BTAS 08) Sep. 29 - Oct. 1, 2008.
- Liu X., Bowyer K., and Flynn P. Experiments with an improved iris segmentation algorithm. Fourth IEEE Workshop on Automatic Identification Advanced Technologies, 17 - 18:118 – 123, 2005.
- Cui J., Wang Y., Tan T., Ma L., and Sun Z. A fast and robust iris localization method based on texture segmentation. Proc. SPIE, 5404:401 – 408, 2004.
- Abiyev R. and Altunkaya K. Neural Network Based Biometric Personal Identification with fast iris segmentation. Int. Journal of Control, Automation and Systems. Vol.7, No.1, 2009.
- Abiyev R. and Altunkaya K. Iris recognition for biometric personal identification using neural networks. Lecture Notes in Computer Sciences, Springer-Verlag, 4669, 2007.
- Abiyev R. and Altunkaya K. Personal Iris Recognition Using Neural Networks. International Journal of Security and its Applications, vol.2. No.2, April, 2008.
- Abiyev R. and Kilic K. Adaptive Iris segmentation. Lecture Notes in Computer Sciences, Springer-Verlag, Berlin Heidelberg, CS press, 2009.
- Daugman J. and Downing C. Recognizing iris texture by phase demodulation. IEEE Colloquium on Image Processing for Biometric Measurement, 2:1 – 8, 1994.
- Miyazawa K., Ito K., Aoki T., Kobayashi K. and Nakajima H. An effective approach for iris recognition using phase-based image matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, 30(10):1741 – 1756, October 2008.
- Sanchez-Avila C. and Sanchez-Reillo R. Iris-based biometric recognition using dyadic wavelet transform. IEEE Aerospace and Electronic Systems Magazine, pages 3 – 6, 2002.
- Noh S., Bae K. and Kim J. A novel method to extract features for iris recognition system. Proc. 4th Int. Conf. Audio and Video Based Biometric Person Authentication, pages 838 – 844, 2003.
- Mallat S. Zero crossings of a wavelet transform. IEEE Trans. Inf. Theory, 37(4):1019 – 1033, April 1992.
- Park C., Lee J., Smith M. and Park K. Iris based personal authentication using a normalized directional energy feature. Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication, pages 224 – 232, 2003.
- Lim S., Lee K., Byeon O. and Kim T. Efficient iris recognition through improvement of feature vector and classifier. ETRI J., 23(2):61 – 70, 2001.
- Ma L., Tan T., Zhang D., and Wang Y. "Local intensity variation analysis for iris recognition", *Pattern Recognition*, vol. 37, no. 6, pp. 1287-1298, 2005.
- Scotti F., "Computational intelligence techniques for reflections identification in iris biometric images", *CIMSA 2007 - IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications*, Ostuni - Italy, 27-29 June 2007
- Wang Y. and Han J. Q. Iris feature extraction using independent component analysis. Proc. 4th Int. Conf. Audio and Video Based Biometric Person Authentication, pages 838 – 844, 2003.

- Wang Y. and Han J. Q.. Iris recognition using independent component analysis. Proc. of the Fourth Int. Conf. on Machine Learning and Cybernetics, Guangzhou, 2005.
- Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans. Sys., Man., Cyber.*, 9:62 - 66, 1979.
- Sezgin M. and Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146 - 165, 2004.
- Trier I. D. and Taxt T. Evaluation of binarization methods for document images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1995.
- CASIA iris database. Institute of Automation, Chinese Academy of Sciences <http://www.cbsr.ia.ac.cn/IrisDatabase.htm>
- Abiyev R.H. and Kilic K. An efficient Fractal Measure for Image Texture Recognition. International Conference on Soft Computing and Computing with Words in System Analysis, Decision and Control ICSCCW-2009, North Cyprus, Turkey, 2009

# Iris Recognition System Using Support Vector Machines

Hasimah Ali and Momoh J. E. Salami

*University Malaysia Perlis (UniMAP), International Islamic University Malaysia (IIUM)  
Malaysia*

## 1. Introduction

In the modern world, a reliable personal identification infrastructure is required to control the access in order to secure areas or materials. Conventional methods of recognizing the identity of a person by using passwords or cards are not altogether reliable, because they can be forgotten, stolen, disclosable, or transferable (Zhang, 2000). Biometric technology, which is based on physical and behavioral features of human body such as face, fingerprint, hand shapes, iris, palmprint, keystroke, signature and voice, (Lim et al., 2001, Zhang, 2000, Zhu et al., 1999) has now been considered as an alternative to existing systems in a great deal of application domains such as bank Automatic Teller Machines (ATM), telecommunication, internet security and airport security.

Each biometric technology has its own advantages and disadvantages based on their usability and security. Among the various traits, iris recognition has attracted a lot of attention. Iris is an internal (yet externally visible) organ of the eye, which is well protected from the environment and its patterns are apparently stable throughout the life. The iris consists of variable sized hole called pupil. The average diameter of the iris is 12 mm, and the pupil size can vary from 10% to 80% of the iris diameter. It has the great mathematical advantage that its pattern variability amongst people is enormous (Daugman, 2002).

The number of features in human iris is large. Its complex pattern can contain many distinctive features such as arching, ligaments, furrows, ridges, crypts, rings, corona, freckles and zigzag collarette (Wildes, 1999, Daugman, 2002) for personal identification. Fig. 1 is an example of human iris. That is because every iris has fine unique texture and does not change over time. In addition, iris pattern can have up to 249 independent degrees of freedom. Because of high randomness in the iris pattern, it has made the technique more robust and it is very difficult to deceive an iris pattern (Daugman, 2003). Unlike other biometric traits, iris recognition is the most accurate and non-invasive biometric for secure authentication and positive identification. This proposed system use a publicly available library for iris recognition written in MATLAB (Masek, 2003).

Due to the advantages of iris recognition systems which offer reliable and effective security in the present day, this research proposed the use of iris-based as verification system system to identify the person's identity. This research work adopts Support Vector Machines (SVMs) as pattern classification techniques which are based on iris code model which the feature vector size is transformed to one-dimension vector which reduces to  $1 \times 480$  by using averaging techniques (each segment is divided by 20) contains the average value to

recognize an authorized user and unauthorized user. The effectiveness of the proposed system is evaluated based on False Rejection Rate (FRR) and False Acceptance Rate (FAR).

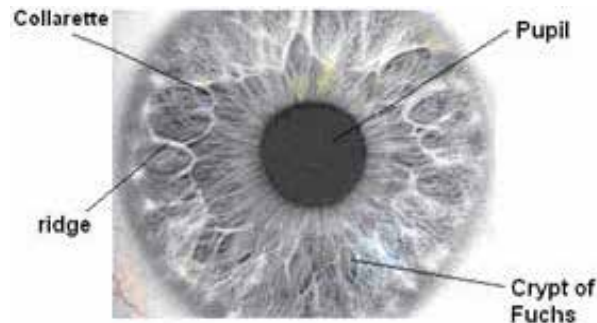


Fig. 1. An iris which has highly complex unique texture.

## 2. Iris image acquisition

Image acquisition is meant to produce image of user's eye region. A high-quality image of the iris has to be captured in order for the iris recognition system to work efficiently. The acquired image of the iris must have sufficient resolution and sharpness to support recognition. Besides, it is important to have good contrast in the interior iris pattern. The brightness of illumination which is not uniformly distributed will result in poor quality images with lots of spatial reflections. Images captured using infrared camera has good quality with high contrast and low reflections (Wildes, 1999).

## 3. Iris image preprocessing

The acquired image that contains irrelevant parts (e. eyelid, eyelash, pupil, etc) should be removed. For the purpose of analysis, the original image needs to be preprocessed. The preprocessing is composed of two steps: iris localization/ segmentation and normalization.

### 3.1 Iris localization / segmentation

The first stage of iris recognition system is to isolate the actual iris region in a digital eye. The purpose of iris localization is to localize an acquired image that corresponds to an iris. The iris region, shown in Fig. 2, can be approximated by two circles, one for the iris/sclera boundary and the other is interior to the first, for the iris/pupil boundary (Daugman, 2002). The eyelids and eyelashes are normally occluding the upper and lower parts of the iris region. Eyelids and eyelashes are isolated from the detected iris image by considering them as noise because they degrade the performance of the system.

Daugman (2002) proposed integro-differential operator to detect the centre and diameter of the iris and used the differential operators to detect the pupil. That is,

$$\max_{(r, x_p, y_o)} \left| G_o(r) * \frac{\partial}{\partial r} \oint_{r, x_o, y_o} \frac{I(x, y)}{2\pi r} ds \right| \quad (1)$$

where  $I(x, y)$  is the eye image,  $r$  is the radius to search for,  $G_o(r)$  is a Gaussian smoothing function,  $s$  is the contour of the circle given by  $r, x_o, y_o$  (Masek, 2003). Wildes (1999) used

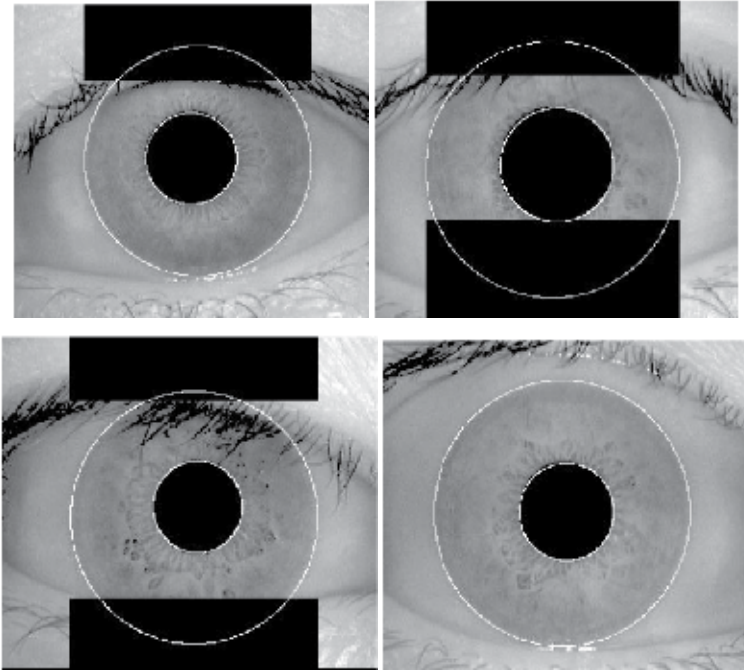


Fig. 2. Automatic segmentation of various images from the CASIA database. Black regions denote detected eyelid and eyelash regions (Masek, 2003).

the first derivative of image intensity to find the location of edges corresponding to the borders of the iris. This approach explicitly models the upper and lower eyelids with parabolic arcs whereas Daugman (2002) excludes the upper and lower portion of the image in its modal.

### 3.2 Iris normalization

The localized iris part is transformed into polar coordinates system so that it has fixed dimensions and also to overcome imaging inconsistencies. The annular iris region is transformed into rectangular region where the iris texture is analyzed. The Cartesian to polar transform of the iris region is based on the Daugman's Rubber Sheet model as shown in Fig. 3. The rubber sheet model takes into account pupil dilation and size dimensions. Therefore, the iris region is modeled as a flexible rubber sheet anchored at the boundary with the pupils centre as the reference point (Masek, 2003). Daugman remap 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  $[0, 2\pi]$ . The remapping of the iris region is modeled as,

$$I(x(r,\theta), y(r,\theta)) \rightarrow I(r,\theta) \quad (2)$$

with

$$x(r,\theta) = (1-r)x_p(\theta) + rx_1(\theta) \quad (3)$$

$$y(r,\theta) = (1-r)y_p(\theta) + ry_1(\theta) \quad (4)$$

where  $I(x, y)$  is the iris region image,  $(x, y)$  are the original Cartesian coordinates,  $(r, \theta)$  are corresponding normalized polar coordinates, and  $x_p, y_p$  and  $x_i, y_i$  are the coordinates of the pupil and iris boundaries along the  $\theta$  direction.

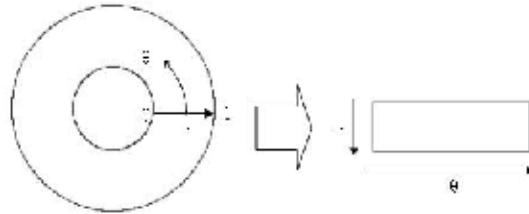


Fig. 3. Daugman's Rubber Sheet Model.

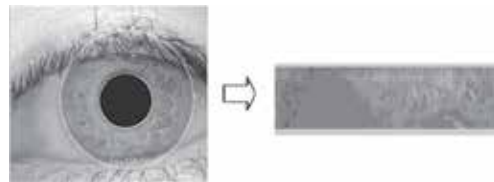


Fig. 4. Iris normalized into polar coordinates.

For normalization of iris regions, a technique based on Daugman's rubber sheet model was employed in the proposed system. The centre of pupil was considered as the reference point and radial vectors pass through the iris region (Masek, 2003).

#### 4. Feature extraction / Encoding

The iris has a particularly interesting structure and provides abundant texture information. In order to recognize the individuals accurately, the most discriminating features that present in the region must be extracted. Only the significant features of the iris must be encoded. There are many algorithm that are available for feature extraction such as wavelet encoding, Gabor filters (Daugman, 2002) , Log-Gabor filter (Field, 1987), zero-crossings of the 1D wavelet (Boles et. al, 1998), Haar wavelet (Lim et. al, 2001) and Laplacian of Gaussian filters (Widles, 1996).

In the proposed system, Masek's algorithm was used for feature encoding by convolving the normalized iris pattern with 1D Log-Gabor wavelets. Log-Gabor filters are constructed using,

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right) \quad (5)$$

where  $f_0$  represents the centre frequency ,  $\sigma$  gives bandwidth of the filter. The 2D normalized pattern is broken up into a number of 1D signal, and then these signal are convolved with 1D Gabor wavelet. The rows of the 2D normalized pattern are taken as the 1D signal. Each row corresponds to a circular ring on the iris region. The phase information in the pattern only is used because the phase angles are assigned regardless of the image contrast. Extraction of the phase information, according to Daugman, is done using 2D Gabor wavelets. It determines which quadrant the resulting phasor lays using the wavelet,



$$h\{\text{Re}, \text{Im}\} = \text{sgn}_{\{\text{Re}, \text{Im}\}} \iint_{\rho \phi} I(\rho, \phi) e^{-i\omega(\theta_0 - \phi)} e^{-(r_0 - \rho)^2 / \alpha^2} e^{-(\theta_0 - \phi)^2 / \beta^2} \rho d\rho d\phi \quad (6)$$

where,  $h\{\text{Re}, \text{Im}\}$  has the real boundary and imaginary part, each having the value 1 or 0, depending on which quadrant it lies in (Daugman, 2002).  $I(\rho, \phi)$  is the raw iris image in a dimensionless polar coordinate system that is size and translation-invariant, and which also corrects for pupil dilation;  $\alpha$  and  $\beta$  are the multi-scale 2D wavelet size parameters, spanning an 8-fold range from 0.15 mm to 1.2 mm on the iris;  $\omega$  is wavelet frequency, spanning 3 octaves in inverse proportion to  $\beta$ ; and  $(r_0, \theta_0)$  represent the polar coordinates of each region of iris for which the phasor coordinates  $h\{R_e, I_m\}$  are computed.

## 5. Matching

To verify a person's identity, the calculated iris template need to be matched with the stored template. Matching algorithm that normally used are Hamming Distance, Weighted Euclidean Distance and Normalized Correlation. In the proposed system, SVM is used as pattern matching method to verify a person's identity based on the iris code. The following section will describe the SVM used as pattern matching.

### 5.1 Hamming distance (HD)

The Hamming distance (HD) gives a measure of how many bits are the same between bit patterns. Using the Hamming distance of two bit patterns, a decision can be made as to whether the two patterns are generated from different irises or from the same ones.

The Hamming distance,  $HD$ , is defined as the sum of disagreeing bits (sum of the exclusive-OR between  $X$  and  $Y$ ) over  $N$ , the total number of bits in the bit pattern, that is,

$$HD = \frac{1}{N} \sum_{j=1}^N X_j (XOR) Y_j \quad (7)$$

### 5.2 Weighted Euclidean Distance

Zhu et al. (2000) employed Weighted Euclidean Distance ( $WED$ ) for comparing two templates, especially if the template is composed of integer values. The weighting Euclidean distance gives a measure of how similar a collection of values are between two templates by using the following equation,

$$WED(k) = \sum_{i=1}^N \frac{(f_i - f_i^{(k)})^2}{(\delta_i^{(k)})^2} \quad (8)$$

where  $f_i$  is the  $i^{th}$  feature of the unknown iris, and  $f_i^{(k)}$  is the  $i^{th}$  feature of iris template  $k$ , and  $\delta_i^{(k)}$  is the standard deviation of the  $i^{th}$  feature in iris template  $k$ . The unknown iris template is found to match iris template  $k_o$ , when  $WED(k_o)$  has a minimum value.

### 5.3 Normalized Correlation

The normalized correlation (NC) between the acquired and database representation has been reported by Wildes et al. to ascertain the accuracy of matching (Masek, 2003). This is expressed as,

$$NC = \frac{\sum_{i=1}^n \sum_{j=1}^m (p_1[i, j] - \mu_1)(p_2[i - j] - \mu_2)}{nm\sigma_1\sigma_2} \quad (9)$$

where  $p_1$  and  $p_2$  are two images of size  $nm$ ,  $\mu_1$  and  $\sigma_1$  are the mean and standard deviation of  $p_1$ , and  $\mu_2$  and  $\sigma_2$  are the mean and standard deviation of  $p_2$ .

## 6. Iris-based verification system

Iris recognition utilizes the distinctive features of human iris in order to identify or verify the identity of individuals. Traditionally, iris recognition system has been used in high-security physical access application, for example in ATMs and kiosks for banking and travel application. Iris recognition's strength includes the following:

1. It has the potential for exceptionally high levels of accuracy.
2. It is capable of providing reliable identification as well as verification.
3. It maintains stability of characteristic over a lifetime.

Fig. 5(a) and Fig. 5(b) depict the essential steps of the proposed iris verification system. The system has two sub-systems: the iris training or enrollment phase and iris testing (operational) phase. The training phase is shown in Fig. 5(a).

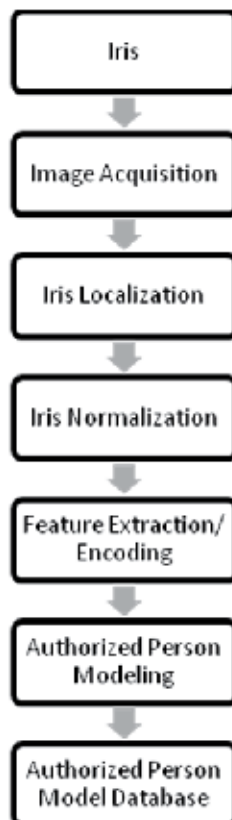


Fig. 5(a). Basic structure of iris-based verification system: Training phase.

In this phase, the authorized persons registered their iris images. Fig. 5(b) shows the process involved in the second phase. Thus, a person who wants to access the system is required to enter the claimed identity and his/her iris image. Furthermore, the captured iris image is processed and compared with the claimed person model to verify his claim. The iris testing phase has a decision process in which the system decides whether the extracted features from the given iris image matches with the model of the claimed person.

In order to give access or reject, a threshold is set. If the degree of similarity between a given iris image and its model than a given threshold, then the user will access the system, otherwise the user is rejected. The system computes the following decisions: the authorized person is accepted, the authorized person is rejected, the unauthorized person (impostor) is accepted and the unauthorized person (impostor) is rejected.

Taking the advantages of iris recognition systems which offer reliable and effective security in the present day, the use of iris-based verification system to identify the person's identity is also examined in this research work. This research adopts Support Vector Machines (SVMs) as pattern classification techniques which are based on iris code model to recognize an authorized user. The effectiveness of the proposed system is evaluated based on False Rejection Rate (FRR) and False Acceptance Rate (FAR).

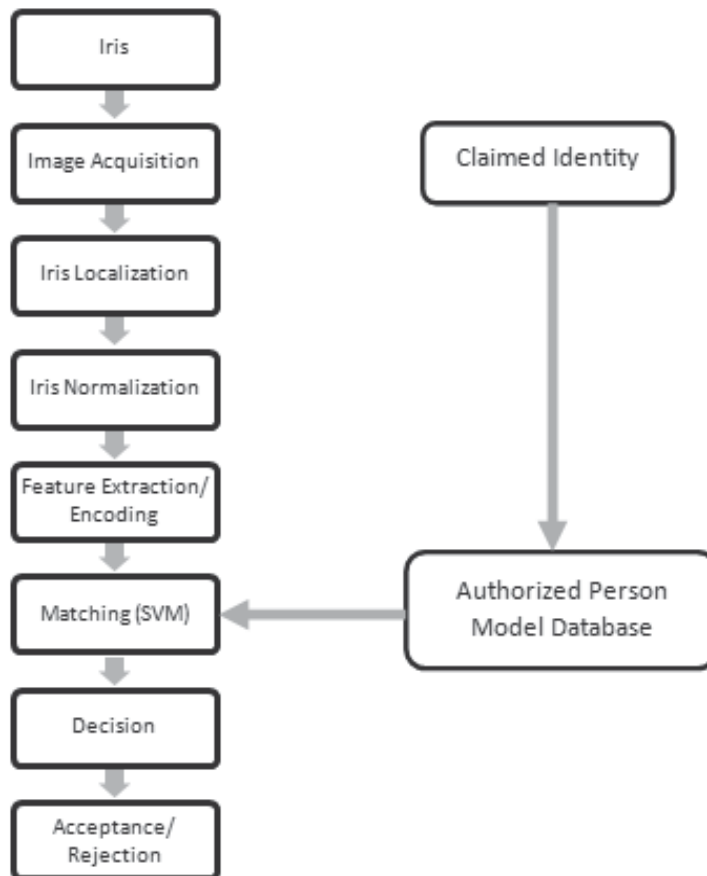


Fig. 5(b). Basic structure of iris-based verification system: Testing phase.

Furthermore, this research aims to develop a system that would have both low *FRR* and *FAR* so as attain both high usability and high security.

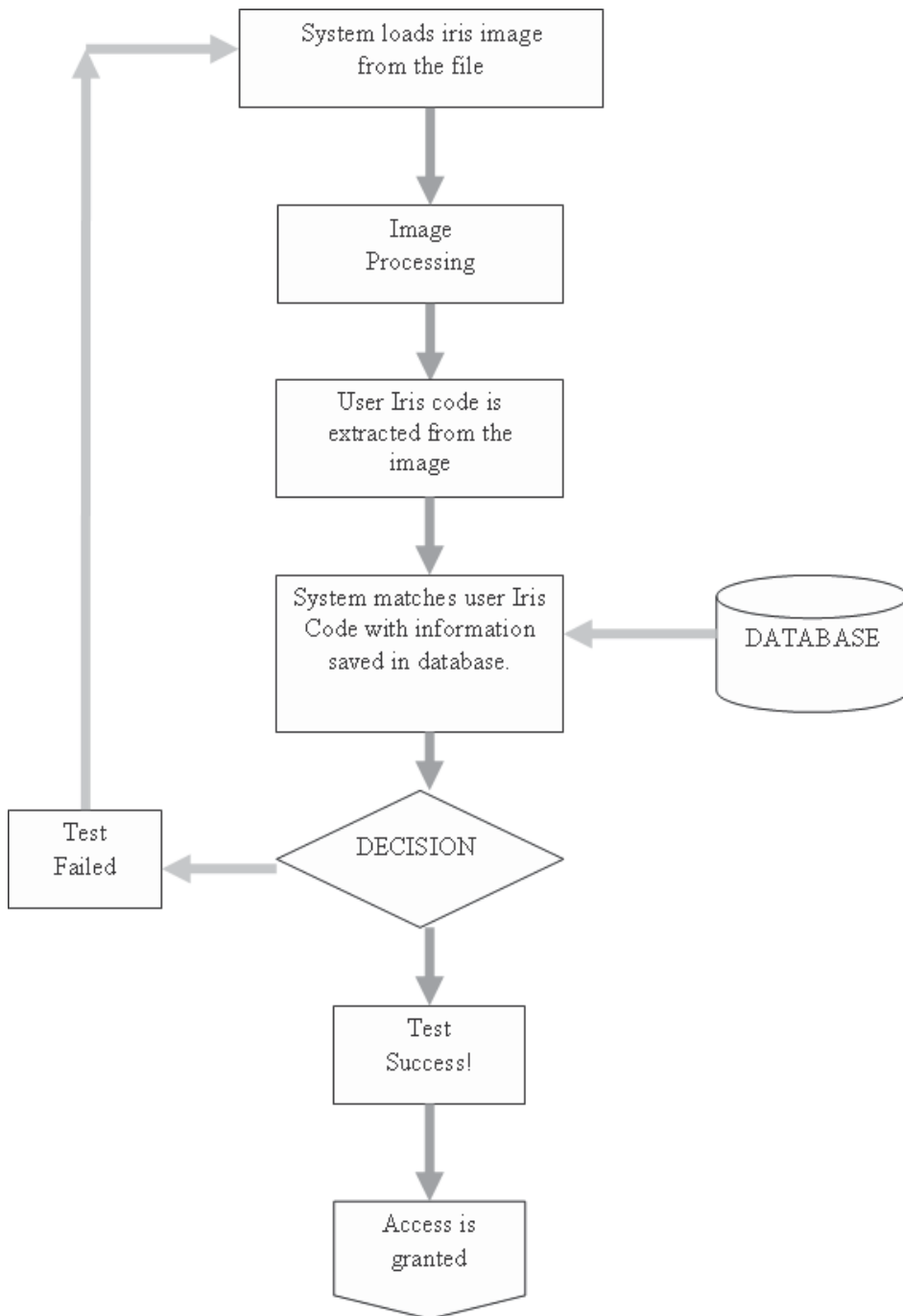


Fig. 6. Flowchart of iris-based authentication system.

## 7. Support vector machines

SVM is a relatively new learning machine technique, which is based on the principle of structural risk minimization (minimizing classification error). A SVM is binary classifier that optimally separates the two classes of data (Burges, 1998). There are two important aspects in the development of SVM as classifier. The first aspect is determination of the optimal hyperplane which will optimally separate the two classes and the other aspect is transformation of non-linearly separable classification problem into linearly separable problem. This section will discuss in brief the two aspects of the SVM development. Detail discussion on the SVM can be found in the introductory text by (Burges, 1998), for more detail description, (Cristianini, 2000). Fig. 7 shows linearly separable binary classification problem with no possibility of miss-classification data. Let  $x$  and  $y$  be a set of input feature vector and the class label respectively. The pair of input feature vectors and the class label can be represented as tuples  $\{x_i, y_i\}$  where  $i = 1, 2, \dots, N$  and  $y = \pm 1$ . In the case of linear separable problem, there exists a separating hyperplane which defines the boundary between class 1 (labeled as  $y = 1$ ) and class 2 (labeled as  $y = -1$ ). The separating hyperplane is,

$$w \cdot x + b = 0 \quad (10)$$

which implies

$$y_i(w \cdot x_i + b) \geq 1, i = 1, 2, \dots, N \quad (11)$$

Basically, there are numerous possible values of  $\{w, b\}$  that create separating hyperplane. In SVM only hyperplane that maximizes the margin between two sets is used. Margin is the distance between the closest data to the hyperplane.

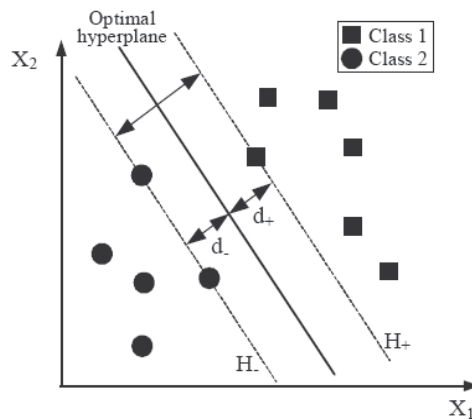


Fig. 7. SVM with Linear separable data.

Referring to Fig. 7 the margins are defined as  $d_+$  and  $d_-$ . The margin will be maximized in the case  $d_+ = d_-$ . Moreover, training data in the margins will lie on the hyperplanes  $H_+$  and  $H_-$ . The distance between hyperplane  $H_+$  and  $H_-$  is,

$$d_+ + d_- = \frac{2}{\|w\|} \quad (12)$$

As  $H^+$  and  $H^-$  are the hyperplane in which the closest training data to the optimal hyperplane, then there is no training data which fall between  $H^+$  and  $H^-$ . This means the hyperplane that separates optimally the training data is the hyperplane which minimizes  $\|w\|^2$  so that the distance of (12) is maximized. However, the minimization of  $\|w\|^2$  is constrained by (11). When the data is non-separable, slack variables,  $\xi_i$ , are introduced into the inequalities for relaxing them slightly so that some points allow to lie within the margin or even being misclassified completely. The resulting problem is then to minimize,

$$\frac{1}{2}\|w\|^2 + C(\sum_i L(\xi_i)) \quad (13)$$

where  $C$  is the adjustable penalty term and  $L$  is the loss function. The most common used loss function is linear loss function,  $L(\xi_i) = \xi_i$ . The optimization of (13) with linear loss function using Lagrange multipliers approach is to maximize,

$$L_D(w, b, \alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle x_i \cdot x_j \rangle \quad (14)$$

subject to

$$0 \leq \alpha_i \leq C \quad (15a)$$

and

$$\sum_i \alpha_i y_i \quad (15b)$$

where  $\alpha_i$  is the Lagrange multipliers. This optimization problem can be solved by using standard quadratic programming technique. Once the problem is optimized, the parameters of optimal hyperplane are,

$$w = \sum_i \alpha_i y_i x_i \quad (16)$$

As matter of fact,  $\alpha_i$  is zero for every  $x_i$  except the ones that lie on the margin. The training data with non-zero  $\alpha_i$  are called as support vectors. In the case of a non-linear separable problem, a kernel function is adopted to transform the feature space into higher dimensional feature space in which the problem become linearly separable. Typical kernel functions commonly used are listed in Table 1.

| Kernel       | $K(x, x_i)$   |
|--------------|---|
| Linear       | $x^T \bullet x_j$                                   |
| Polynomial   | $(x^T \bullet x_j + 1)^d$                           |
| Gaussian RBF | $\exp\left(\frac{-\ x - x_j\ ^2}{2\sigma^2}\right)$ |

Table 1. Formulation for Kernel function.

## 8. Performance measures

This proposed system in general makes four possible decisions; the authorized person is accepted, the authorized person is rejected, the unauthorized person (impostor) is accepted and the unauthorized person (impostor) is rejected. The accuracy of the proposed system is then specified based on the rate in which the system makes the decision to reject the authorized person and to accept the unauthorized person. False Rejection Rates (FRR) is used to measure the rate of the system to reject the authorized person and False Acceptance Rates (FAR) used to measure the rates of the system to accept the unauthorized person. Both performances are can be expressed as:

$$FRR = \frac{NFR}{NAA} \times 100\% \quad (17)$$

$$FAR = \frac{NFA}{NIA} \times 100\% \quad (18)$$

NFR is referred to the numbers of false rejections and NFA is referred to the number of false acceptance, while NAA and NIA are the numbers of the authorized person attempts and the numbers of impostor person attempts respectively (Zhang, 2000). Furthermore, low FRR and low FAR is the main objective in order to achieve both high usability and high security of the system.

## 9. Data set and experimental results

The Chinese Academy of Sciences-Institute of Automation (CASIA) eye image database is used in the experiment. To evaluate the effectiveness of the proposed system, a database of 42 grayscale eye images (7 eyes with 6 different images for each eye) was employed. About 30 grayscale eye images with 5 unique eyes are considered as authorized users and the others are impostors as shown in Fig. 8.

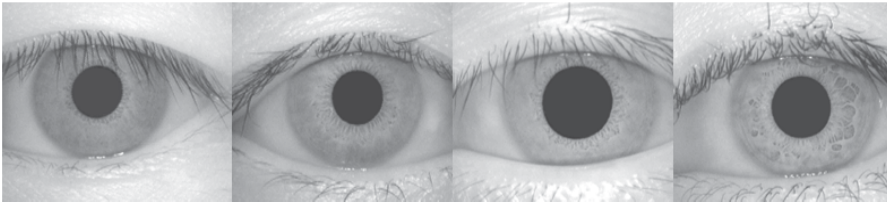


Fig. 8. Example of the iris images for different users.

For each eye, 6 eye images were captured in two different sessions with one month interval between sessions (three samples are collected in the first session and others three in second sessions) using specialized digital optics developed by the National Laboratory of Pattern Recognition, China. Infra-red lighting was used in acquiring the images, hence features in the iris region are highly visible and there is good contrast between pupil, iris and sclera regions.

As previously discussed, the performance of biometric systems is usually described by two error rates: FRR and FAR. Hence, the effectiveness of the proposed system in testing (operational) phase is evaluated based upon FRR and FAR values. The FAR is calculated

based on the close set and open set. In the close set, the typing biometric of an authorized person uses other authorized person identity. On other hand, the open set is referred as typing biometric of the impostors use authorized person.

The obtained feature vector of iris code comprises matrix of  $20 \times 480$ . This feature vector consists of bits 0 and 1. It was observed in all the experiments conducted that the feature vector size which containing high dimensionality often contributed to high FRR and FAR values with long processing time.

To overcome this problem, feature vector size is transformed to one-dimension vector which reduces to  $1 \times 480$  by using averaging techniques (each segment is divided by 20) contains the average value. Experimental results are subsequently discussed in the following section.

### 9.1 SVM-based iris code model

SVMs are classifiers which have demonstrated high capability in solving variety of problems that include the object recognition problems. Experimental results of training and testing based on iris code using SVMs are discussed.

In developing user models based on iris code, a SVM with polynomial kernel function of order 8 is used. Each authorized user has its own SVM-based model characterized by a set of support vectors. By using quadratic programming in the MATLAB environment, appropriate support vectors are determined. The penalty term  $C$  of  $10^{15}$  is used to anticipate misclassified data. Table 2 shows the training performance when the SVM is employed to develop user's models based on their iris code.

| Authorized User | Training Time(sec) | Classification Result (%) |
|-----------------|--------------------|---------------------------|
| User 1          | 0.0781             | 100                       |
| User 2          | 0.0781             | 100                       |
| User 3          | 0.0625             | 100                       |
| User 4          | 0.1094             | 100                       |
| User 5          | 0.0781             | 100                       |
| Average         | 0.0812             | 100                       |

Table 2. Training performances of iris code.

These results indicate that all of the SVM-based user models give perfect classifications as there are no errors in recognizing all the users. Besides, all of the SVM-based models can be trained in a very short time of about 0.1 second. Consequently, the SVM model should be further investigated and adopted for use in the proposed system.

A series of experiments is conducted using the testing data which have not been used during the training phase. Table 3 shows the testing performances of the SVM-based authorized user models.

The SVM-based authentication gives very good results for FAR of close set and open set conditions. This implies that the proposed system is well protected from attacking by impostors. In contrast, the FRR values are very high percentage with an average value of about 19.80%. Hence, the system seems to have poor usability. Experimental results show that the first, second and fourth users produce maximum FRR values of about 33%. Further study and improvement should be done before incorporating this model to the system.



| Authorized User | FRR (%) | FAR (%)<br>Close Set | FAR (%)<br>Open Set |
|-----------------|---------|----------------------|---------------------|
| User 1          | 33      | 0                    | 0                   |
| User 2          | 33      | 0                    | 0                   |
| User 3          | 0       | 0                    | 0                   |
| User 4          | 33      | 0                    | 0                   |
| User 5          | 0       | 0                    | 0                   |
| Average         | 19.80   | 0                    | 0                   |

Table 3. Testing performances of iris code.

## 10. Conclusion

This chapter has presented an iris recognition system, which was tested using database of grayscale eye images in order to verify the authorized user of iris recognition technology. Firstly segmenting method was used to localize the iris region from the eye image. Next, the localized iris image was normalized to eliminate dimensional inconsistencies between iris regions using Daugman's rubber sheet model. Finally features of the iris region were encoded by convolving the normalized iris region with 1D Log-Gabor filters and phase quantizing the output in order to produce a bit-wise biometric template.

The Support Vector Machine was adopted as classifier in order to develop the user model based on his/her iris code data. Experimental study using CASIA database is carried out to evaluate the effectiveness of the proposed system. Based on obtained results, SVM classifier produces excellent FAR value for both open and close set condition. Thus, the proposed system seems in a good level of security. However, further study has to be done to improve level of usability by reduce the value of FRR.

## 11. Acknowledgment

Portion of the research in this project work use CASIA iris image database (CASIA, 2003) collected by Institute of Automation, Chinese Academy of Science.

## 12. References

- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2 (2), 121-167.
- Chinese Academy of Sciences-Institute of Automation (2003). *Database of greyscale eyes images*. Version 1.0.  
<http://www.sinobiometrics.com>
- Daugman, J. (2002). How iris recognition works. Paper presented at the International Conference on Image Processing, Vol. 1, pp. 33-36.
- Daugman, J. (2003). The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36 (2), 279-291.
- Field, D. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, pp. 42-72.

- Hasimah, A. (2008). Design and development of integrated biometric system for high level security.(Master Thesis). Kulliyah of Engineering. International Islamic University Malaysia.
- Lim, S., Lee, K., Byeon, O., & Kim, T. (2001). Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal*, 23 (2), 61-70.
- Masek, L. (2003). Recognition of human iris patterns for biometric identification [B]. School of Computer Science and Software Engineering. University of Western Australia. Public source of the *MATLAB source code for iris recognition software* is available: <http://www.csse.uwa.edu.au/~masek101/>
- Shiavi, R. (1991). *Introduction to applied statistical signal analysis*. Homewood III.
- Widles, R., Asmuth, J., Green, G., Hsu, S., Kolezynski, R., Matey, J.,& McBride, S. (1996). A machine-vision system for iris recognition. *Machine Vision and Application*, Vol. 9, pp. 1-8.
- Wildes, R. (1999). Iris recognition: an emerging biometric technology. *IEEE*, 85 (9).
- Zhang, D. (2000). *Automated biometrics technologies and system*. Kluwer Academic Publishers.
- Zhu, Y., Tan, T., & Wang, Y. (2000). Biometric personal identification based on iris patterns. Paper presented at the 15<sup>th</sup> International Conference on Pattern Recognition, Spain, Vol. 2, pp. 801-804.

## **Part 4**

### **Other Biometrics**



# Verification of the Effectiveness of Blended Learning in Teaching Performance Skills for Simultaneous Singing and Piano Playing

Katsuko T. Nakahira<sup>1</sup>, Yukiko Fukami<sup>2</sup> and Miki Akahane<sup>3</sup>

<sup>1</sup>*Nagaoka University of Technology*

<sup>2</sup>*Kyoto Women's University*

<sup>3</sup>*Tokyo College of Music  
Japan*

## 1. Introduction

While difficulty in acquiring a skill in simultaneous singing and piano playing is attributable to both instructors and students, instructors have been taking a variety of approaches to remedy the situation at least in the piano playing skill. One hardware-based approach to improving teaching has been to develop a system that is installed in a music laboratory (ML) and enables the instructor to check the level of skill of each student even in a group lesson. A music laboratory is usually equipped with some tens of keyboards. In an ML, the instructor can listen to the piano playing of individual students, and give private advice to each student. As such, an ML was considered a pioneering educational system for group training of music. MLs have been used for the training of not only piano playing but also elementary music theory, including harmony, which can be tried out using keyboards. It was thought that an ML, in which each keyboard is used by one or two students, was effective in teaching elementary music theory and piano playing because the instruction through an ML can efficiently substitute for private lessons in these areas. However, the use of ML equipment is so complicated that it imposes a considerable burden on the instructor. Also, since students' performance can be checked only during the class hour, an ML is not practical at all for classes with a hundred or so students. Consequently, there is still considerable reliance on face-to-face training when it comes to the teaching of music.

Research efforts to improve teaching using software have mainly focused on a combination of two approaches: the development of appropriate computer software and the improvement of the teaching method itself. Approaches that involve the development of computer software include the following. In 1990, (Dannenberg et al., 1990) developed Piano Tutor, a computer-based piano tutoring system for beginner-level piano students. In 2000, (Hosaka, 2001) used audio-visual material. In 2001, (Matsumoto, 2001) used piano instruction material for computer-assisted instruction (CAI). In 2005, (Suzuki, 2005) developed Net-CAPIS, which introduced a connection to a network in a music laboratory (ML).

Attempts to improve teaching methods include the use of practice record cards ((Imaizumi, 2004)) and observation of others ((Nakajima, 2002)) . (Ogura, 2006) introduced blended learning using MIDI audio sources.

In spite of the various initiatives mentioned, the method most often used for teaching the still important subject of simultaneous singing and piano playing is one that uses self-learning records and the observation of others in classroom lessons, including face-to-face group lessons. This method is used not only in Japan but also in many other countries, such as China and Germany. There have been many practice-based studies on the advantages and disadvantages of group lessons as a means of teaching a musical performance skill. (Li & Kenshiro, 2003) reported on the status of group piano lessons given in teacher training colleges in China. They undertook a questionnaire survey with 169 students in teacher training colleges in China, and found that students preferred individual lessons to group lessons in learning piano playing. The reasons given were that (1) individual lessons provide detailed instructions, and (2) students can learn a wider variety of subjects more efficiently. Studies in Japan have been inconclusive about the advantages of group lessons. (Furukawa, 2003) recognized the advantages of group lessons but indicated that ultimately it was necessary to rely on individual lessons. (Nakagawa, 2007) suggested that group lessons have positive effects on students.

The above studies indicate that it is difficult to entirely eliminate some form of group lessons for the teaching of simultaneous singing and piano playing. The authors have considered that it may be possible to make group lessons as effective as individual lessons in improving students' skills by combining them with other methods or by modifying the way they are given.

In this chapter, we report on a training method in which students not only took group lessons (hereafter referred to as "face-to-face" training) but also were required to view e-learning material and to submit videos of their performance. We study the effects of the e-learning material by examining how the students' performance and perception changed after viewing the e-learning material. In addition, we indicate the limitations of non-face-to-face training and the need to combine such training with face-to-face training in what we call "blended" training.

## 2. Practice environment

We applied our method to the course "Music for Children I" at the Faculty of Developmental Education in Kyoto Women's College. The course lasts from April to July. This paper uses data for the courses conducted in 2006 and 2009. In both years, the number of students who took mid-term or final exams in piano-playing and singing was 102 for 2006 and 105 for 2009. Figure 1 shows the teaching model used in the course, which consists of (1) singing, (2) chord progression, and (3) simultaneous singing and piano playing. Each lesson lasted for 180 minutes. For the first half of the four-month course, one lesson consisted of 90 minutes of singing, 45 minutes of group lesson and 45 minutes of self-learning (individual performance practice). In the latter half of the course, one lesson consisted of 60 minutes of singing, 60 minutes of group lesson, and 60 minutes of lecture on chord progression. A midterm exam and a final exam were given at the end of May and in the middle of July respectively to examine performance skills for simultaneous singing and piano playing. Since 2006, KS20 (hereinafter referred to as "Kenshukun") from Company Fujifilm, has been used to enable students to submit videos of their own performance before and after performance exams, as was proposed by (Yokoyama et al., 2004) (There have been many papers reporting on the use of Kenshukun, such as (Nakahira et al., 2009)). Kenshukun is a video recording and playing device, and can be used for the creation of video content. The video format used is MPEG2. The filename of the video file in which the student recorded her performance can be printed as a bar code on paper. The student can review her performance right after she has recorded it by having the bar code of her file read by Kenshukun. For this experiment, endoscopic

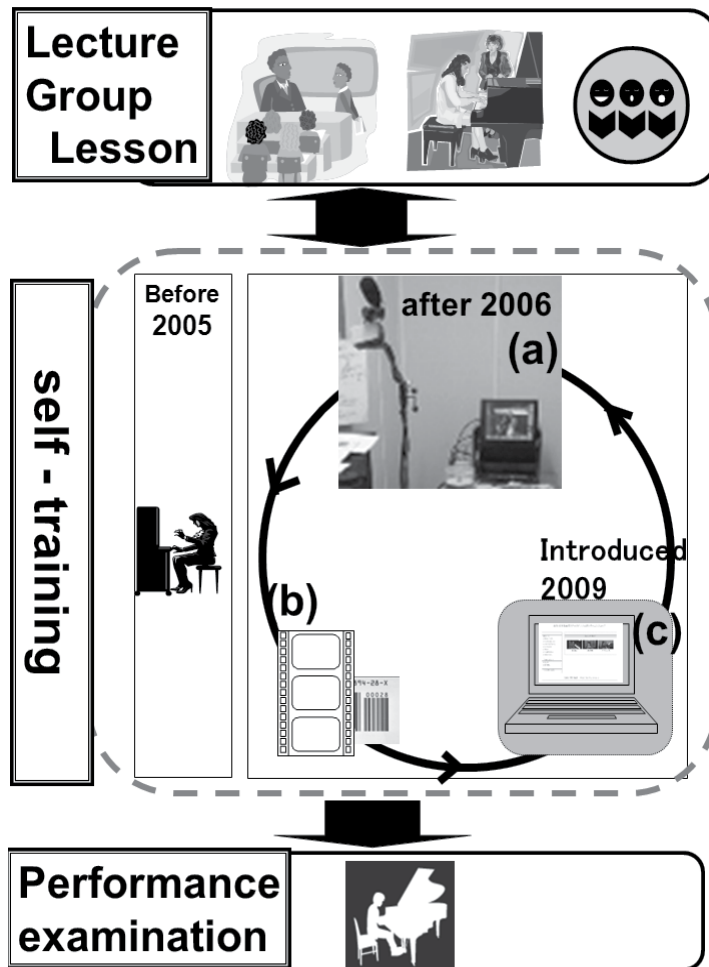


Fig. 1. Class Design

technology was adopted in the camera so that the position of its lens can be extended by 3.5 m to take a video of the student's facial expressions and hand movements on the keyboard of an upright piano from above. The students submitted paper with the bar code printed on it instead of actually submitting a video. Instructors can have a bar code read and in this way play the corresponding video.

The e-learning material (entitled "e-learning course on piano performance for teachers and pre-school teachers", developed by (Nakahira et al., 2008)) consists of four parts: (1) model performance of simultaneous singing and piano playing, (2) model performance of singing, (3) musical scores with annotations, and (4) FAQs for better singing. Part (1) contains videos of model performances of the seven carefully selected songs. Figure 2 shows the sample of e-Learning contents.

The model performance of each song is presented in the form of three videos, showing, respectively, finger movements, facial expressions, and the overall appearance. Part (2) contains videos of model singing. Additionally, it also provides short advice on important points in videos, in text and on scores. Part (3) contains PDF files of musical scores

教員・保育者養成のためのピアノ実技e-ラーニングコース  
ピアノ弾き歌い模範演奏



ピアノ弾き歌い模範演奏

- ・おててですくって
- ・おててですくって
- ・おててですくって
- ・おててですくって
- ・おててですくって
- ・おててですくって
- ・おててですくって

歌い模範演奏

- ・おててですくって
- ・おててですくって
- ・おててですくって


おててですくって


おててですくって

ワンポイントアドバイス

おててですくってのみました



まず、「すくって」の「す」や「のみました」の「の」のような無声音、そして「すくって」の促音(結まる音)などを意識して歌うようにしましょう(よりの歌唱へのFAG「日本語の発音のコツ」を参照してください)。このほかのフレーズ、たとえば1番の歌詞の「あとからあとからうってきて」や、5番の歌詞の「かきでもかき」や「あなたまは」などを部分的にしても、促音(結まる音)によく気をつけて発音するようお願いします。

また最初の高音をきれいに歌うには、その直前の「のみました」の部分で、息の支えをゆるめず、歌いきるまでしっかり保ち、次の「おてて」の低い音に声の響きがつかないように意識して歌いましょう。そしてフレーズは「おてて」に入る直前に、「のみました」の「た」の声をしっかりその小節に押しつけてから、鼻と口を両方閉じ、胸を開いた状態でさっと自然な息を入れるようにしましょう(よりの歌唱へのFAG「呼吸のトレーニング法」を参照してください)。

[Video](#) 動画による解説

Fig. 2. Contents example of piano skill e-Learning course for pre-teacher course students. (upper side)Model playing. (lower side)Guidance of singing.

with annotations. The information in these files is extraordinarily rich in detail for scores designed for teachers and pre-school teachers. Part (4) contains text with photos and singing performance videos both of which explain how to ensure good vocalization. The length of time for which each student viewed the e-learning material was recorded as part of his/her learning activity log. In this chapter, we analyze the performance videos and reports submitted before and after the viewing of the e-learning material by the top 15 students ranked in terms of the length of viewing time.



### 3. Analysis

| <i>S</i> | <i>LH</i> | <i>IP</i> | <i>CS</i>  |
|----------|-----------|-----------|--|
| A        | 15        | B         | Tempo became correct. Opened the mouth more widely and tried to sing more expressively. Looked more relaxed. Improvement in singing was greater than in piano playing.   |
| B        | 10        | B         | Tempo became correct. As a whole, looked anxious to sing carefully.  |
| C        | 1         | G         | While looking somewhat anxious to open the mouth widely, overall there was no significant improvement.   |
| D        | 9         | G         | Although the problem of not remembering lyrics was solved, overall there was no significant improvement.   |
| E        | 7         | B         | Voice became louder although expression was still poor.  |
| F        | 16*       | B         | Although voice became clearer and more articulate, singing became childish. Played the piano more carefully.   |
| G        | 10        | B         | Made improvements in the expression of glottal stops (a feature of the Japanese language), in distinction between loud and soft parts, and in expression of the words, "Annakoto Konnakoto."   |
| H        | 0*        | A         | Made a remarkable improvement in all aspects, particularly in singing. Despite the steady improvement, fluctuations in tempo were not corrected.   |
| I        | 7         | G         | The attempt to play the endings of phrases carefully backfired, causing fingers to move less smoothly. There was no significant improvement.   |
| J        | 4         | B         | While voice was still too soft, made an improvement overall.   |
| K        | 13        | A         | Learned a great deal from the e-learning material, and made a remarkable improvement. However, expression was too strong, which caused tempo to fluctuate. Looked anxious to mimic the model performance of simultaneous singing and piano playing. However, excessive eagerness to play well caused the body to sway back and forth. Singing reflected good understanding of the lyrics. Showed improved balance in volume between piano and singing. |
| L        | *         | B         | Made improvements, particularly in singing. Awkward movements of the left hand on the piano were not corrected.  |
| M        | 7         | B         | Looked anxious to sing expressively and to make a clear distinction between loud and soft parts.   |

Table 1. Examples of changes in quality of performance after viewing the e-learning material for more than 30 minutes. The index *S* means "students". The index *LH* means Learning History(total years). The students with (\*) in the learning history column took the course, "Introduction to Piano", and private lessons when they were first year students. Student H has carrire for studying piano before entrans to the University. Student L has no data. The index *IP* means improvement, which the rank *A* is excellent, rank *B* is good, rank *G* is newtral. The index *CS* means changes in the second recording compared with the first recording(students viewed the e-learning between the two recordings).

N 13 B Sang expressively. Tempo improved.

O 3 G Although there was some improvement in breathing in some parts of singing, still had breathing problems elsewhere.

### 3.1 Change after viewing the e-learning material

Table 1 summarizes changes that occurred in the case of the 15 students after viewing the e-learning material. Two students achieved a remarkable improvement, and 9 showed some improvement while the remaining 4 made little improvement. Generally, students made more improvement in singing than in piano playing. The learning history column shows the number of years for which the student had learned piano playing. The students with (\*) in the learning history column took the course, "Introduction to Piano", and private lessons when they were first year students.

The characteristics of the two students who made a remarkable improvement are as follows. Student H was a beginner in the piano. Although she mastered almost everything she could learn from the e-learning material, she seemed to be lost about how to practice in the future. Student K was an advanced learner of the piano. Watching the model performance enabled her to take a new look at her performance. As a result, she made remarkable progress in areas where she was able to build an image of how the performance should be. However, there were also areas where she was able to build an ideal image but failed to express it. Both students tended to be insensitive to fluctuations in tempo (Specifically, they gradually increased tempo).

### 3.2 Changes in scores from the mid-term exam to the final exam

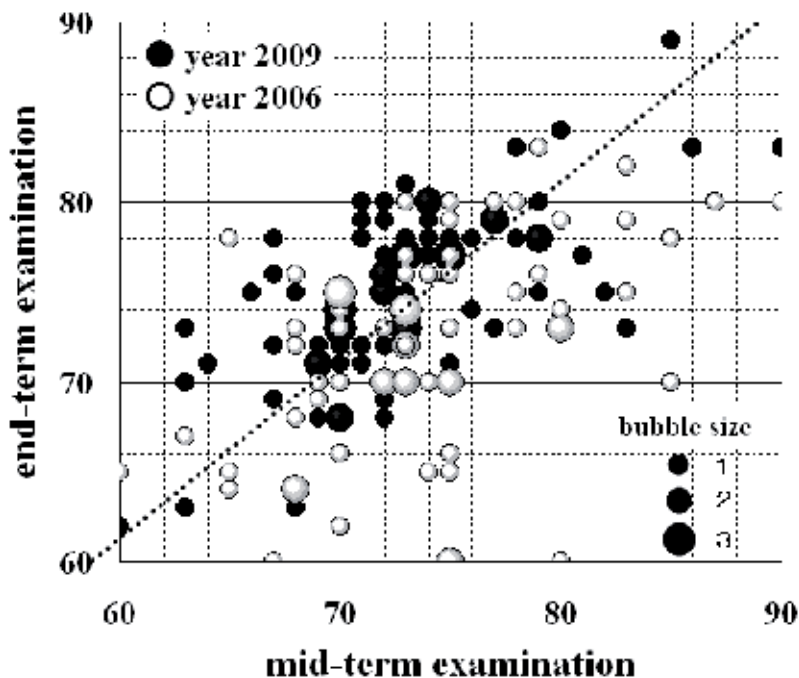


Fig. 3. distribution of  $s_m$  and  $s_f$ . The bubble size shows the number of person who got the score pair.

Through the experience, we show the change of the students' skill between mid-term examination and the end-term examination ((Nakahira et al., 2010a;b)). To examine how

|   | Length of time spent viewing the e-learning material |              |                  |       |
|---|--|--------------|------------------|-------|
|   | 0  | up to 30 min | more than 30 min |       |
| number of students                              | 34   | 53           | 18               |       |
| mid-exam  | $\bar{S}C$   | 72.26        | 73.08            | 73.78 |
|   | $\sigma$   | 6.21         | 4.20             | 7.32  |
| end-exam  | $\bar{S}C$   | 70.32        | 74.85            | 75.39 |
|   | $\sigma$   | 18.56        | 4.07             | 5.76  |
| Percentage of those whose score went up or down | up   | 53%          | 68%              | 67%   |
|   | down   | 29%          | 21%              | 28%   |

Table 2. Relationship between the results of the mid-term and final exams and the length of time of viewing the e-learning material.  $\bar{S}C$  means average score.

individuals fared, Figure 3 shows a two-dimensional distribution of the scores of the mid-term and final exams of individual students. Let  $s_m$  be the score of the mid-term exam, and  $s_f$  the score of the final exam. Define  $\delta$  as  $s_f - s_m$ . The horizontal axis in the figure shows  $\delta$  while the vertical axis shows the fraction of students for each value of  $\delta$ . The horizontal axis shows  $s_m$  while the vertical axis shown  $s_f$ . Therefore, the coordinate of a student can be expressed as  $(s_m, s_f)$ . The size of each bubble indicates the number of students on that coordinate. The dotted line shows  $s_m = s_f$ . If a bubble is on the upper left side of the line, its  $\delta$  is positive. The farther away from the line a bubble is, the greater the value of its  $\delta$ . Figure 3 clearly shows that, from 2006 to 2009, the average positions of the bubbles shifted to the upper left side of the line. These data indicate that the change in the teaching model from 2006 to 2009 resulted in a positive improvement in students' performance skills.

Table 2 shows the changes in scores from the mid-term exam to the final exam. Students could be classified into three groups: a group that did not view the e-learning material, a group that viewed the material for up to 30 minutes, and a group that viewed the material for more than 30 minutes.

In the mid-term exam, there was no significant difference between the three groups with regard to the average score. In the final exam, the difference was more pronounced. The group that had not viewed the e-learning material did not show any improvement in the average score in the final exam, and its standard deviation,  $\sigma$ , increased. The groups that had viewed the material achieved average scores which were 4 or 5 percentage points higher, and which had smaller standard deviations, than the group that did not view the material at all. With regard to individual results, a higher percentage of the students in the group that did not view the e-learning material saw no change in their scores between the mid-term and final exams than the percentage of those in the other two groups.

Considering the fact that there was no significant difference in the average score in the mid-term exam between the three groups, that almost all students submitted videos of their performance at least once, i.e., they had the opportunity to review their own performance by watching their videos, and that there was a pronounced difference between the three groups in the average score and the standard deviation,  $\sigma$ , it can be concluded that it was useful for students to view the e-learning material before they submitted their second performance videos.

|   | <i>MPSP</i> | <i>MPS</i> | <i>MS</i> | <i>FAQ</i> |
|---|-------------|------------|-----------|------------|
| tempo   | ✓           |            | ✓         |            |
| articulation                                      | ✓           |            | ✓         |            |
| image of the music                                | ✓           |            |           |            |
| facial expression                                 | ✓           |            |           |            |
| posture in playing                                | ✓           | ✓          |           |            |
| dynamic   | ✓           | ✓          |           |            |
| movement of finger                                |             |            | ✓         |            |
| balance of the volume between singing and playing | ✓           |            |           |            |
| length of note or rest                            | ✓           |            |           |            |
| balance of the volume between each hands plaing   | ✓           |            |           |            |
| pronunciation                                     |             | ✓          |           | ✓          |
| vocalism  |             | ✓          |           | ✓          |
| the mean of lyric                                 |             | ✓          |           |            |
| timing of the breath                              |             | ✓          |           |            |
| aware of missing point                            |             |            | ✓         |            |
| consciousness to breath                           |             |            |           | ✓          |
| recognize their own voice                         |             |            |           | ✓          |

Table 3. The items which evaluate useful by students. *MPSP*, *MPS*, *MS*, *FAQ* means each e-Learning contents title whether *MPSP* means model performance of simultaneous singing and piano playing, *MPS* means model performance of singing, *MS* menas musical scores with annotations, *FAQ* means FAQs for better singing.

## 4. Discussion

### 4.1 Effects of viewing the e-learning material

The reports submitted by the students indicate that those students who viewed the e-learning material for a long time became aware of many important points and were anxious to learn. In their reports, the 15 students whose reports were analyzed identified 17 points that they found useful in the e-learning material, such as dynamics, articulation, facial expression, performance posture, balance in volume between the piano and singing, tempo of each song, and vocalization. Table 3 shows the items which students feels useful.

In particular, they found the model performance of simultaneous singing and piano playing and that of singing especially useful. The comparison of performances recorded on videos before and after the viewing of the e-learning material shows that 11 out of 15 students made progress in performance, particularly in singing, in spite of the fact that they had not received any face-to-face lessons between the two recordings.

The above findings together with the results of the mid-term and final performance exams indicate that the viewing of the e-learning material helped beginners to build basic skills and advanced learners to acquire a higher level of skills in expression. As a whole, the material tended to be useful for improving singing. It is certain that providing appropriate e-learning material to enable students to learn singing theory and to enable a model performance to be imprinted on their memory, and requiring students to submit reports and videos of their performance are very useful in supporting students in self-training. In particular, the very fact that students with no piano playing background before entering the university were able to acquire almost all the basic skills in simultaneous singing and piano playing indicates that the use of e-learning material affords the possibility that it can make up for a shortage of time for face-to-face lessons.

#### **4.2 Needs of blending non-face-to-face training with face-to-face training**

While including the viewing of the e-learning material in the training of simultaneous singing and piano playing helped students improve many skills, there were certain skills that could not be improved just by viewing the e-learning material. Such skills include controlling the tempo of each song, articulation and vocalization, although students reported that the e-learning material did help them improve these skills. Specifically, the e-learning material was not helpful in solving the problems of fluctuations in tempo during performance, pronunciation of the nasal sonant in Japanese, placing of breaths, and excessive staccato in piano playing. In particular, students showed little improvement in solving the problem of fluctuations in tempo in spite of the fact the instructor clearly pointed out this problem by playing the videos of students' performances. Improvement in these skills can be better achieved through direct advice given during face-to-face lessons. While some students made a good improvement, some failed to do so even when they had viewed the e-learning material for the same length of time as the others. Differences in the amount of practice after the viewing may partially explain the difference, but it is essential that students develop the ability to perceive differences between their own performance and the model performance, and to apply what they have learned from the model performance to their own performance. Therefore, in face-to-face lessons, the instructor needs to train students in how to learn from the e-learning material.

### **5. Conclusion**

In this paper, we introduced a method of requiring students to view e-learning material and to submit videos of their performance. In this paper, we evaluate how this method improves students' performance skills, based on the length of time the students spent in viewing the e-learning material, the reports submitted by the students, the videos of the students' performance taken in the course of their regular practice, and the results of their mid-term and final performance exams. It is shown that (1) the combination of the viewing of the e-learning material and the submission of performance videos, which are both non-face-to-face training, encourages the students to undertake self-training and can considerably improve their performance skills, and that (2) it is necessary to provide face-to-face training for certain skills that cannot be improved through non-face-to-face training.

### **6. Acknowledgement**

This work was supported by 2009-2011 Grant-in-Aid for Scientific Research(C)(10283053, chief researcher: Yukiko Fukami), and the contract research fund by FUJINON Corporation. We also thanks for Kazumasa Toyama for grateful helps.

## 7. References

- Dannenberg, S., Jheph, C. & Joseph, S. (1990). A computer-based multi media tutor for beginning piano students, *Interface Journal of New Music Research* 19(2-3): 155–173.
- Furukawa, M. (2003). Considerations on music education in the training of pre-school teachers: mass training of piano beginners using children's songs, *Japan Society of Research on Early Childhood Care and Education Annual Report* 56: .870–871.
- Hosaka, T. (2001). The use of audio visual education materials for piano study, *Journal of Chikushi Jyogakuen Junior College* 35: 115–128.
- Imaizumi, A. (2004). A trial of teaching piano playing to students with no experience–group lesson using keyboard pianos: (2) introduction of practice record cards, *Japan Society of Research on Early Childhood Care and Education Annual Report* 57: 281–282.
- Li, Y. & Kenshiro, T. (2003). A study of piano group lessons in normal universities of the people's republic of china: The history and the present situation, *Bulletin of Tokyo Gakugei University, Arts and sports sciences* 57: 33–46.
- Matsumoto, T. (2001). The status of beginners in piano playing regarding the acquisition of basic skills and the status and issues of cai learning using computer systems, *Yojikyoku Special Issue* pp. 42–56.
- Nakagawa, K. (2007). Effectiveness of group lessons on singing in teacher training colleges from the perspective of the “effectiveness in instruction on performance skills”, *Study of School Music Education* 11: 69–70.
- Nakahira, K. T., Akahane, M. & Fukami, Y. (2008). Development e-learning contents with blended learning for teaching piano singing and playing piano, *JSiSE Research Report* 23(1): 85–92.
- Nakahira, K. T., Akahane, M. & Fukami, Y. (2009). Verification of the effectiveness of blended learning in teaching performance skills for simultaneous singing and piano playing, *the Proceedings of the 17th International Conference on Computers in Education* pp. 975–979.
- Nakahira, K. T., Akahane, M. & Fukami, Y. (2010). Faculty development for playing and singing education with blended learning, *the Journal of the Japan Society for Educational Technology* 34: 45–48.
- Nakajima, T. (2002). A practical study on piano teaching method at the department of education–for acquirement of musical capacity, *Studies on Educational Practice, Center for Educational Research and Training, Faculty of Education, Shinshu University* 3: 31–40.
- Ogura, R. (2006). Making practical use of performance data on midi in music lessons: Uses of network and a floppy disks, *Annual report of the Faculty of Education, Bunkyo University* 40: 43–53.
- Suzuki, H. (2005). “e-learning” in piano teaching –practice in the center for practical education –, *Journal of practical education* 19: 11–22.
- Nakahira et al. (2010) *The Use of Blended Learning in the Teaching of Piano Playing and Singing in Education Faculties, the Proceedings of the 2nd International Symposium on Aware Computing* 19: 227-232.
- Yokoyama, J., Matsuda, S., Nakahira, K. T. & Fukumura, Y. (2004). Development of a lesson support tool allowing easy handling of multimedia, *IPSI SIG Technical Report* 2004(117): 61– 66.

# Portable Biometric System of High Sensitivity Absorption Detection

Der Chin Chen  
*FENG CHIA University,*  
*Taiwan, R.O.C*

## 1. Introduction

The traditional bio-chemical detecting methods include two methods. One method is to use a meter to measure the change of voltage that is converted from the bio-chemical energy. The other method is to use the testing agent applied on the sample. For example, a chemical testing agent or testing paper can be used to measure the density of a target object in the sample. After which, a quantitative analysis can be done by using some optical techniques. The first method is easy, compact, fast, and no pollution generated by the testing agent or paper. However, its disadvantages include low precision and not suitable for repeated testing. The second method is good in the quantitative analysis and suitable for mass detection. It is widely used in different automatic detecting equipments in the bio-chemical industry. Also, it is the major method used in current medical and bio-chemical related fields. But, its disadvantages include the testing device is complex, it is not suitable for dynamic testing, it will cause certain pollution from the testing agent, and its operation is complicated. The basic principle of the optical technique is to measure the absorbance as a scale to determine the density of a specific colored object in the sample. Furthermore, it can be classified into the following methods, such as colorimetric analysis method, spectral analysis method, fluorescent analysis method, turbidity analysis method, etc. The devices about all these methods are quite similar to the conventional spectrometer.

The current commercial spectrometer for bio-chemical testing has many kinds. The large spectrometer is expensive and occupies a huge space. However, the micro spectrometer is light-weighted, small, fast, easy to operate, suitable for mass detection and non-expensive [1-2]. But, its precision and sensitivity is not good due to the technical limitation of its micro light detector. So, it is still not suitable for most bio-chemical testing. The conventional spectrometer main specification was shown in Table 1.

Referring to Fig. 1, it illustrates the structure of a conventional micro spectrometer. It contains two parts, namely the optical system and the electrical system (not shown). The optical system includes a traditional light source having the tungsten filament, a condenser and filter, spatial filter, a self-focusing blazed reflection grating, and a linear CCD detector. The traditional light source provides an enough light (or beam) for the micro spectrometer and will cover the wavelength range of the linear CCD detector. The condenser can collect the incoming light into the micro spectrometer. Light from the input fiber enters the optical bench through this condenser. Also, the numerical aperture (N.A.) value of the condenser should match with the one of the input fiber. The filter is a device that restricts light to pre-determined wavelength

regions. The function of a self-focusing blazed reflection grating is to separate the light with different wavelength ranges and to focus the reflected lights to the linear CCD detector. The linear CCD detector can detect the light intensity and distribution and then converted into electrical signals for further analog output processing. Each pixel on the linear CCD detector responds to the wavelength of light that strikes it, creating a spectral response.

No matter the conventional spectrometer is the large one or the micro one, the container must be the standard rectangular quartz made container (the bottom is one centimeter square and the depth is one centimeter, so the volume is 1 cc as shown Fig. 2). Two sizes are available; 3.5ml (standard) and 1.5ml (semi-micro). All containers have a 10mm path length and are 45mm high. Therefore, they require sample's volume is relative large. The optical requirement is as follows: parallelism <30 arc sec, perpendicularity <10 arc min and flatness < 2λ @ 632.8 nm. Assume that the user is located in an area or country where it is under the danger of Severe Acute Respiratory Syndrome (briefly called SARS hereafter) virus and H1N1. If the user (of a medical organization) needs to collect the sputum of a patient, the user has to collect at least 1 cc to conduct a bio-chemical testing. In case the user only collects 0.5 cc of the patient's sputum, the test cannot be done by the conventional spectrometer because the volume is fewer than the require minimum 1 cc. Hence, the user will miss the best time for determine whether this patient is a SARS patient or not. It not only is disadvantageous for the patient, but also increases the uncertainty and panic for the society. However, if the user simply reduces the required volume down to 0.5 cc, it will cause another problem that the sensitivity of the entire system becomes one half of the original one. Therefore, the testing becomes unreliable. Besides, because different coloring agents have different light absorption characteristics, an additional filter has to be added on the conventional light source so that a particular wavelength range can be selected. However, after using the filter for a long period, it will cause the over-heating problem.

Thus, the disadvantages of the conventional spectrometer can be summarized as follows: [a] the required volume of the sample is relatively large, [b] the entire optical system is complex and expensive, [c] the sensitivity of one-path penetration through the sample is low, and [d] the conventional tungsten-typed light source needs additional filters to solve the over-heating problem. Due to drawbacks mentioned above, we are motivated to explore new technique named portable biometric system of high sensitivity absorption detection. The new technique covers three systems: the double optical path absorbance biometric system that can be applied in practices, the multi optical path absorbance biometric system and the multi-color sensor measurement system which are in experimental stage. The measured sample by the first two systems is blue latex and the measured sample by the last one is color plastic ball.

|                     |   |
|---------------------|---|
| Sensor range:       | 200-1100 nm   |
| Sensor:             | 2048-element linear silicon CCD array                                 |
| Gratings:           | 14 gratings; UV through Shortwave NIR                                 |
| Entrance aperture:  | 5, 10, 25, 50, 100 or 200 mm wide slits or fiber (no slit)            |
| filters:            | Installed long-pass and band-pass filters                             |
| Focal length:       | f/4, 42 mm (input); 68 mm (output)                                    |
| Optical resolution: | 0.3~10.0 nm FWHM (depending on grating and size of entrance aperture) |

Table 1. Conventional spectrometer main specifications.



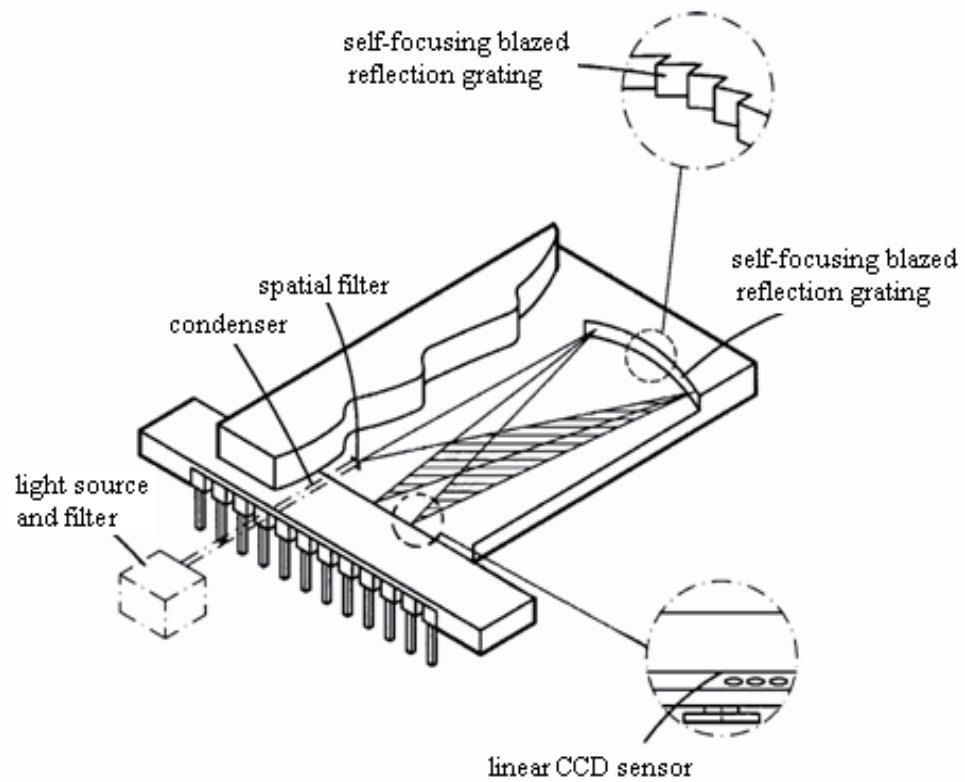


Fig. 1. Spectrometer with components.

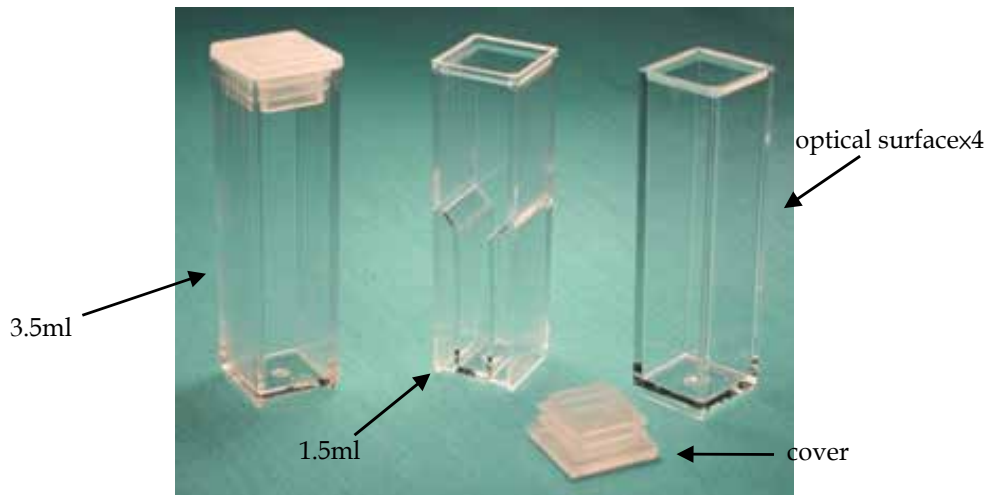


Fig. 2. Containers are designed to fit all standard instruments.

## 2. Principle

The portable biometric system presented in this paper involves two principles. The first principle is the biological principle of Enzyme-Linked Immuno Sorbent Assay (ELISA), the second principle is the electro-optical technique used in this system. Performing an ELISA involves at least one antibody with specificity for a particular antigen. The sample with an unknown amount of antigen is immobilized on a polystyrene microtiter plate either non-specifically (via adsorption to the surface) or specifically (via capture by another antibody specific to the same antigen, in a "sandwich" ELISA). After the antigen is immobilized the detection antibody is added, forming a complex with the antigen. The detection antibody can be covalently linked to an enzyme, or can itself be detected by a secondary antibody which is linked to an enzyme through bioconjugation. The enzyme use in this system is the gold colloid, which has the property of light absorption. The system detects the intensity change of the visible collimated beam, which passes through the sample solution twice, to know the absorbance of the sample solution. The detective object of this system is the gold colloid that the sample produces after the biochemical reaction-immunology. There is a certain relationship between the color of the gold colloid and its concentration. By measuring the particular wavelength of the optical absorbance of the gold colloid, the concentration of the waiting measured sample can be obtained. The concentration of chemical compositions can be further obtained by establishing the relationship table between the optical absorbance of the gold colloid and the concentration of chemical compositions. From the Beer-Lambert law (as shown Fig. 3), the amount of radiation absorbed is represented in the follow equation:  $A = \epsilon bc$ , where  $A$  is absorbance (since  $A = \log_{10} \left( \frac{P_0}{P} \right)$ );  $P_0$  is the beam of radiation entering the sample solution while  $P$  is the beam of radiation exiting the sample);  $\epsilon$  is the molar absorptivity with units of  $L \text{ mol}^{-1} \text{ cm}^{-1}$ ;  $b$  is the path length of the sample-that is , the path length of the container in which the sample is contained, we will express this measurement in centimeter;  $c$  is the concentration of the compound in the solution, expressed in  $\text{mol L}^{-1}$ . From above we can know that increasing the path length of the solution or increasing the monochromatic optical path in the waiting-measured sample solution can increase the optical absorbance. The processes of double optical path increase the opportunities that the collimated beam is absorbed which raise the detection sensitivity of this system. The method that allows the visible collimated beam to pass by the sample solution twice is the use of prism array.

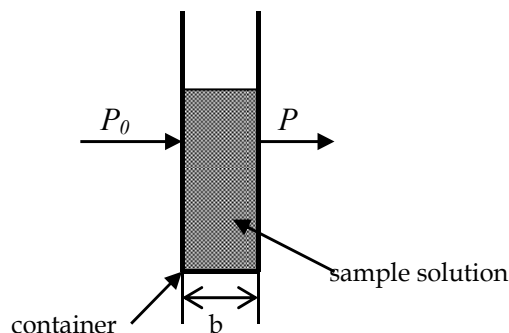


Fig. 3. The absorbance of sample solution.

### 3. The portable biometric system

The systematic construction of portable double optical path absorbance biometric system is shown in Fig. 4[3-4]. Increasing the path length of the sample increases the amount of samples needed resulting in increased testing cost. But sometimes we are absolutely not able to obtain so many samples. However, this defect can be improved by using the double optical path device. The biometric system includes : (a) optical transmitter/receiver unit,(b) micro container as shown in Fig. 5, and (c) double optical path unit as shown in Fig. 6. The optical transmitter/receiver unit includes LED light source, current driver, LED's selector, photodiode, operator amplifier, and signal processor. The double optical path unit includes input fiber, cube beam splitter, two way fiber, collimator, micro container, output fiber, and corner prism array. With regard to the micro container, it has a circular recess with predetermined depth H and with a predetermined diameter D for receiving a sample. The micro container has a first surface and a second surface. The micro container is made from transparent optical plastic plate by UV laser lithography. This optical plate is made by PMMA. The micro container is a circular hole with diameter (D) between 4 to 6 mm and a depth (H) between 1 to 5  $\mu$ m. The corner cube array is disposed near the first surface of the micro container. This corner cube array has several micro corner cubes so as to reflect an incoming beam back along its original path of the incoming beam. The collimator is disposed at a position outside the second surface of the micro container. The beam splitting

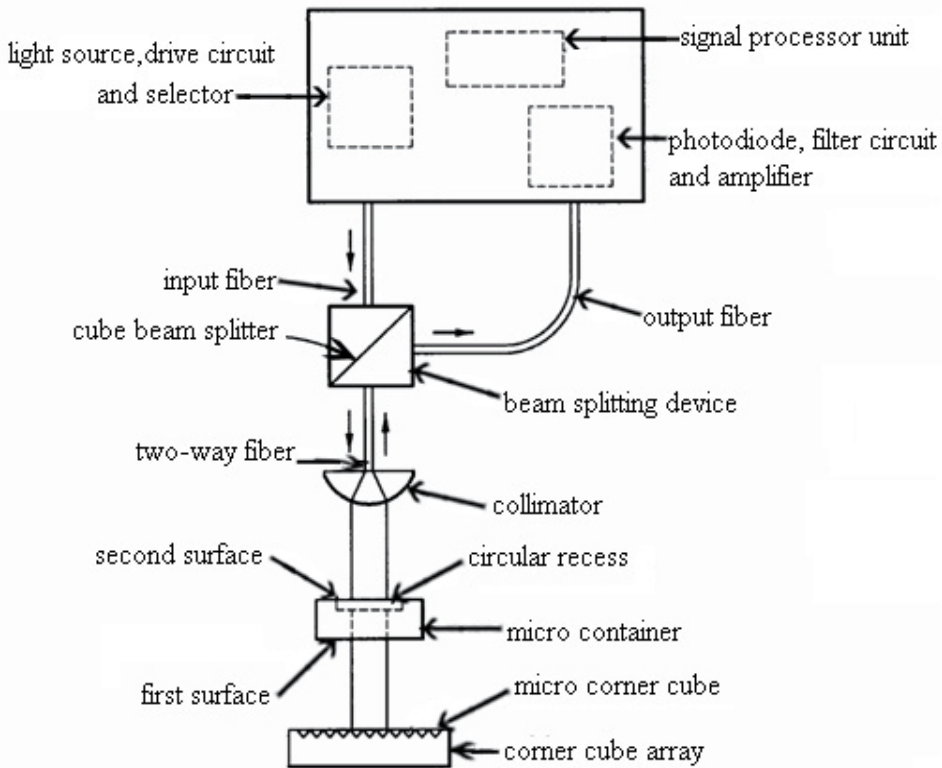


Fig. 4. Portable double optical path absorbance biometric system.

device has a cube beam splitter, an input fiber, an output fiber, and a two-way fiber. The light source selector is used for providing a first beam that has a predetermined wavelength range and intensity.

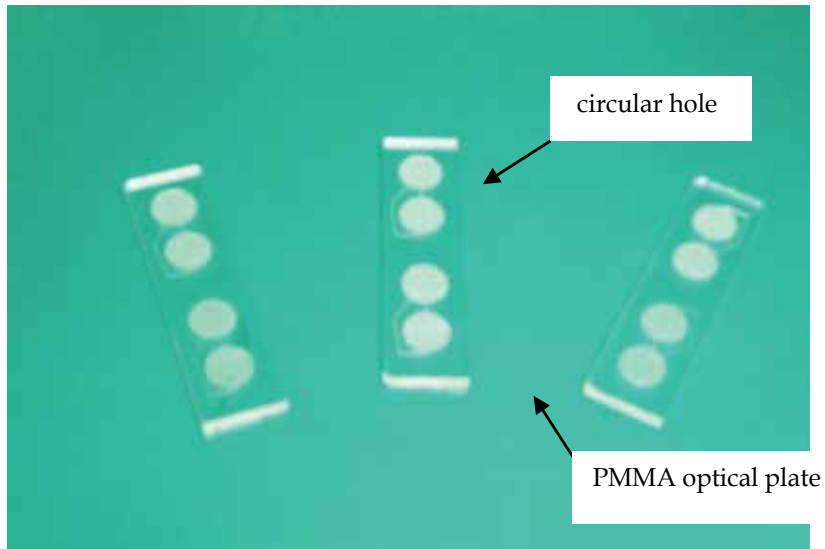


Fig. 5. Micro container.

The light source selector includes an electrical switch, a LED array that has a plurality of different colored LEDs, a plurality of middle fibers, and a light coupler to guide a selected colored LED light into the input fiber so that the first light beam having a desired wavelength, as shown in Fig. 7. Therefore, there are many wavelength ranges that can be chosen. Its advantage is to allow the user to execute one of many different bio-chemical tests by selecting a suitable wavelength range, such as to detect the density of blood sugar,

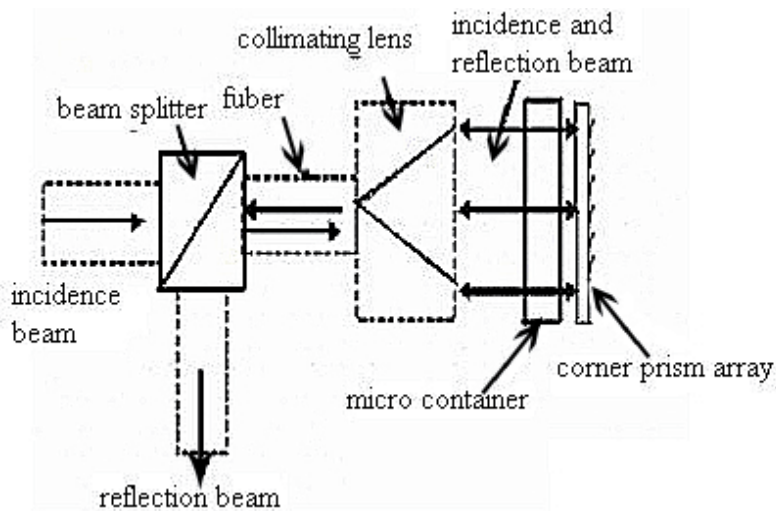


Fig. 6. Double optical path unit.

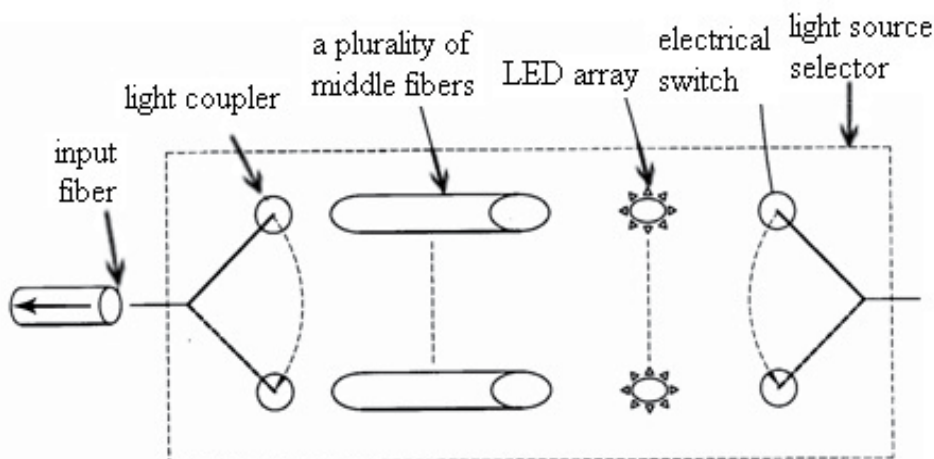


Fig. 7. Light source selector.

the density of a specific influenza virus, etc. Thus, this invention is multi-functional. Concerning the detector, it is used for detecting a final intensity of a received beam. The signal processor unit is used for comparing the detected final intensity of the sample with reference intensity so that an absorbance of the sample can be calculated. Therefore, the input fiber guides a first beam passing through the cube beam splitter of the beam splitting device and then through the two-way fiber to arrive to the collimator so as to expand as a second beam having another diameter approximately equal to said predetermined diameter; then the second beam continues to penetrate the micro container and becomes a third beam (The third beam is weaker than the second beam because some of the beam is absorbed by the sample). After reflecting by the corner cube array and passing through the micro container again, a fourth beam (the fourth beam is weaker than the third beam) is obtained. After passing through the collimator, the fourth beam becomes a collected fifth beam. The fifth beam enters the two-way fiber to arrive to the cube beam splitter of the beam splitting device and then to be reflected aside into the output fiber and finally to the detector so that the signal processor unit can calculate an absorbance of the sample, as shown in Fig. 8. About the measurement principle of this research, once the beam penetrates the sample one time, the passing beam will be weakened. Thus, if it penetrates the sample twice, the total amount of decaying will be doubled. So, by measuring the amount of decaying, the absorbance can be determined. The corner prism array consists of many corner cube refractor. In this system, the sample contains a specific colored material (or label)/agent that already chemically reacted with an object that the user wants to measure. The shade of the colored sample is proportional to the degree of density of such object. By establishing the relationship between the shade and the density of the colored sample after chemically reacted with a suitable coloring agent (or label), the density of the sample can be determined by detecting its actual shade of the sample. In a bio-chemical test, usually a suitable coloring agent will be added to react with the object that we want to measure. Therefore, the shade of such color means the density of the object. The shade can be precisely measured by the

voltage output of the detector. Of course, a graph about such relationship between the shade and the density can be established. A corner cube reflector is a retroreflector consisting of three mutually perpendicular, intersecting flat surfaces, which reflects parallel light ray back towards the source as shown Fig. 9. Fig. 9 (a) is contour of corner prism array. Fig. 9 (b) shows that corner cube are used to reflect light beam back to the original direction.

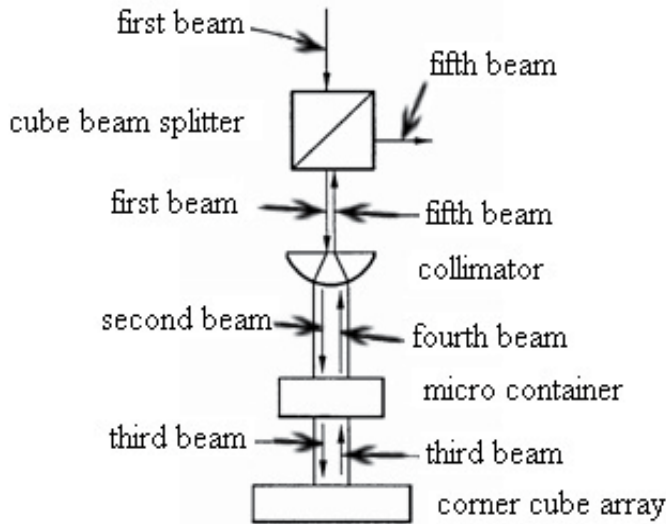


Fig. 8. The layout of double optical path.

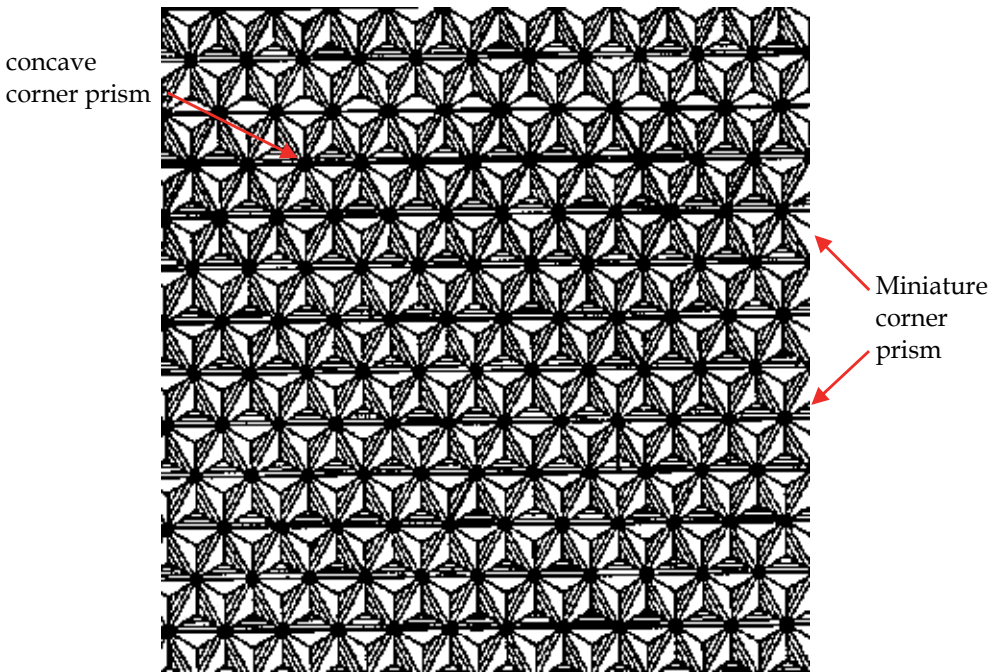


Fig. 9. (a) Contour of corner prism array (from US patent 6,206,565 B1).

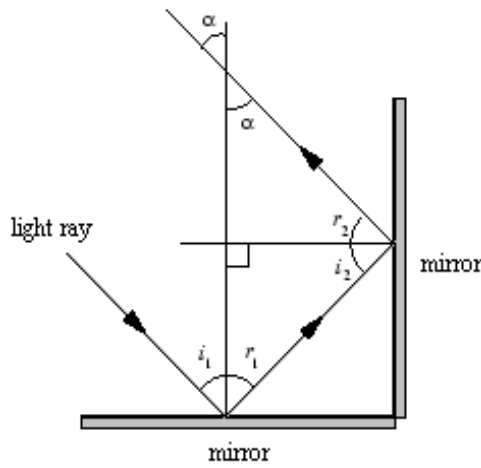


Fig. 9. (b) Parallel light path reflected by a corner cube.

A light ray incident from any direction in the plane perpendicular to both mirrors is reflected through 180°. We are effectively being asked to prove that  $\alpha=i_1$ , for any value of  $i_1$  from trigonometry and the law of reflection. The light emitting diode (LED) is used as the light source in the optical transmitter/receiver unit. The LED emits the red light which passes through the collimating lens and becomes a collimated beam. After this collimated beam passes through the sample, it will be reflected back by corner prism array, and pass through the sample again. The reflected beam is focalized by collective lens of the receiver and be detected by the photodiode. By using the collimated beam with the corner prism array, the system does not require a very sensitivity optical alignment. The collective lens used in this system can form a sharper spot size and image all of the light emitted into the much larger area of it, therefore we can predict that the image illuminance will increase dramatically. We begin to calculate the axial image illuminance. An object region of area  $A_O$  emits rays that are focused into an image region with area  $A_I =M^2A_O$  is shown in Fig. 10.  $M$  is linear magnification of collective lens. The image illuminance  $E_I$  is the total flux divided by the image area. The flux is the product of the object’s axial luminance,  $L_O$ , the object area,  $A_O$ , and, the solid angle subtended by the lens,

$$\Omega_L = \frac{\pi D^2}{4S_O^2} \tag{1}$$

Thus the image illuminance is

$$E_I = \frac{L_O A_O \Omega_L}{A_I} = \frac{\pi D^2 L_O}{4M^2 S_O^2} = \frac{\pi D^2 L_O}{4S_I^2} \tag{2}$$

Using thin lens equation, the linear magnification,  $M$ , and the F-number,  $N$ , we can transform Equation (2) into a more useful form which incorporates the lens F-number  $N$

$$E_I = \frac{\pi D^2 L_O}{4f^2(1-M)^2} = \frac{\pi N^2 L_O}{4(1-M)^2} \tag{3}$$

Image illuminance is proportional to the object's axial luminance. The proportionality factor is a square of the lens F- number ( $F/\#$ ), which can usually be read off the side of the lens barrel, and the magnification, which can be calculated by dividing the image size by the object size.

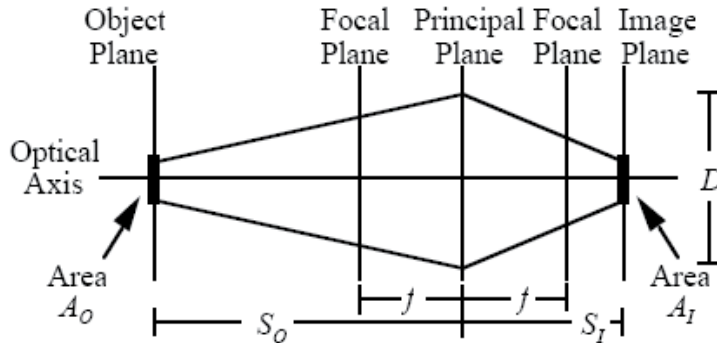


Fig. 10. Geometry of image illuminance (from "Introductory Optic" Ver. 2.04 1986-1987 Norman Wittels).

The system will also not suffer the effect of environment interference because of the use of pulse width modulation LED light. The peak emission's wavelength of the LED is matched with the peak absorption's wavelength of the Blue-Latex. The advantages for using the LED over the general light source such as household tungsten lamp are the small size, short response time, long lifetime, stable performance, etc. Also, the cost of design for the driver circuit using LED is lower than that of the laser-diode. As the peak emission wavelength  $\lambda_p$  of LED has to be matched with the peak absorption wavelength  $\lambda_p$  of the sample, we use the peak emission wavelength of the red light of 656 nm as the light source in this optical-transmitter. LED is a kind of current-driven device, not a voltage-driven device. So, in order to generate stronger light beam for LED, we should provide larger forward bias current  $I_F$ , but the light beam intensity is not proportional to the provided current. As the forward bias current  $I_F$  becomes larger, the dark current also becomes larger. When  $I_F$  becomes larger and larger, the LED will go into the saturation state, and then it will be burnt out. Also, if LED works under the larger current for a long time, it will become overheated and age too soon. To prevent LED from the aging problem, we use the pulsed current to drive LED, i.e. at the fixed duration, the constant current is provided. We use the inverter (CD4069) with the capacitors to form the multi-vibrating circuit. The pulse light output is 46 $\mu$ sec pulse duration which is obtained with the LED driving circuit, as shown in Fig. 11. [5] Fig. 12 shows the light output obtained with the LED circuit. The collimated beam of this device passes by the sample back and forth. Both the incident beam and the reflected beam are guided by the optical fiber. The reflected light will be focused on the detector by collective lens. The optical received terminal includes photodiode, filter circuit and amplifier, as shown in Fig. 13. [6] A potentiometer is placed before the filter circuit. By changing this, we can trim the sensitivity of the detector. To prevent the noise, such as the signal from the lighting lamp, from being amplified with the signal, we use a high pass filter comprised of capacitors and resistors to filter out the 120 Hz signal. Owing to the current generated by the photodiode is very small, and in order to be easily analyzed, we need to amplify the output voltage. The signal is amplified by using the amplifier LM324M and negative feedback. The



ambient electric-magnetic interference (EMI) will effect the detection in the system. To prevent the EMI noise, we add a metal shield outside the circuit. The beam of the visible LED will be split into two beams. The first beam will go directly into the collimator to form the collimated beam. The collimated beam will pass through the samples, be reflected by the prism arrays to pass though the sample a second time before being finally detected by the photodiode. The other beam will go into the collimator, form collimated beam, and then be detected by the other photodiode without passing through the sample.

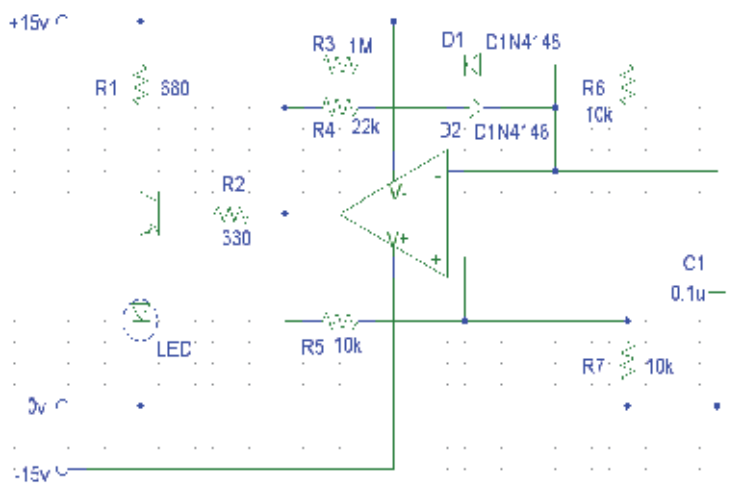


Fig. 11. LED drive circuit.



Fig. 12. The light output obtained with the LED circuit.

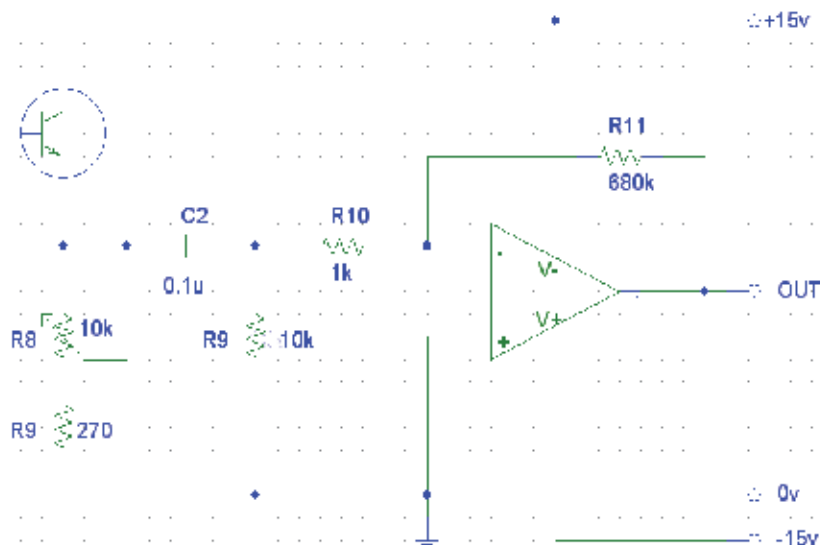


Fig. 13. Receiver circuit.

#### 4. Experimental results

The new technique covers three important systems; these experiments are described as follows:

Experiment 1: the double optical path absorbance biometric system

The double optical path absorbance biometric system is illustrated as Fig. 4 and double optical path unit as Fig. 6. The experimental parameters are as follows: The diameter of the optical fiber used as the light guidance is 1.00 mm. The NA (numerical aperture) of multi-model plastic fiber is 0.44. The peak emission wavelength ( $\lambda_p$ ) of LED is 656nm and spectral bandwidth ( $\Delta\lambda$ ) is 30nm, as shown in Fig. 14. The detector is Silicon photodiode. Blue-Latex solution has two peaks which absorption wavelength ( $\lambda_p$ ) is 608nm and 656nm; the rapidly declining curve of 400nm is resulted from the Blue-Latex particles loss caused by scattering. The absorption spectrum of the Blue-Latex solution was shown in Fig. 15. As a result, the peak emission wavelength ( $\lambda_p$ ) of LED has to be matched with the peak absorption wavelength  $\lambda_p$  of the sample. The procedures of the experiment are stated below: (1) Reagent: Blue-Latex solution is diluted by RO distilled water. (2) Using 500 micro liter ( $\mu$ l) micro container to load it, the effective optical path length of this solution is 5 mm. (3) the peak absorption wavelength of the Blue-Latex solution is located on the peak emission wavelength of LED. As in Fig. 16(a), the result in this experiment reveals: if the concentration is above 10% or under 0.1%, there is no detected signal. The situation is shown as Fig. 16(b) shows that the linear range is only from 0.1% to 1% of concentration. From this experiment, we can learn the following: sensitivity of this device is not only related to the design, but also related to (1) the quality of optical alignment of double optical path unit, and (2) the parallel and flatness of the front and back of the micro container. From the results of the experiment, we see that the system achieved a sensitivity of 0.1%, which is a five fold increase over a single optical path system with sensitivity of 0.5%. The overall cost of the system is also significantly cheaper than a normal optical system that uses a spectrometer. A normal spectrometer system has a

wide variety of uses and consists of complex and costly circuitry to support the diversity of the system. However, by designing the system that is more focused on the particular task of absorption detection, resulting in elimination of a microprocessor system, precision grating system, and the need of complex software, the cost of the system can be reduced by approximately 95%. The reduction in the complexity and the reduction of the total amount of component in the system, the system has the additional benefit of reduced system size to that of one-fifth of a normal system. From now on, by changing its LED, we can detect different sample solutions without redesigning another detection circuit. The circuit of the double optical path absorbance biometric system is shown in Fig. 17 and the prototype of the double optical path absorbance biometric system, shown in Fig. 18, consists of double optical path unit, electronic unit, fiber and optic mount. The prototype of the single optical path absorbance biometric system is shown in Fig. 19, and the size of which is bigger than the double optical path absorbance biometric system because the optical path of former is more distant than the latter's.

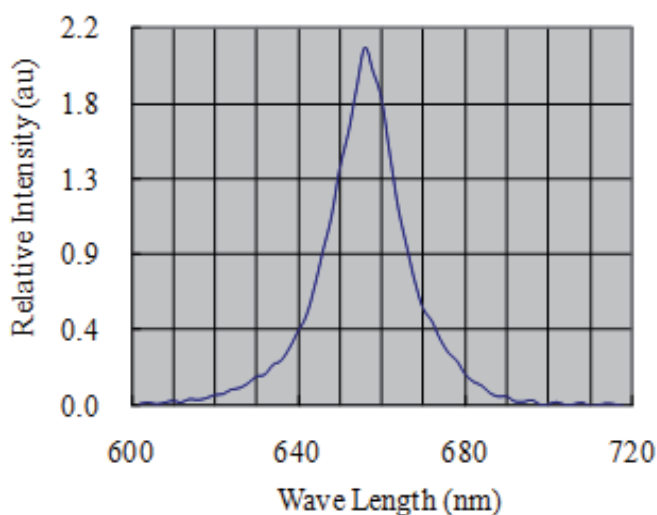


Fig. 14. Red LED emission spectrum.

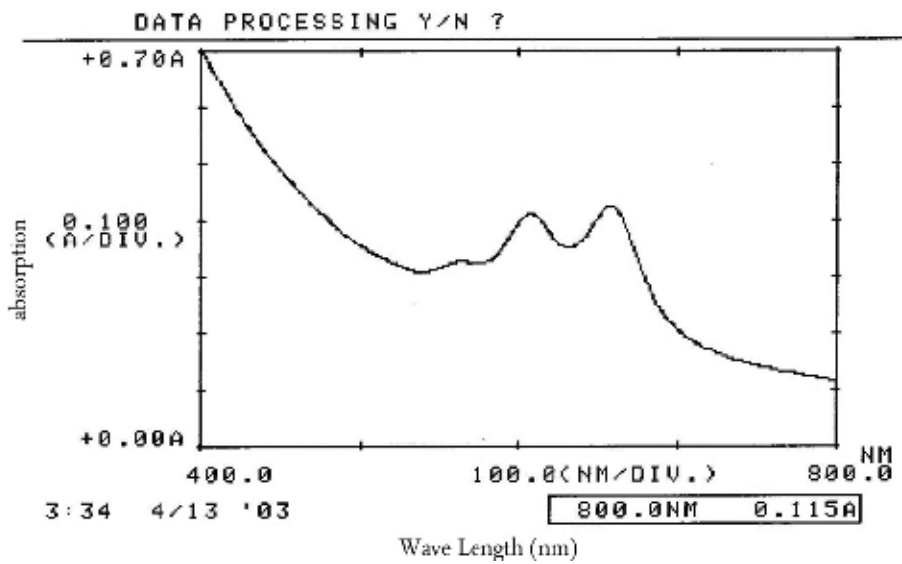


Fig. 15. Blue-Latex absorption spectrum.

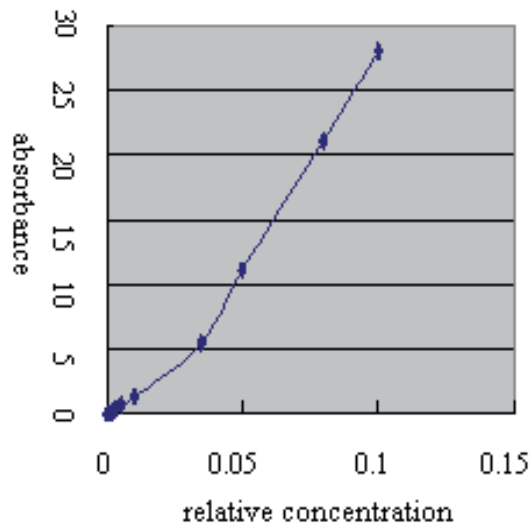


Fig. 16(a). Absorbance vs. relative concentration.

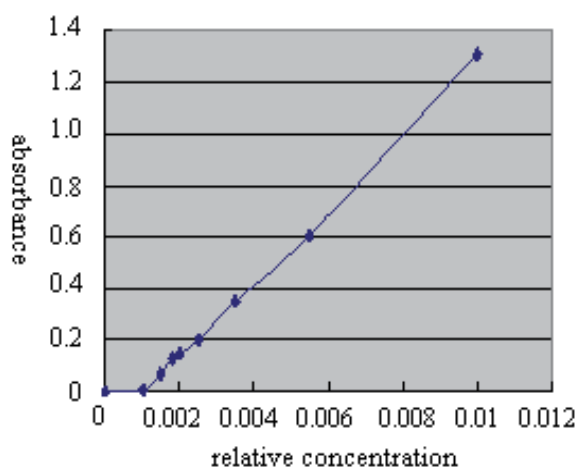


Fig. 16(b). Absorbance vs. relative concentration.

Experiment 2: the multi optical path absorbance biometric system

After finishing the study of the double optical path biometric system, we continued the study of multi optical path biometric system for improving the detective sensitivity. Those circuits are all the same. The design parameters of multi optical path biometric system are not yet optimized at present, so the performance is not good. The design concept of multi optical path biometric system has described as follow.

The multi-optical path length absorbance biometric system designed in the way that the light emitted from light source goes through a specific path at least two times can detect the reduction amount of collimated beam of visible light passing through the label to know the absorbance of the sample. The object detected by this system is the label produces after biochemical reaction. After measuring absorbance of the label we can obtain the concentration of the sample. The objective for design of multi-optical path length absorbance is to increase absorption opportunities of the collimated beam and raise detective sensitivity of this system.

The system comprise of a transmitter and receiver, a sample fluid and a multi-optical path unit as shown Fig.19. A transmitter includes a light source and driver. The receiver includes a detector, O/P amplifier, A/D converter, and processor. A multi- optical path unit includes back plane mirror, beam splitter, fold mirror, and output mirror.

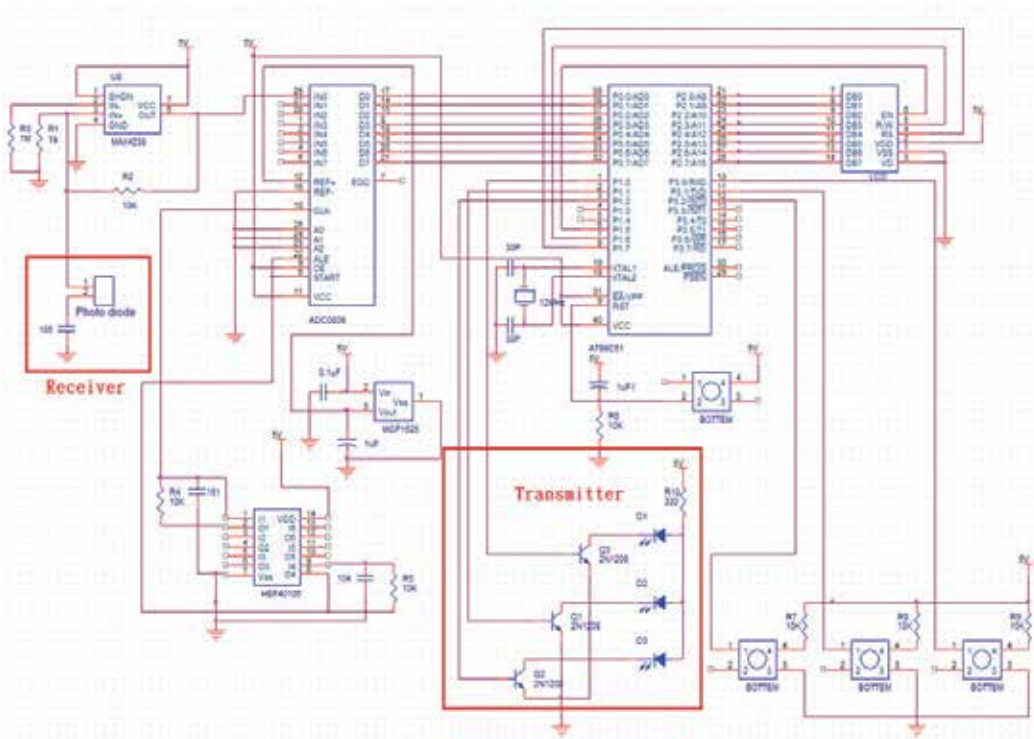


Fig. 17. The circuit of the double optical path absorbance biometric system.

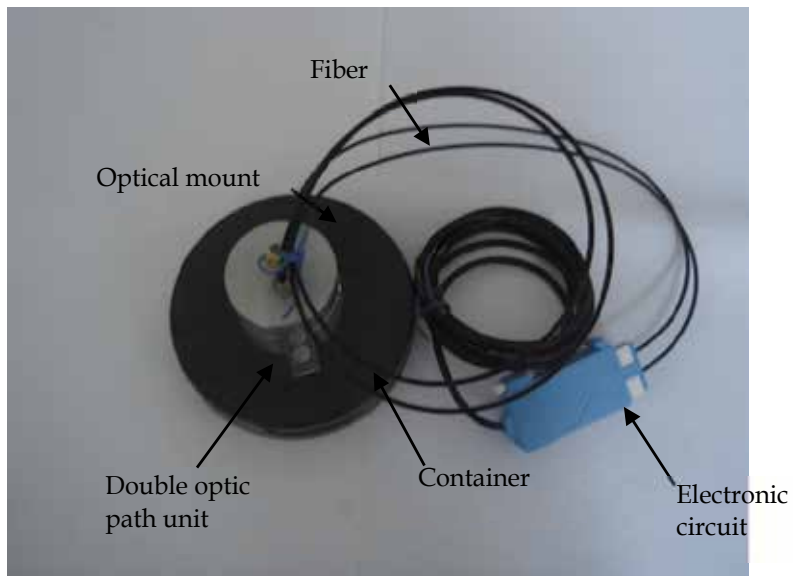


Fig. 18. Prototype of the double optical path absorbance biometric system.

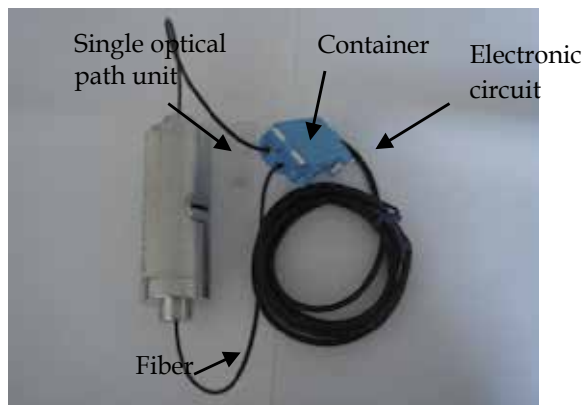


Fig. 19. Prototype of the single optical path absorbance biometric system.

The multi-optical path increases absorption opportunities of the collimated beam by sample. Comparing with one- or two-light path, the multi-optical path has highest detective sensitivity. Thus, it has the detective capability of the low concentration of the sample. This system has three advantages: (a) the multi-light path absorbance has low cost in terms of passive detective amplification, (b) the beam splitter and output mirror has combinations of different R and T, which derive different detective sensitivity, (c) it need less the sample. Following you could find the light path within multi-optical path device as solid and dashed line trip in Fig. 20. The solid line trip is defined the collimated beam passing through beam splitter, fold mirror, sample, output mirror and then output to the detector of receiver.

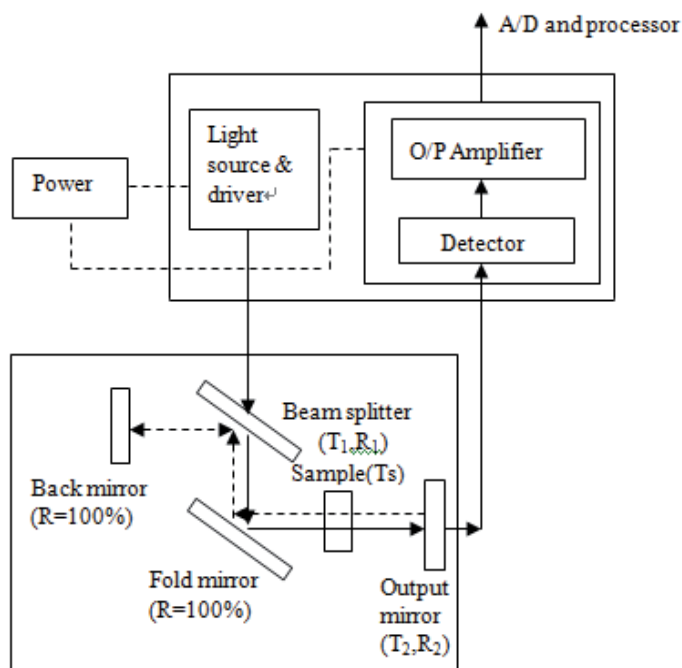


Fig. 20. The multi-optical path absorbance biometric system.

The reflected beam will be back from output mirror and goes through sample, fold mirror, beam splitter and back mirror as a dashed line trip. The beam will undergo round trips and then the intensity of the beam will be reduced due to absorption of round trips as shown in Fig.20. The transmission of multi-optical path simulation is shown at Table 2. The simulation parameters  $T_1$ ,  $T_2$ ,  $R_1$ ,  $R_2$ , and  $T_s$  are 50%, 30%, 50%, 70% and 70% respectively. The transmission of multi-optical path is 0.012. The transmission of one trip path is 0.105. From simulation results, absorbance is significantly increasing.

|          | Solid line  | Dash line   |
|----------|---|---|
| 1.       | $(T_1 T_s)(R_2 T_s R_1)^0$                          | $T_1 T_s T_2^0$   |
| 2.       | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s R_2 T_s)$         | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s T_2^1)$                               |
| 3.       | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s R_2 T_s)^2$       | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s)(R_2 T_s R_1)(R_1 T_s T_2)$           |
| n.       | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s R_2 T_s)^{(n-1)}$ | $(T_1 T_s T_2)(R_1^2 T_s^2 R_2)^n$                                      |
|          |   | $T_1 T_s T_2 +$   |
|          |   | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s T_2) +$                               |
|          |   | $[(T_1 T_s)(R_2 T_s R_1)](R_1 T_s)(R_2 T_s R_1)(R_1 T_s T_2) + \dots +$ |
|          |   | $(T_1 T_s T_2)(R_1^2 T_s^2 R_2)^{n-1}$                                  |
| $\Sigma$ |   | $[(T_1 T_s T_2)/(1 - R_1 R_2 T_s^2 R_2)]$                               |

Table 2. The transmission of multi-optical path simulation.

### Experiment 3: the multi-RGB color sensor measurement system

Unlike previous two experiments detecting the sample of single wavelength, the multi-RGB color sensor measurement system detects the sample of multi-wavelength. We study the design of the multi-RGB color sensor absorbance biometric system for the visible range absorbance biometric detection. The color samples were scarcely obtained and the color ball is to substitute it. Color can be regarded as an intrinsic physical property of an object or as a visual sensation. As a sensation, it results from three different types of the receptor cells in the retina each of which responds to a different portion of the visual spectrum. As a physical property, color is determined by the wavelength distribution of the transmitted or reflected light. Given any three different colored sources of light, it is possible to mix them in proportions that will match any sample. Thus, a color transmission of sample can be specified in terms of  $X_R$ ,  $Y_G$ , and  $Z_B$ , which are the amounts of the three primaries colors required to match the sample. For convenience, the ratio of the output voltage is commonly expressed in percent, as follows:

$$X_R = R / (R + G + B)$$



$$Y_G = G / (R + G + B)$$

$$Z_B = B / (R + G + B) \tag{4}$$

where R,G, and B is the red, green, and blue light output voltage in RGB color photodiode sensor. This system uses seven RGB color sensors to detect the color sample and recognize the sample with data comparison technique. The white LED emits the visible light which passes through the color sample and detect by the RGB color sensor. The emission's spectrum of the LED is matched with the peak absorption's wavelength of the RGB color sensor. The advantages for using the LED are the small size, short response time, long lifetime, and stable performance. The three sub pixel sensor of the seven RGB color sensor has a total of 21 analog output channels. To manage the massive amounts of analog signal, the analog Multiplexers/De-multiplexer is connected to the output of the RGB color sensor. An operation amplifier is connected to the output of the analog Multiplexers/De-multiplexer to raise the small output signal to about tenfold, raising the contrast of this system. Before the signal enters microprocessor, the analog signals are converted into the digital signal by the A/D converter. The software in the microprocessor plans and arranges the digital signal, as well as compares the data for color recognition. The data base of the microprocessor has  $X_R$ ,  $Y_G$ , and  $Z_B$  of seven balls at the different environment conditions. When the microprocessor receives the interrupt command of the push bottom, the white LED emits visible light. The voice IC enable by the subprogram, produces a series of sounds corresponding to the color sample. The system in Fig. 21 is composed of the followings: (a) the seven RGB sensors show in Fig. 26, are group according to Reds, Greens and Blues. (b) The Analog Multiplexers/Demultiplexer and Operational amplifier, (c) the A/D converters and 2-input multiplexer (d) 8051 single chip microcontroller, single chip voice record and playback device and seven push button pieces(S1....S7).

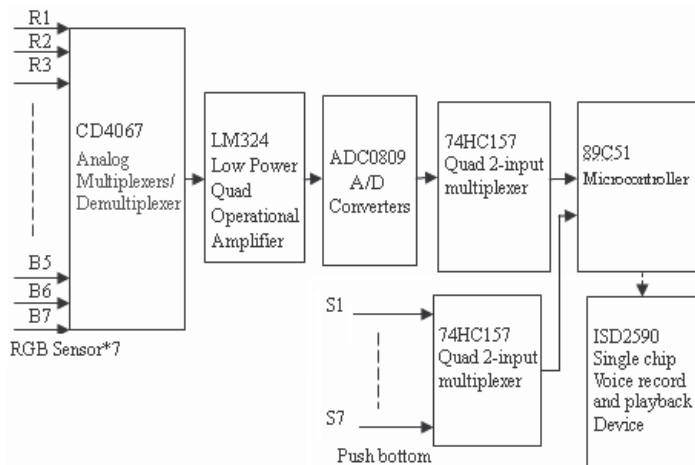


Fig. 21. Block diagram of system.

A photodiode has two terminals, a cathode and an anode. It has a low forward resistance (anode positive) and high reverse resistance (anode negative). Normal biased operation of most photodiodes calls for negative biasing. The active area of the device is the anode or positive biasing. The backside of the device is the cathode. Fig. 22 shows method of

measuring light by measuring the photocurrent. Fig. 22 shows a basic circuit connection of an operational amplifier and photodiode. The output voltage  $V_{out}$  from DC through the low-frequency region is 180 degrees out of phase with the input current. The feedback resistance  $R_f$  is determined by input current and the required output voltage  $V_{out}$ . If, however,  $R_f$  is made greater than the photodiode internal resistance  $R_{sh}$ , the operational amplifier's input noise voltage input equivalent noise voltage and offset voltage will be multiplied by  $\left(1 + \frac{R_f}{R_{sh}}\right)$ . This is superimposed on the output voltage  $V_{out}$ , and the operational amplifier's

bias current error will also increase. It is therefore not practical to use an infinitely large  $R_f$ . If there is an input capacitance  $C_t$ , the feedback capacitance  $C_f$  prevents high-frequency oscillations and also forms a low pass filter with a time constant  $C_f \times R_f$  value.

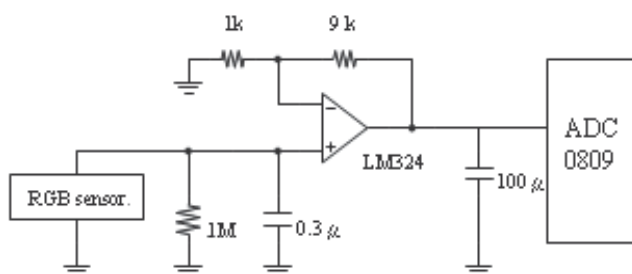


Fig. 22. Photodiode operation circuit.

The optical measurements software is capable of measuring chromaticity and luminance. The flow chart of software program is stated below:

1. Initiate 89C51 microprocessor and execute main program (as shown Fig. 23).
  - Step.1: check to see if button is push.  
 $R3=0$ , the INTT0 interrupt button not push, Re-check register.  
 $R3=1$ , the INTT0 interrupt button is push, Processor to next step.
  - Step.2: Read the digital signal from the Analog-to-Digital Converters into microprocessor memory.
  - Step.3: Calculate  $X_R$ ,  $Y_G$ , and  $Z_B$  for each sensor and compare it with  $X_R$ ,  $Y_G$ , and  $Z_B$  Data Table.
  - Step.4: After calculation, the Processor sends signal to the 74HC157 to switch the input signal from ADC0809 to the button.
2. When INTT1 interrupt button is push (see Fig. 24):
  - Step.1: Read the signal from the push button.
  - Step.2: Determine the specific button of the signal by matching the signal with the data in memory.
  - Step.3: ISD2590 produces the sounds that match the sensor reading for 3 second.

The conditions and specifications used for the experiment are listed in Table 3. This experiment use seven RGB sensors, seven high power white LED, seven difference color plastic balls, one polishing steel ball, and optoelectronic circuit board. The coloring of plastic ball is uniform. The spectral response of RGB sensors is shown Fig. 25. The hardware is shown as Fig. 26. The RGB Sensor and white LED is set in the two sides of the hole of color balls test device. The distance between them is about 40 mm.

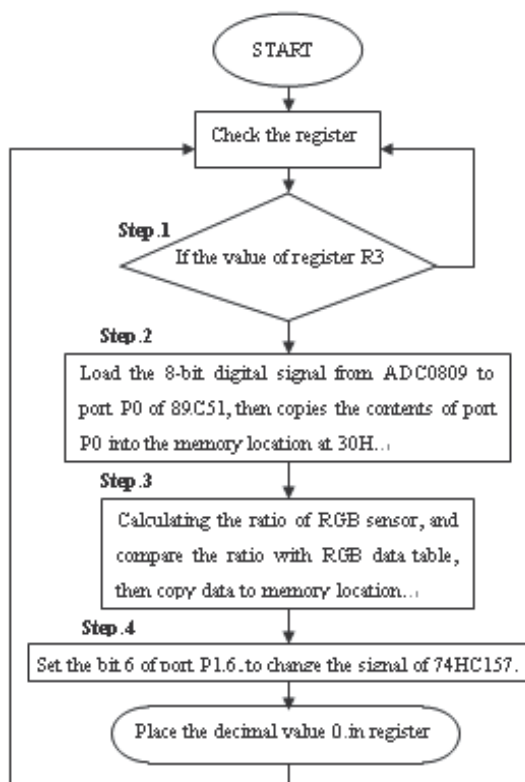


Fig. 23. Main program flow chart.

|   |  |
|---|--|
| Operating wavelength                    | 400nm~700nm  |
| RGB Sensor                              | (a) 3-channel (RGB) photodiode sensitive to the blue ( $\lambda_p=460$ nm), green ( $\lambda_p=540$ nm) and red ( $\lambda_p=620$ nm) regions of the spectrum.<br>(b) Active area: 3-segment (RGB) circular active area of $\phi 2$ mm.<br>(c) spectral response as shown Fig.5.<br>(d) Non-reflective black sleeve. |
| Plastic color ball                      | The $X_R : Y_G : Z_B$ of seven balls is showed as Table 4.<br>Some of colour balls have four $X_R : Y_G : Z_B$ due to the surface colour is not uniform. But it won't affect systems recognition on colour balls.  |
| white LED                               | (a) Emitting Color: White,<br>(b) Luminous Intensity 1100 mcd@ Forward Current 20mA, Reverse Voltage 5volt.  |
| Dynamic range                           | -0.3V~3.4V   |
| Noise                                   | Dark current~10nA  |
| Environmental intensity of illumination | Luminous Intensity $\geq 900$ mcd  |

Table 3. Experiment's conditions and specifications of this measurement system.

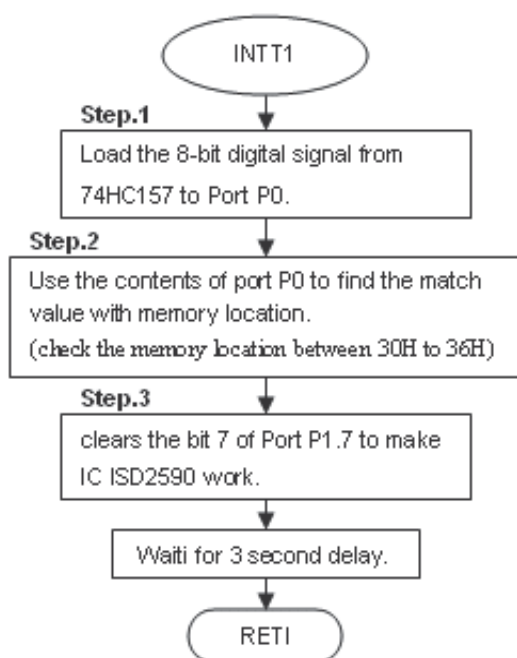


Fig. 24. Interrupts 1 subprogram flow chart.

Experiment procedure describes as follow:

1. Put color balls in the hole of color balls test device and ICE tester set up as shown in Fig. 7.
2. Built and record the look up table of  $X_R$ :  $Y_G$ :  $Z_B$  (see tab.4 of color balls at the different environmental intensity of illumination).
3. When the system power on, the detection of color balls are ready.
4. When one color ball is press down, the push bottom switch is enabling and the color ball is detected.
5. The loudspeaker will send out corresponding sounds of the color ball.
6. Repeat fifty times from (4) to (5).

According to the arrangement, there are seven kinds of color ball producing 5040 combinations in this system. The detection error of the system is less than 0.1%.

It is necessary, in general, to take into account the geometry of the source-detector system to calculate the signal incident on the detector. The usual case is that the detector only intercepts a small fraction of the radiated signal, however. Assuming that even a rough surface can be treated as a collected of plane surfaces that the incremental power incident on the portion of the detector considered ( $dA_d$ ) due to the portion of the source under consideration ( $dA_s$ ) is given by  $d\Phi = [L_s dA_s \cos\theta_s dA_d \cos\theta_d] / R^2$ . If the field of view of the detector system is greater than the solid angle subtended by the source at the detector, a large uniform background of radiance  $L_B$  contributes a radiant power of  $\Phi_b = L_B A_d (\Omega_B - \Omega_s)$ . When  $\Omega_B > \Omega_s$  then  $\Omega_B$  represents the solid angle field of view intercepting the background. The effect of field of view of the detector system on the irradiance due to an extended uniform background is mathematically similar to the calculation, except that it is carried out over a symmetrical field of view of half-angle  $\theta_{1/2}$  rather than over a full hemisphere:

$$E=L_s \int_{0}^{1/2} \int_{0}^{2\pi} \cos\theta \times \sin\theta \, d\theta \, d\phi, \quad E=\pi L_s \sin^2\theta_{1/2}$$

Consider the two cases (A and B) of a detector directly irradiated by (A) an extended source with field of view restricted by a circular stop and (B) a detector radiated by a concentrating lens that has  $\tau=1$  and no aberrations.

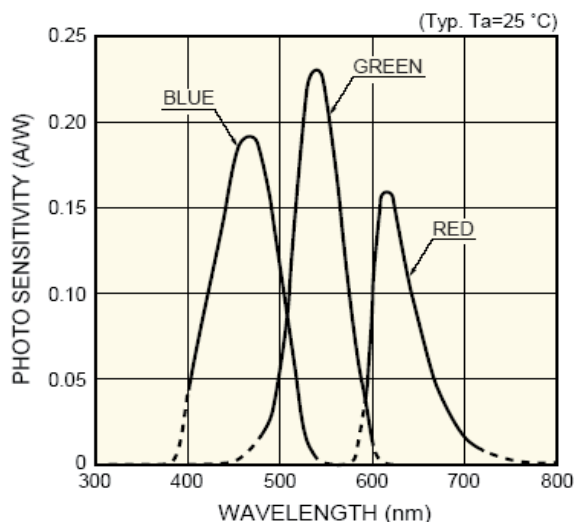


Fig. 25. The spectral response of RGB Sensor.

| Color ball | $X_R:Y_G:Z_B$ |       |       |       |
|------------|---------------|-------|-------|-------|
| Red        | 4:3:3         | 4:2:4 | 3:3:4 | 3:3:4 |
| Orange     | 5:2:2         | 4:2:3 | 4:3:3 | 4:2:3 |
| Yellow     | 4:3:2         | 3:3:2 | 4:3:2 | 3:3:2 |
| Green      | 3:3:3         | 4:3:4 | 4:4:3 | 4:4:4 |
| Blue       | 2:2:4         | 2:3:5 | 2:3:4 | 2:2:4 |
| Indigo     | 2:2:5         | 1:2:6 | 1:1:7 | 1:1:7 |
| Purple     | 3:2:4         | 3:2:4 | 3:2:4 | 3:2:4 |

Table 4. The  $X_R: Y_G: Z_B$  table of colour balls.



Fig. 26. The color balls test device.

The key features of the system have :(1) Non-contact luminance and chromaticity measurement for color recognition.(2) Memory for storing 7 channels of reference color data recognizable.(3) High-quality single-chip voice IC to increase the recognizable data output. (4) Convenient user interface that switches the mode select by a single button. (5) Noise immunity: RFI and light source. Developed with the most advanced micro-processor, data comparison technique and the technology of optoelectronic detection and circuit design, the device is capable of performing accurate stable and high speed color tests. According to the arrangement, there are seven kinds of color ball producing 5040 combinations in this system. The detection error of the system is less than 0.1%.

## 5. Conclusion

This paper designed the high-resolution visible portable biometric system to measure the specific portions of the spectrum that feature sharper and better-resolved bands—the first time to use double optical path technique. This spectrum region is key for the qualitative and quantitative analysis of the gold colloid and blue latex solution. Miniature, rugged, and reliable, portable biometric technology has immediate application in industrial quality assurance and control, particularly in the emerging field of distributed process analytical spectroscopy in the chemical, food industry and pharmaceutical industries. The portable biometric system offer high speed, signal-to-noise, and resolution comparable to a traditional spectrophotometer.

## 6. Acknowledgments

This research was supported by National Science Council of Taiwan (ROC) under contract No. NSC100-2623-E-035-001-D.

## 7. References

- [1] Reinoud F. Wolffenbuttel, "State-of-the-Art in Integrated Optical Microspectrometers", *IEEE Trans. Instrum. Meas.*, vol.53, pp.197-202,2004.
- [2] P. C. Montgomery, D. Montaner, O. Manzardo, M. Flury and H. P. Herzig, "The metrology of a miniature FT spectrometer MOEMS device using white light scanning interference microscopy", *Thin Solid Films*, vol.450, pp.79-83,2004.
- [3] Der-Chin Chen, "Portable Biometric system of High sensitivity absorption detection", *Journal of Microwave and Optical Technology letters* ,vol.50 (4), pp.868-871, 2008.
- [4] Der-Chin Chen, Chi-Chieh Hsieh, Huang-Tzung Jan and Huei-Kai Lai, "High precision multi-color sensor measurement system", *International Conference on Manufacturing and Engineering Systems*, pp.79-83,2009(MES2009)
- [5] Mark I. Montrose, "EMC and the printed circuit board : design, theory, and layout made simple" , *IEEE Press*,pp.13-14, 1999.
- [6] Y.D.Lin, C.D. Tsai, H .H. Huang, D.C. Chiou, and C .P. Wu " Preamplifier with a Second-Order High-Pass Filtering Characteristic", *IEEE Trans. Biomed. Eng.*, Vol. 46, pp. 609-612, 1999.

# Texture Analysis for Off-Line Signature Verification

Jesus F. Vargas<sup>1,2</sup> and Miguel E. Ferrer<sup>1</sup>

<sup>1</sup>*Universidad de Las Palmas de Gran Canaria - IDETIC,*

<sup>2</sup>*Universidad de Antioquia - GEPAR*

<sup>1</sup>*Spain*

<sup>2</sup>*Colombia*

## 1. Introduction

It has been suggested that verification of identity is a crucial part of the current information society. The number of situations where it is necessary a quick and inexpensive document authentication for access or information sharing, and even more for electronic commerce grows daily.

In the case of personal Verification, it is possible consider two types of biometric means: *Physiological*, which are derived from direct measurements of human body parts, and *Behavioral*, which are derived from measurements taken from an action performed by an individual to be described in an indirect manner. As examples of the former can include fingerprint, face, palm, and retina among others. In the second group may find the voice, signature, and the rhythm of typing on a computer Liu & Silverman (2001).

The handwritten signature is still one of the most commonly used and widely accepted ways for the authentication of the identity of a person. Every day thousands of documents are signed by someone in order to authorize a banking transaction, access to information or a building, a legal representation or just a contract. Despite the technology available, the vast majority of verification processes for these signatures are performed manually by human beings through visual inspection. Consequently, there is great interest in the development of automatic signature verification systems which are effective and have the ability to make quick and accurate verification of the handwritten signature of an individual. This implies that the verification of handwritten signatures (VFM) is not just a problem of pattern recognition theoretically interesting, but a real world problem with very significant commercial implications.

Handwritten signature verification is a very complex problem and scientifically attractive. Generally there are few samples to train the classification model, and also has a large intraclass variability. This represents a challenge to the scientific community. Given the importance from an economic point of view represented by this task, there is a great dependence on the effectiveness of security systems intended to prevent fraudulent access to information systems.

Currently the number of documents containing a signature as a means of identifying the person who signs is enormous, for example, in the United States extend around 17 billion

checks a year TowerGroup (Online). The verification of these documents by expert reviewers would mean astronomical costs in time and money.

Trying to increase the information available for biometric verification systems based on handwritten signatures, without this means the construction or purchase of additional capture devices, implies the possibility of developing more reliable systems with an aggregate value in constant growth. As mentioned above, in the case of off-line signature verification systems, the information available comes from a static image of the signature. This implies the non-availability of dynamic information corresponding to the signature. We can say that it is possible to characterize the information according to three approaches: a global analysis of the shape of the signature, a local analysis, or an analysis that moves away from the shape to focus primarily on the reconstruction of dynamic information.

Although, with enough training, a forger can reproduce with great skill the shape and distribution of the strokes that make up the signature, with respect to the velocity, pressure and the writing order of strokes, a forger will always find features very difficult to reproduce. This suggests that the ability to recreate and/or represent dynamic information from static images of a signature, represents a very attractive challenge for the scientific community. However, considering that the main source of information for this purpose is a grayscale image, it is necessary to develop the procedures necessary to identify the effect of perceived changes in brightness when analyzing images of signatures that have been written on different paper types (color, weight), and using different types of pen (color ink).

Also, despite the off-line systems are still disadvantaged compared to on-line systems in terms of percentages of success, still present the need for verification of identity of a person who is not present in the moment that it performs this task, as in the case of payments of checks, powers of attorney, contract signings, and any other transaction involving documentation that has been signed by someone as proof of their commitment.

This implies the need to continue working on the development of off-line systems more reliable and better suited to actual needs. The emergence of international conferences with specific focus on the analysis of documents, shows the high interest of the scientific community on this issue. Consequently, it has generated a large number of activities aimed at the automatic verification of handwritten signatures with very promising results. However, there are unexploited sources of information. In that sense, this work raises, it seeks to advance the study of the gray levels as a source of information for the characterization of a handwritten signature oriented towards his classification as genuine or fake.

This chapter is organized as follows: Section 2 describes characterization based on gray level information containing signature images. This section includes information about ink analysis and statistical texture analysis. Texture analysis is also described for the transformed domain when Wavelet transform is used. Section 3 presents the procedure proposed here for feature extraction. It describes how block analysis is used for image characterization. Datasets used for experiments are also presented in this section. Section 4 presents the experiments and results obtained here, and finally, the conclusions and remarks.

## **2. Characterization based on graylevel information**

A neglected aspect in the off-line signature verification is the use of the information in the gray levels of the signature, although a number of studies have been carried out in search of a off-line signature verification system, results have been few satisfactory for the case of those systems based on information from the gray levels. While this weakness has been offset by the combination of parameters of different nature, the information contained in the gray level image of a signature is still an untapped potential. The main difficulty encountered by



researchers has been the influence of the type of ink on the distribution of gray levels, this has meant that the developed systems do not perform the verification of a signature, but the classification of the type of ink used for realization.

It is assumed that the signature depends on the neuromotor apparatus of the person whose development is unique and determines his writing, both in the way of writing the strokes as the manner to handle the pen, and the latter is manifested on how is deposited ink on paper. Depending on the type of ink, a masking is performed for graylevel histogram of the image, is then necessary to develop an analysis that derives the information obviating the masking caused by the ink. In this way you can characterize the distribution patterns of each signer enabling the signature verification stage. In that sense, we propose that the signer information is less masked in the relationship between the gray levels of pixels of the strokes of a handwritten signature than in its absolute value.

## 2.1 Ink-type analysis

There are many methods for the reconstruction of dynamic information from the analysis of ink in handwritten samples. In the field of forensic document analysis, these methodologies are based mainly on microscopic inspection of the strokes and on assumptions about the writing process. This has resulted in the development of ink distribution model proposed by Franke Franke & Grube (1998). This first model, permitted its adaptation to specific properties of the pen used (solid ink, liquid or viscous). Its weakness was that this will be done manually by selecting the type of ink distribution according to the type of pen. In Franke et al. (2002), the authors propose an evolution to the model. It automatically determines the type of pen used, by analyzing the static strokes. This work opens the possibility to develop procedures that minimize the effect of using different types of pen on automatic verification systems.

### 2.1.1 Pen types and their properties

There are several types of pen. From a technical standpoint, the pens can be classified according to the mechanical principle, the type of tip and the ink used Franke et al. (2002). Maybe it is the type of ink used that have the greatest impact on the outcome of a manuscript stroke. Degree of liquidity of the ink significantly determines the final visual appearance of a stroke made on paper. For example, the ink used on a pencil is graphite and therefore we say it is *solid*. Pen ink is a *viscous* paste made from resin, glycerine and other additives. The *fluid* ink of the other writing instruments, is mainly composed of water that is added to the color pigments. Depending on the additives used, change the liquidity of the ink. We note that the ink of a roller-ball pen is more liquid than a Gel-ink pen. Figure 1 shows strokes made with the three types of ink mentioned.

### 2.1.2 Forensic analysis of the type of ink

As mentioned, the ink is composed mainly of water to be mixed with various additives in order to generate the different types of ink that are usually found in the marketplace. Within these additives used are oils, solvents and resins, whose composition affects the characteristics of flow and drying of the ink. Additionally, other substances such as dryers, plasticizers, waxes, fats, soaps and detergents are also used to finely alter the characteristics of the ink.

The large number of elements used in different types of ink, and possible contamination of the writing surface, become a complex problem for forensic specialists. The goal of most of the analysis is to determine if two pieces of written text have been made with the same ink. The techniques used for this purpose can be divided into non-destructive methodologies and destructive. Although it is preferred in non-destructive, the number of such techniques is limited available Thanasoulis et al. (2003).

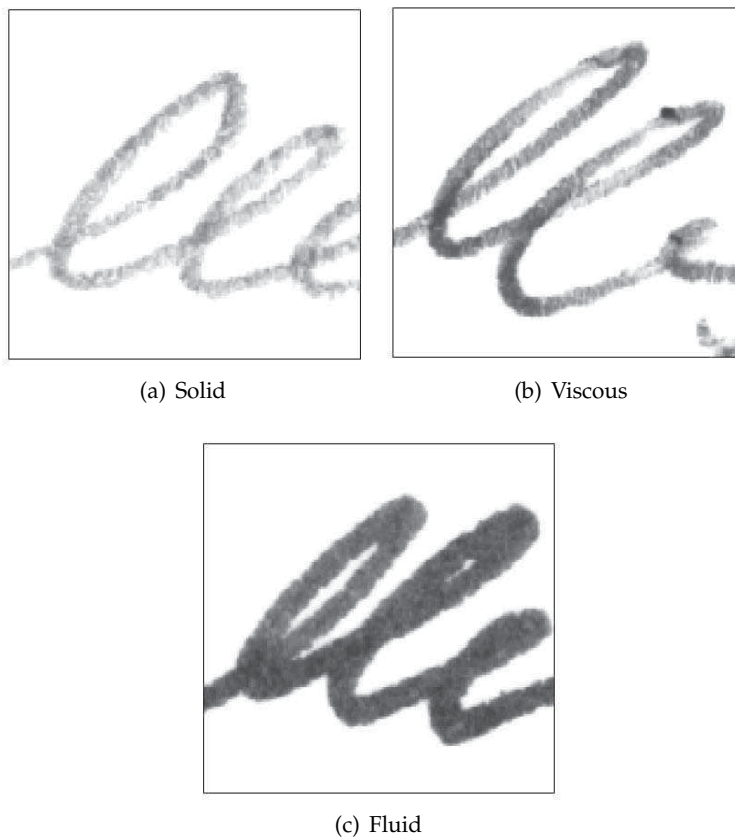


Fig. 1. Samples made with different kind of ink Franke et al. (2002).

Chromatographic separation of the ink in its components has proved to be a highly productive, since it not only allows the comparison of inks but also to do match on a database of chromatograms. Thin layer chromatography (TLC) is widely used for its speed, low cost and minimal destruction of the documents reviewed. The chromatograms can be scanned using a densitometer, but unfortunately the signal to noise ratio is low due to the large size of the area scanned.

### 2.1.3 Type of ink analysis using DSP

Use the techniques of digital image processing in forensic document analysis is relatively new Ellen (1997). The image processing offers significant cost benefits by eliminating or at least minimizes the need for expensive instrumentation and the use of a destructive methodologies. More recent works done by forensic experts show that these specialists have begun using common software packages for the analysis of documents. Most of the techniques used for this purpose are the basic procedures of digital image processing, for example some forms of contrast enhancement Bhagvati & Haritha (2005). In this sense, there are many benefits to the expert in documents analysis from the image processing community.

Similarly, the forensic field allows us to define and develop a new line of research to be exploited by experts in image processing. An example is the work presented by Bhagvati and others Bhagvati & Haritha (2005), where it is studied the problem of identifying different

types of ink techniques using digital image processing. According to the conclusions of the authors, the method presented has a better performance than human perception in the case of printed inks. Used HSV space information, in particular saturation levels, to determine variations in patterns of absorption of the paper according to the ink. Figure 2 shows the estimated models.

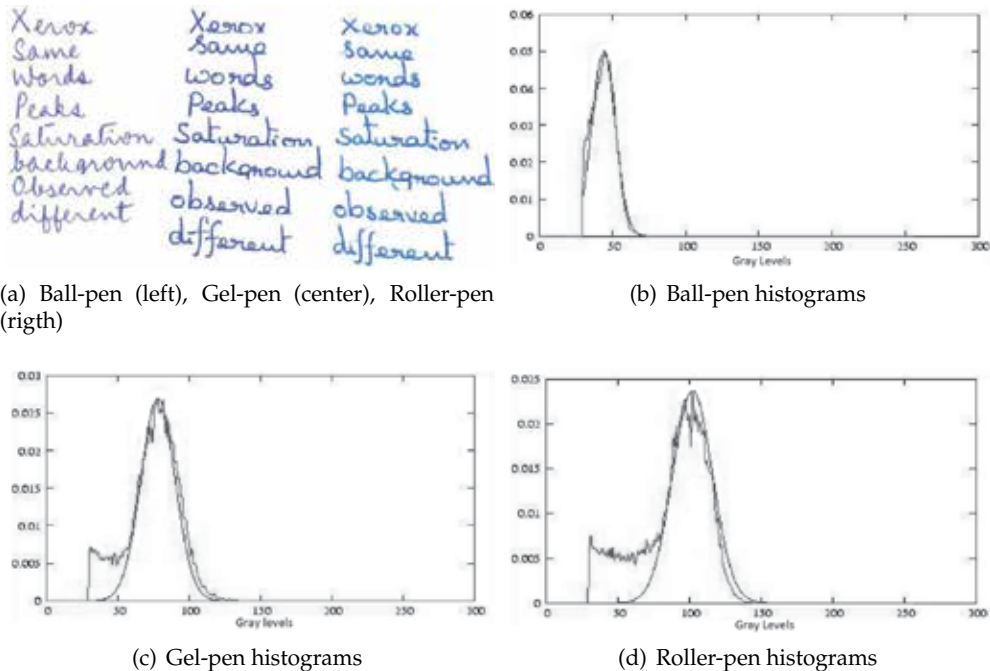


Fig. 2. Saturation histograms and Gaussian Model estimated for different types of ink, Bhagvati & Haritha (2005).

For the specific case of handwriting analysis, Frank and others Franke et al. (2002) presented a paper focusing on developing a methodology to automatically determine the type of pen used by the analysis of the strokes in a static image. In that study, there was more attention to the characteristics that describe the visual appearance of the ink distribution along the handwritten lines. Given that the shape of the lines is dependent on each writer and provides little information about the type of ink, this feature was not taken into account by the authors. As shown in Figure 1, texture in images of handwritten strokes is determined by the physical properties of the ink used. Therefore it is suggested that is possible to determine the type of ink used from an analysis of textures in the image. To this end, the authors use a classical approach in the area of texture analysis, the Gray Level Co-occurrence Matrix (GLCM), or array of spatial dependence of gray levels present in an image. To represent each texture, were used second-order statistical features calculated from GLCM, which allows independence from the lighting. Figure 3 presents the GLCM matrices obtained for the three types of ink shown in Figure 1. The results obtained by the authors in the classification of these three types of ink showing an error less than 1%. In another work, Frank and Rose Franke & Rose (2004) describe their study of the influence of physical and biomechanical processes on the ink strokes in order to provide a solid foundation that allows the improvement of signature analysis systems. Using a robot writer, able to handle different types of writing tools under

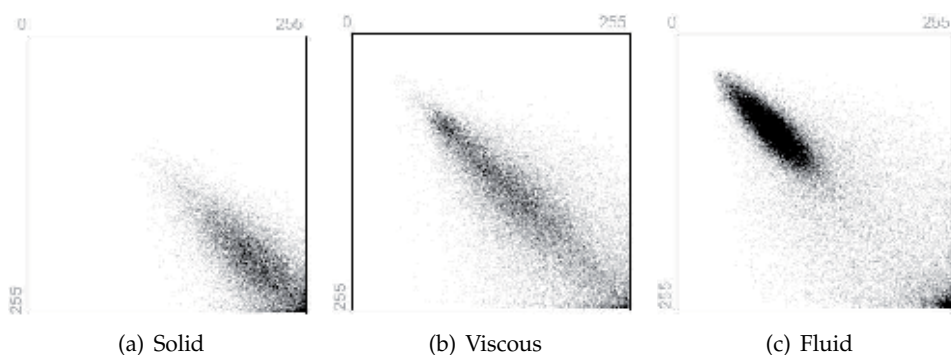


Fig. 3. GLCM with 255 graylevels for different type of ink Franke et al. (2002).

controlled conditions, they simulate the motions of writing in order to study the relationship between the characteristics of the writing process and the deposition of ink on paper. As a result of the analysis of these artificial strokes, corresponding to the use of 30 different pens, the authors proposed an Ink Deposition Model (IDM). This model describes analytically the relationship between the force applied to the pen and the relative intensity distribution of three types of ink: Solid, Viscous and fluid.

Figure 4 shows histograms for the three types of ink taken into account in the study, when applying different values of force on the pen. It may see a shift to the left (darker graylevels) when increasing the force exerted on the pen. The authors also mention that changes in the color of the ink only represent a shift of the histogram without changes in the distribution.

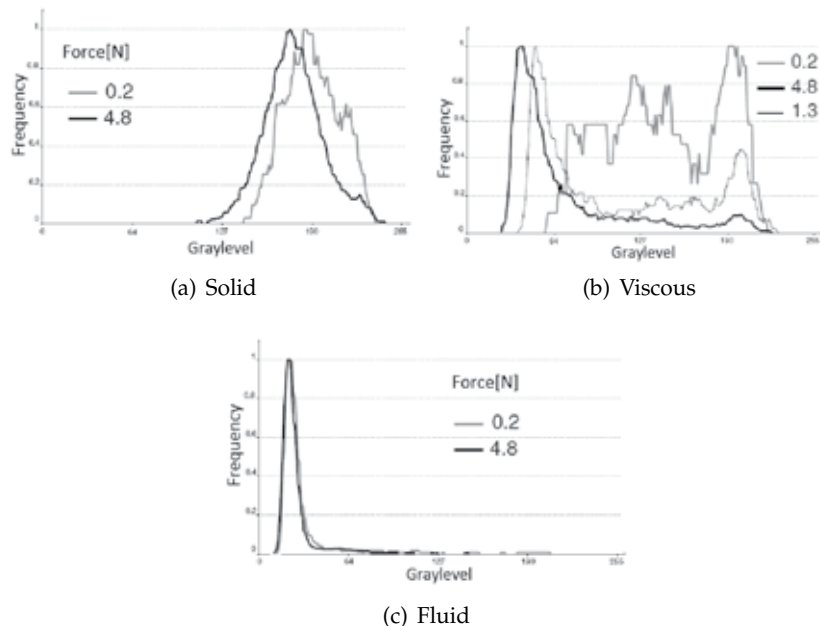


Fig. 4. Histograms applying different force values on pen Franke & Rose (2004).

## 2.2 Statistical texture analysis

Biometric systems based on signature verification, in conjunction with textural analysis, can reveal information about ink-pixels distribution which reflects personal characteristics from the signer i.e. pen-holding, writing speed and pressure. But we do not think that only ink distribution information is sufficient for signer identification. So, in the specific case of signature strokes, we have also taken into account, for the textural analysis, the pixels in the stroke contour. By this we mean those stroke pixels that are in the signature-background border. These pixels will include statistical information about the signature shape. So this distribution data may be considered as a combination of textural and shape information.

### 2.2.1 Grey Level Co-occurrence Matrices

The Grey Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features from the image Connors & Harlow (1980). This approach has been used in a number of applications, including ink type analysis, e.g. Franke et al. (2002); Haralick (1979); He et al. (1987); Trivedi et al. (1984).

A GLCM of an image  $I(x,y)$  is a matrix  $P(i,j|\Delta_x,\Delta_y), 0 \leq i \leq G-1, 0 \leq j \leq G-1$ , where the number of rows and columns is equal to the number of grey levels  $G$ . The matrix element  $P(i,j|\Delta_x,\Delta_y)$  is the relative frequency with which two pixels with grey levels  $i$  and  $j$  occurs separated by a pixel distance  $\Delta_x,\Delta_y$ . For simplicity, in the rest of the paper, we will denote the GLCM matrix as  $P(i,j)$ .

For a statistically reliable estimation of the relative frequency we need a sufficiently large number of occurrences for each event. The reliability of  $P(i,j)$  depends on the grey level number  $G$  and the  $I(x,y)$  image size. In the case of images containing signatures, instead of image size, this depends on the number of the pixels in the signature strokes. If the statistical reliability is not sufficient, we need to reduce  $G$  to guarantee a minimum number of pixels transitions per  $P(i,j)$  matrix component, despite losing texture description accuracy. The grey level number  $G$  can be reduced easily by quantifying the image  $I(x,y)$ .

The classical feature measures extracted from the GLCM matrix (see Haralick et al. Haralick (1979) and Connors et al. Connors & Harlow (1980)) are the following:

- Texture homogeneity H:

$$H = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{P(i,j)\}^2 \quad (1)$$

A homogeneous scene will contain only a few grey levels, giving a GLCM with only a few but relatively high values of  $P(i,j)$ . Thus, the sum of squares will be high.

- Texture contrast C:

$$C = \sum_{n=0}^{G-1} \left\{ n^2 \cdot \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \right\}, |i-j| = n \quad (2)$$

This measure of local intensity variation will favour contributions from  $P(i,j)$  away from the diagonal, i.e  $i \neq j$ .

- Texture entropy E:

$$E = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \cdot \log \{P(i,j)\} \quad (3)$$

Non-homogeneous scenes have low first order entropy, while a homogeneous scene reveals high entropy.

- Texture correlation O:

$$O = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{i \cdot j \cdot P(i,j) - (m_i \cdot m_j)}{\sigma_i \cdot \sigma_j} \tag{4}$$

where  $m_i$  and  $\sigma_i$  are the mean and standard deviation of  $P(i,j)$  rows, and  $m_j$  and  $\sigma_j$  the mean and standard deviation of  $P(i,j)$  columns respectively. Correlation is a measure of grey level linear dependence between pixels at the specified positions relative to each other.

### 2.2.2 Local Binary Patterns

The Local Binary Pattern (LBP) operator is defined as a grey level invariant texture measure, derived from a general definition of texture in a local neighbourhood, the centre of which is the pixel  $(x,y)$ . Recent extensions of the LBP operator have shown it to be a really powerful measure of image texture, producing excellent results in many empirical studies. LBP has been applied in biometrics to the specific problem of face recognition Marcel et al. (2007); Nikam & Agarwal (2008). The LBP operator can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its invariance to monotonic grey level changes. Equally important is its computational simplicity, which makes it possible to analyse images in challenging real-time settings T. (2003). The local binary pattern operator describes the surroundings of the pixel  $(x,y)$  by generating a bit-code from the binary derivatives of a pixel as a complementary measure for local image contrast. The original LBP operator takes the eight neighbouring pixels using the centre grey level value  $I(x,y)$  as a threshold. The operator generates a binary code 1 if the neighbour is greater than or equal to the central level, otherwise it generates a binary code 0. The eight neighbouring binary codes can be represented by an 8-bit number. The LBP operator outputs for all the pixels in the image can be accumulated to form a histogram which represents a measure of the image texture. Figure 5 shows an example of LBP operator.

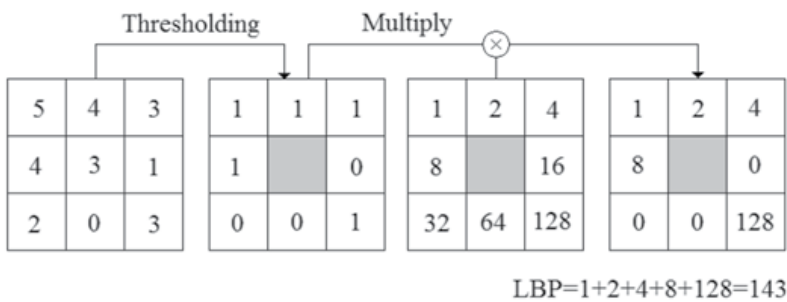


Fig. 5. Working out the LBP code of pixel  $(x,y)$ . In this case  $I(x,y) = 3$ , and its LBP code is  $LBP(x,y)=143$ .

The above LBP operator is extended in Ojala et al. (2002) to a generalised grey level and rotation invariant operator. The generalised LBP operator is derived on the basis of a circularly symmetric neighbour set of  $P$  members on a circle of radius  $R$ . The parameter

$P$  controls the quantisation of the angular space and  $R$  determines the spatial resolution of the operator. The LBP code of central pixel  $(x, y)$  with  $P$  neighbours and radius  $R$  is defined as:

$$LBP_{P,R}(x, y) = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p \quad (5)$$

where  $s(l) = \begin{cases} 1 & l \geq 0 \\ 0 & l < 0 \end{cases}$ , the unit step function,  $g_c$  the grey level value of the central pixel:  $g_c = I(x, y)$  and  $g_p$  the grey level of the  $p^{th}$  neighbour, defined as:

$$g_p = I \left( x + R \cdot \sin \frac{2\pi p}{P}, y - R \cdot \cos \frac{2\pi p}{P} \right) \quad (6)$$

If the  $p^{th}$  neighbour does not fall exactly in the pixel position, its grey level is estimated by interpolation. An example can be seen in Figure 6

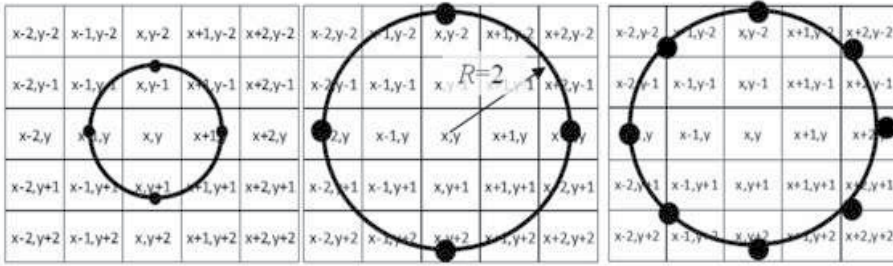


Fig. 6. The surroundings of  $I(x, y)$  central pixel are displayed along with the  $p^{th}$  neighbours, marked with black circles, for different  $P$  and  $R$  values. Left:  $P = 4, R = 1$ , the  $LBP_{4,1}(x, y)$  code is obtained by comparing  $g_c = I(x, y)$  with  $g_{p=0} = I(x, y - 1)$ ,  $g_{p=1} = I(x + 1, y)$ ,  $g_{p=2} = I(x, y + 1)$  and  $g_{p=3} = I(x - 1, y)$ . Centre:  $P = 4, R = 2$ , the  $LBP_{4,2}(x, y)$  code is obtained by comparing  $g_c = I(x, y)$  with  $g_{p=0} = I(x, y - 2)$ ,  $g_{p=1} = I(x + 2, y)$ ,  $g_{p=2} = I(x, y + 2)$  and  $g_{p=3} = I(x - 2, y)$ . Right:  $P = 8, R = 2$ , the  $LBP_{8,2}(x, y)$  code is obtained by comparing  $g_c = I(x, y)$  with  $g_{p=0} = I(x, y - 2)$ ,  $g_{p=1} = I(x + \sqrt{2}, y - \sqrt{2})$ ,  $g_{p=2} = I(x + 2, y)$ ,  $g_{p=3} = I(x + \sqrt{2}, y + \sqrt{2})$ ,  $g_{p=4} = I(x, y + 2)$ ,  $g_{p=5} = I(x - \sqrt{2}, y + \sqrt{2})$ ,  $g_{p=6} = I(x - 2, y)$  and  $g_{p=7} = I(x - \sqrt{2}, y - \sqrt{2})$ .

In a further step, Ojala et al. (2002) defines a  $LBP_{P,R}$  operator invariant to rotation as follows:

$$LBP_{P,R}^{riu2}(x, y) = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(x, y) \leq 2 \\ P + 1 & \text{otherwise} \end{cases} \quad (7)$$

where

$$U(x, y) = \sum_{p=1}^P |s(g_p - g_c) - s(g_{p-1} - g_c)|, \text{ with } g_c = g_0 \quad (8)$$

Analysing the above equations,  $U(x, y)$  can be calculated as follows:

1. Work out the function  $f(p) = s(g_p - g_c), 0 < p < P$  considering  $g_p = g_0$ ;
2. Obtain its derivate:  $f(p) - f(p - 1), 1 \leq p \leq P$ ;

3. Calculate the absolute value:  $|f(p) - f(p - 1)|, 1 \leq p \leq P$ ;
4. Obtain  $U(x, y)$  as the integration or sum  $\sum_{p=1}^P |f(p) - f(p - 1)|$

If the grey levels of the pixel  $(x, y)$  neighbours are uniform or smooth, as in the case of Figure 7(a),  $f(p)$  will be a sequence of '0' or '1' with only two transitions. In this case  $U(x, y)$  will be zero or two and the  $LBP_{P,R}^{riu2}$  code is worked out as the sum  $\sum_{p=0}^{P-1} f(p)$ .

Conversely, if the surrounding grey levels of pixel  $(x, y)$  vary quickly, as in the case of figure 7(b),  $f(p)$  will be a sequence containing several transitions '0' to '1' or '1' to '0' and  $U(x, y)$  will be greater than 2. So, in the noisy case, a constant value equal to  $P + 1$  is assigned to  $LBP_{P,R}^{riu2}$  making it more robust to noise than previously defined LBP operators.

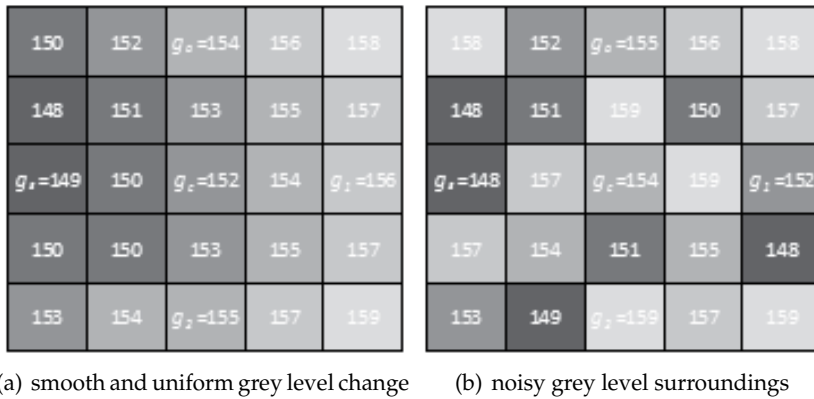


Fig. 7. Calculating the  $LBP_{P,R}^{riu2}$  code for two cases, with  $P = 4$  and  $R = 2$ : Left:  $g_c = 152$ ,  $g_0, g_1, g_2, g_3 = 154, 156, 155, 149, f(0), f(1), f(2), f(3), f(4) = 1, 1, 1, 0, 1$ ,  $U(x, y) = 0 + 0 + 1 + 1 = 2 \leq 2$ , therefore  $LBP_{P,R}^{riu2}(x, y) = 1 + 1 + 1 + 0 = 3$ . Right:  $g_c = 154$ ,  $g_0, g_1, g_2, g_3 = 155, 152, 159, 148, f(0), f(1), f(2), f(3), f(4) = 1, 0, 1, 0, 1$ ,  $U(x, y) = 1 + 1 + 1 + 1 = 4 \leq 2$ ,  $LBP_{P,R}^{riu2}(x, y) = P + 1 = 5$ .

The rotation invariance property is guaranteed because when summing the  $f(p)$  sequence to obtain the  $LBP_{P,R}^{riu2}$ , it is not weighted by  $2^p$ . As  $f(p)$  is a sequence of 0 and 1,  $0 \leq LBP_{P,R}^{riu2}(x, y) \leq P + 1$ . As textural measure, we will use its  $P + 2$  histogram bins of  $LBP_{P,R}^{riu2}(x, y)$  codes.

### 2.3 Texture analysis in transformed domain

We describe the use of Wavelet transform as a complement to features already described and with the aim of achieving a reduction in the variance of the results for different databases due to different type of ink used by the signers.

When performing the Wavelet decomposition of an image 4 matrices obtained: 1 with the coefficients of approximation and other 3 with the detail coefficients. For these 3 last, we make use of 3 wavelet which measure functional changes (changes in intensity or gray levels) in different directions.  $\psi^H$  measures variations along the columns (eg horizontal edges),  $\psi^V$  corresponds to variations along the rows (eg vertical edges) and  $\psi^D$  corresponds to diagonal variations. Each of these wavelet is the product of a unidimensional scale  $\varphi$  and



the corresponding wavelet  $\psi$ . Excluding those products yielding a one-dimensional result, as  $\varphi(x)\psi(x)$ , the four remaining products define the scaling function

$$\varphi(x, y) = \varphi(x)\varphi(y) \quad (9)$$

and three wavelet with directional sensitivity

$$\begin{aligned} \psi^H(x, y) &= \psi(x)\varphi(y) \\ \psi^V(x, y) &= \varphi(x)\psi(y) \\ \psi^D(x, y) &= \psi(x)\psi(y) \end{aligned} \quad (10)$$

wavelet *Haar* was used.

Generally, for image compression tasks, most of the relevant information is contained in the matrix of approximation coefficients. Here we propose that the features based on texture analysis, described above, should be calculated from the matrices of detail. This is based on the fact that the matrix of approximation coefficients mainly contain information corresponding to the type of ink used (low frequencies, little changed from gray level), while the detail coefficients contain information on the personal characteristics of the signer (high frequencies, many changes in levels of gray). Figure 8 shows the histograms of the original image and the matrices resulting from the Wavelet decomposition, and serve as the basis of the hypothesis. This analysis corresponds to a sample made using viscose ink, and whose original histogram is similar to the patterns presented by Franke & Rose (2004) which were presented in Figures 4.

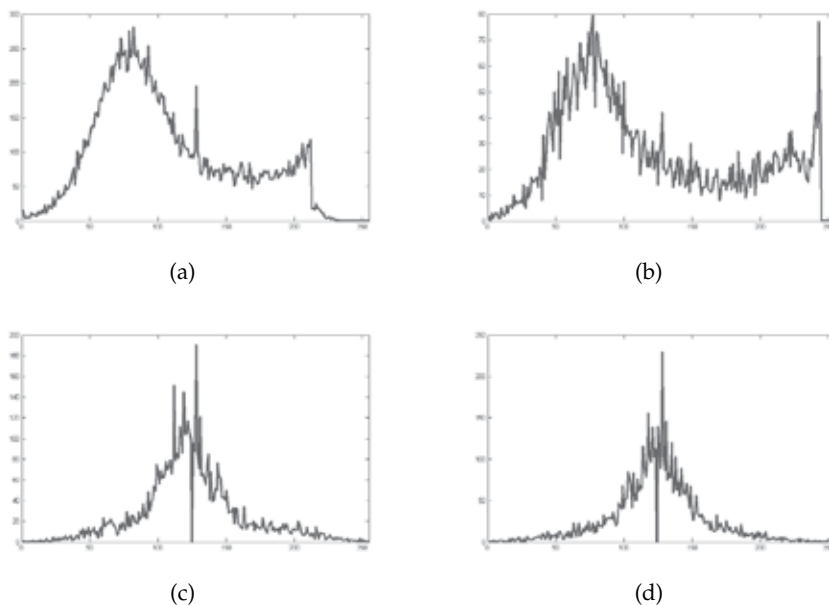


Fig. 8. Wavelet Decomposition. (a) Original image histogram. (b) Histogram of approximation coefficients. (c) Histogram of vertical detail coefficients. (d) Histogram of horizontal detail coefficients. For sake of visualization, background values has been remove from histograms.

### 3. Feature extraction

After the wavelet decomposition of the original image, one proceeds with the calculation of textural features. Each one of the three matrix of approximation coefficients is characterized using block analysis. For GLCM, characteristics were calculated: Homogeneity, Contrast, Entropy, Energy and Correlation, and a combination of them was used for verification (CHEEC). Have been used 8 graylevels for GLCM calculation and Offsets vector  $(\Delta_x \Delta_y)$  with values  $[0 \ 1, -1 \ 1, -1 \ 0, -1 \ -1]$  was used. For LBP, a combination of two LBP operators ( $R = \{1, 2\}$  y  $P = \{8, 16\}$ ) was used. These two settings allow us to analyze the image at different resolutions.

Finally, the three vectors are concatenated to form the final vector of signature features. Figure 9 shows the procedure to construct the final features vector using the vector of characteristics of the three matrices of approximation coefficients.

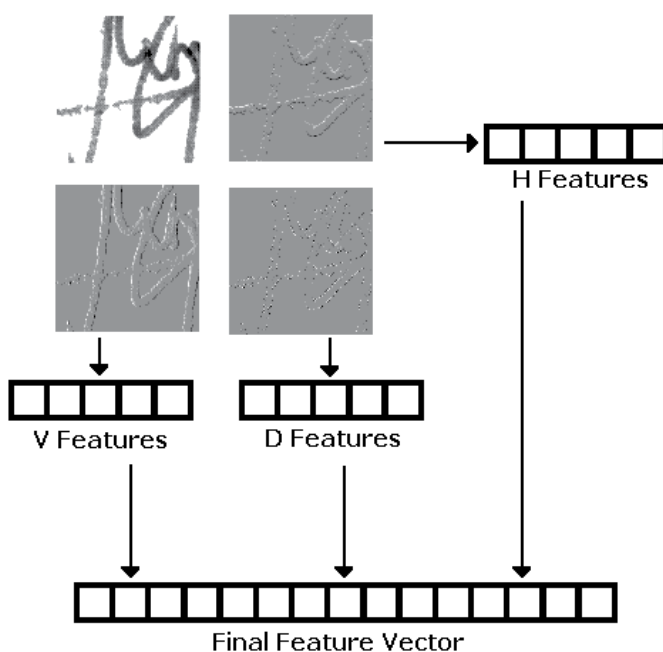


Fig. 9. Feature vector construction.

### 3.1 Datasets

#### 3.1.1 GPDS corpus

Digital Signal Processing Group (GPDS) of the Universidad de Las Palmas de Gran Canaria, has devoted efforts to create a large scale database called GPDS-850 Corpus Vargas et al. (2007). This database contains samples of 850 signers with 24 genuine samples and 24 forgeries for each. Signers used their own pen so the dataset contains samples made with different type of ink. Images have a resolution of 600dpi. This dataset was constructed on two stages. In that way, the corpus was divided in two subcorpus (GPDS100 and GPDS750) for the experiments carried out on this work.

### 3.1.2 MCYT corpus

With the encouragement of having a reference database, this work takes into account a subcorpus of the database created by the Multimodal Biometric Group - ATVS of the Autonomous University of Madrid. It is available in <http://atvs.ii.uam.es/databases.jsp>.

The corpus *MCYT-SignatureOff-75* Fierrez-Aguilar et al. (2004), contains samples of 75 signers. It has 15 genuine samples and 15 forgeries for each signer. The images have a resolution of 600dpi. It is noteworthy that all the signatures were made with the same type of pen. This database has been used in different works Fierrez-Aguilar et al. (2004), Ferrer et al. (2006), Freire et al. (2007), Alonso-Fernandez et al. (2007), Güler & Meghdadi (2008), Gilperez et al. (2008), Prakash & Guru (2009), which allows us to suggest some kind of comparison with the results obtained in this study.

## 4. Experiments and results

Once preprocessed, each sample is characterized and represented by a vector of parameters that feeds the verification stage. In this work, each signer is modeled using Least Squares-Support Vector Machines (LS-SVM).

For this work, we used only samples of the first 75 signers of each database. This in order to make a fair comparison of the results obtained with the 3 corpuses. Model training is performed with 5 and 10 genuine samples as positive samples. This was done in order to carry out an analysis of system performance respect to the number of original samples used for model building. These genuine samples are chosen randomly. Random forgeries were used as negative samples (genuine samples of other signers). The use of these samples was proposed in Bertolini et al. (2009), and for this work it was taken a genuine sample of each of the other signers of the database, ie 74. Given the small number of samples available for training was used the Leaving-one-out cross-validation procedure (LOOCV), to determine the value of the parameters ( $\gamma, C$ ) for LS-SVM classifier with RBF kernel.

For the test were used random forgeries and skilled forgeries. In the case of random forgeries, it took a genuine signature of each of the other users of the database (ensuring to be different than the sample used in training). In the case of skilled forgeries, we used all the available samples in the database, ie, 15 in the MCYT and 24 in the GPDS corpuses. To obtain more reliable results, the training and testing procedures were repeated 10 times with different sets of training and test data. As threshold for determining the values of FAR and FRR, and taking into account that the LS-SVM classifier was trained with samples labeled as '+1' for genuine samples and '-1' for forgeries, the value was set at zero for all signatories. That is, if the LS-SVM provides a value greater than *zero* at its output, the signature is accepted as genuine. If the LS-SVM gives a value less than zero as output, the signature is considered as a forgery and therefore is rejected.

### 4.1 Results

Tables 1 y 2 show results when GLCM and LBP features are used independently for verification.

Table 3 shows results using a feature level combination of GLCM and LBP features. This combination represents a system performance improvement when compared with tables 1 and 2.

The following, we present the results obtained when using the Wavelet transform as a complement to the combination of features referred above. First Wavelet decomposition is performed, then LBP and GLCM features are calculated, and finally these are concatenated to form the feature vector.

| #  | Dataset | FAR(%) | FRR(%) | EER(%) | $\sigma$ -FAR | $\sigma$ -FRR |
|----|---------|--------|--------|--------|---------------|---------------|
| 5  | MCYT    | 5,99   | 20,57  | 11,82  | 3,25          | 8,89          |
|    | GPDS100 | 10,82  | 19,33  | 14,58  | 4,68          | 6,92          |
|    | GPDS750 | 16,02  | 20,93  | 18,19  | 6,41          | 7,48          |
| 10 | MCYT    | 11,41  | 12,85  | 11,76  | 3,91          | 9,21          |
|    | GPDS100 | 18,66  | 11,86  | 16,16  | 5,54          | 7,84          |
|    | GPDS750 | 26,69  | 12,10  | 21,31  | 7,20          | 7,75          |

Table 1. Results using GLCM features

| #  | Dataset | FAR(%) | FRR(%) | EER(%) | $\sigma$ -FAR | $\sigma$ -FRR |
|----|---------|--------|--------|--------|---------------|---------------|
| 5  | MCYT    | 4,34   | 31,18  | 15,07  | 3,36          | 9,04          |
|    | GPDS100 | 5,37   | 27,81  | 15,28  | 3,61          | 9,20          |
|    | GPDS750 | 4,74   | 30,60  | 16,17  | 4,34          | 9,92          |
| 10 | MCYT    | 9,14   | 16,53  | 10,98  | 3,74          | 10,18         |
|    | GPDS100 | 10,86  | 14,25  | 12,11  | 4,17          | 6,89          |
|    | GPDS750 | 8,25   | 15,35  | 10,87  | 5,64          | 6,82          |

Table 2. Results using LBP features.

| #  | Dataset | FAR(%) | FRR(%) | EER(%) | $\sigma$ -FAR | $\sigma$ -FRR |
|----|---------|--------|--------|--------|---------------|---------------|
| 5  | MCYT    | 6,72   | 25,79  | 14,34  | 3,92          | 9,20          |
|    | GPDS100 | 6,20   | 25,98  | 14,94  | 3,73          | 7,96          |
|    | GPDS750 | 6,00   | 27,67  | 15,57  | 4,34          | 8,79          |
| 10 | MCYT    | 11,51  | 13,71  | 12,06  | 3,22          | 9,76          |
|    | GPDS100 | 11,95  | 11,52  | 11,79  | 4,06          | 5,53          |
|    | GPDS750 | 8,29   | 14,16  | 10,45  | 4,40          | 6,74          |

Table 3. Results using a combination of GLCM and LBP features (GLCM+LBP).

| #  | Dataset | FAR(%) | FRR(%) | EER(%) | $\sigma$ -FAR | $\sigma$ -FRR |
|----|---------|--------|--------|--------|---------------|---------------|
| 5  | MCYT    | 3,94   | 25,77  | 12,67  | 2,90          | 9,20          |
|    | GPDS100 | 6,30   | 18,55  | 11,71  | 3,27          | 7,45          |
|    | GPDS750 | 6,69   | 19,63  | 12,41  | 5,50          | 7,87          |
| 10 | MCYT    | 9,12   | 11,83  | 9,80   | 3,42          | 9,42          |
|    | GPDS100 | 11,25  | 6,27   | 9,42   | 3,63          | 4,40          |
|    | GPDS750 | 8,51   | 6,79   | 7,88   | 4,31          | 4,73          |

Table 4. Results using Wavelet decomposition and a combination of GLCM+LBP.

## 5. Conclusions

The use of Wavelet transform as a complement to the features based on texture analysis, has improved the system performance. A single level of decomposition is sufficient to achieve acceptable results. A multilevel Wavelet decomposition significantly increases the computational cost without providing improved results.

Regarding the combination of features, it can be concluded that the joint use of features based on texture analysis improves system performance. The combination made at the level of features, offers better results with respect to use of individual traits. Additionally, this combination of features is enhanced when using the Wavelet transform as a complement to

it. Using Wavelet transform as a previous step to the characterization of the signature using the combination of features provides the best results presented in this paper in terms of EER values and variability for different databases containing samples made with different type of ink.

In an image containing a handwritten signature, most of the pixels belongs the background. This large percentage of elements not provide information about the signature. Using the blocks analysis it is possible to perform a local analysis, allowing to detect areas where there are no traces of the signature. In this way, characterization is done mainly on the elements belonging to different strokes, improving system performance.

### 5.1 Future research

It is important to conduct a study to verify the results obtained in this study for more complex environments, such as images in which the signature is contaminated with seals, or characters, especially case the checks. It is necessary to consider whether the procedures for the signature segmentation in such images affects gray level information.

As seen, the combination of features can improve system performance. As mentioned, this combination was performed at feature level, in this way, would be highly relevant to study the performance of the system when the combination is done at score level. It is also of interest to study the combination of features based on gray levels with those based on binary information. If takes into account that so far the best results were obtained in works where it is used features of the latter type, combining features of different nature could improve the overall performance of a signature verification system.

## 6. References

- Alonso-Fernandez, F., Fairhurst, M. C., Fierrez, J. & Ortega-Garcia, J. (2007). Automatic measures for predicting performance in off-line signature, *IEEE Proceedings International Conference on Image Processing*, Vol. 1, pp. 369–372.
- Bertolini, D., Oliveira, L., Justino, E. & Sabourin, R. (2009). Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers, *Pattern Recognition* 43(1): 387–396.
- Bhagvati, C. & Haritha, D. (2005). Classification of Liquid and Viscous Inks using HSV Colour Space, *Proceedings of the Eighth International Conference on Document Analysis and Recognition*, IEEE Computer Society, Washington, DC, USA, pp. 660–664.
- Connors, R. W. & Harlow, C. A. (1980). A theoretical comparison of texture algorithms, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 2(3): 204–222.
- Ellen, D. (1997). *The scientific examination of documents : methods and techniques*, 2nd ed. edn, Taylor & Francis, London ; Bristol, PA :.
- Ferrer, M., Travieso, C., J.F.Vargas & Alonso, J. (2006). Aplicación del Kernel de Fisher para la verificación de firmas manuscritas, *Terceras Jornadas de Reconocimiento Biométrico de Personas*, pp. 23–36.
- Fierrez-Aguilar, J., Alonso-Hermira, N., Moreno-Marquez, G. & Ortega-Garcia, J. (2004). An off-line signature verification system based on fusion of local and global information, *Proceedings of the Workshop on Biometric Authentication*, Springer LNCS-3087, pp. 298–306.
- Franke, K., Bünnemeyer, O. & Sy, T. (2002). Ink texture analysis for writer identification, *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition*, IEEE Computer Society, Washington, DC, USA, p. 268.

- Franke, K. & Grube, G. (1998). The automatic extraction of pseudodynamic information from static images of handwriting based on marked grayvalue segmentation (extended version), *Journal of Forensic Document Examination* 11(3): 17–38.
- Franke, K. & Rose, S. (2004). Ink-deposition model: The relation of writing and ink deposition processes, *Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition*, IEEE Computer Society, Washington, DC, USA, pp. 173–178.
- Freire, M., Fierrez, J., Martinez-Diaz, M. & Ortega-Garcia, J. (2007). On the applicability of off-line signatures to the fuzzy vault construction, *Proceedings of the Ninth International Conference on Document Analysis and Recognition*, Vol. 2, IEEE Computer Society, pp. 1173–1177.
- Gilperez, A., Alonso-Fernandez, F., Pecharroman, S., Fierrez, J. & Ortega-Garcia, J. (2008). Off-line signature verification using contour features, *Proceedings International Conference on Frontiers in Handwriting Recognition*.
- Güler, I. & Meghdadi, M. (2008). A different approach to off-line handwritten signature verification using the optimal dynamic time warping algorithm, *Digital Signal Processing* 18(6): 940–950.
- Haralick, R. M. (1979). Statistical and structural approaches to texture, *Proceedings of the IEEE* 67(5): 786–804.  
URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1455597](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1455597)
- He, D., Wang, L. & Guibert, J. (1987). Texture feature extraction, *Pattern Recognition Letters* 6(4): 269–273.
- Liu, S. & Silverman, M. (2001). A practical guide to biometric security technology, *IEEE IT Professional* 3(1): 27–32.
- Marcel, S., Rodriguez, Y. & Heusch, G. (2007). On the recent use of local binary patterns for face authentication, *International Journal on Image and Video Processing Special Issue on Facial Image Processing*.
- Nikam, S. & Agarwal, S. (2008). Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems, *Emerging Trends in Engineering and Technology, 2008. ICETET '08. First International Conference on*, pp. 675–680.
- Ojala, T., Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7): 971–987.
- Prakash, H. & Guru, D. (2009). Relative orientations of geometric centroids for off-line signature verification, *International Conference on Advances in Pattern Recognition*, 1: 201–204.
- T., M. (2003). *The local binary pattern approach to texture analysis - extensions and applications*, PhD thesis. Dissertation. Acta Univ Oul C 187, 78 p + App.  
URL: <http://herkules oulu.fi/isbn9514270762/>
- Thanasoulas, N., Evmiridis, N. & Parisi, N. (2003). Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their vis spectra, *Forensic science international* 138: 75–84.
- TowerGroup (Accessed: October, 2008). Press releases.  
URL: <http://www.towergroup.com/research/news/news.htm?newsId=4500>
- Trivedi, M., Harlow, C., Connors, R. & Goh, S. (1984). Object detection based on gray level co-occurrence, *Computer Vision, Graphics, and Image Processing* 28(3): 199–219.
- Vargas, J., Ferrer, M., Travieso, C. & Alonso, J. (2007). Off-line Handwritten Signature GPDS-960 Corpus, *Proceedings of the Ninth International Conference on Document Analysis and Recognition 2007*, pp. 764–768.

# Design and Evaluation of a Pressure Based Typing Biometric Authentication System

Momoh J. E. Salami, Wasil Eltahir and Hashimah Ali  
*International Islamic University Malaysia (IIUM)*  
*Malaysia*

## 1. Introduction

Although a variety of authentication devices to verify a user's identity are in use today for computer access control, passwords have been and probably would remain the preferred method. Password authentication is an inexpensive and familiar paradigm that most operating systems support. However, this method is vulnerable to intruder access. This is largely due to the wrongful use of passwords by many users and to the unabated simplicity of the mechanism which makes such system susceptible to unsubstantiated intruder attacks. Methods are needed, therefore, to either enhance or reinforce existing password authentication techniques.

There are two possible approaches to achieve this, namely by measuring the time between consecutive keystrokes "latency" or measuring the force applied on each keystroke. The pressure-based biometric authentication system (PBAS) has been designed to combine these two approaches so as to enhance computer security. PBAS employs force sensors to measure the exact amount of force a user exerts while typing. Signal processing is then carried out to construct a waveform pattern for the password entered. In addition to the force, PBAS measures the actual timing traces, which are often referred to as "latency".

Two approaches to construct user typing pattern have been implemented with PBAS. First approach utilizes a waveform acquired for user keystroke pressure along with time between each keystroke "latency" to create a unique user password typing pattern for authentication. An auto-regressive (AR) classifier is used for the pressure pattern, while a latency classifier is used for the time between keystrokes. The results of both classifiers are combined to authenticate the user typing pattern.

The second approach combines the pressure and latency by creating a pattern of peak keystroke force and latency. By combining the force and time features other classifiers have been tested with PBAS, namely support vector machines (SVM), artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS). Figure 1 illustrates how these classifiers are integrated to develop the system.

As compared to conventional keystroke biometric authentication systems, PBAS has employed a new approach by constructing a waveform pattern for the keystroke password. This pattern provides a more dynamic and consistent biometric characteristics of the user. It also eliminates the security threat posed by breaching the system through online network as the access to the system is only possible through the pressure sensor reinforced keyboard "biokeyboard".

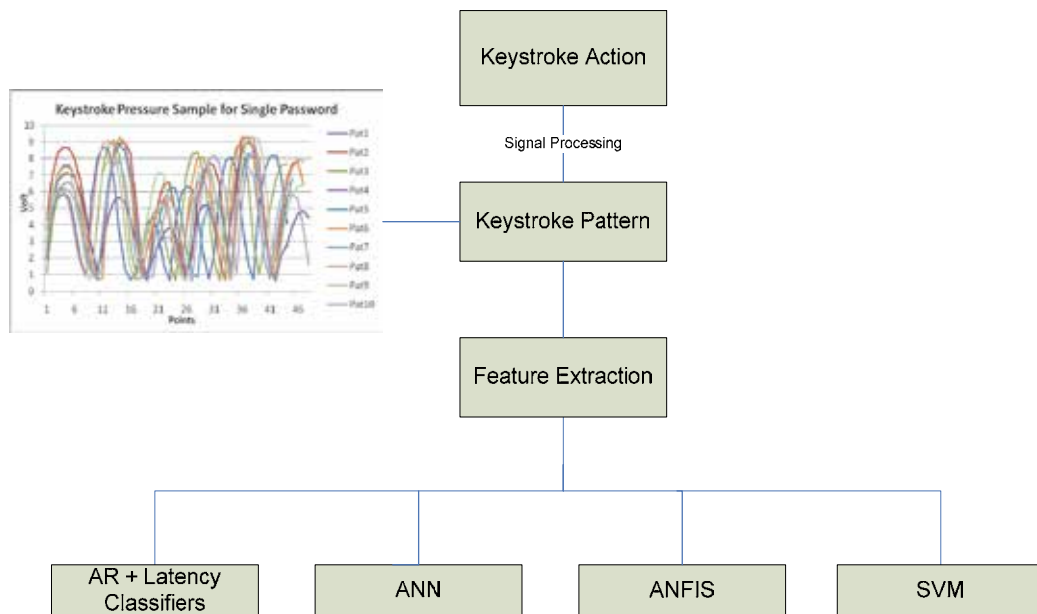


Fig. 1. Keystroke feature extraction and classification.

In order to further reinforce user authentication, PBAS has been integrated with iris system. Figure 2 shows the integrated system which can authenticate users by combining the iris and keystroke authentication mechanisms to achieve optimum results. Preliminary tests on the proposed system have produced promising results. The detailed design of PBAS system and data classifiers would be subsequently discussed. Some preliminary testing results obtained from the system would also be presented.

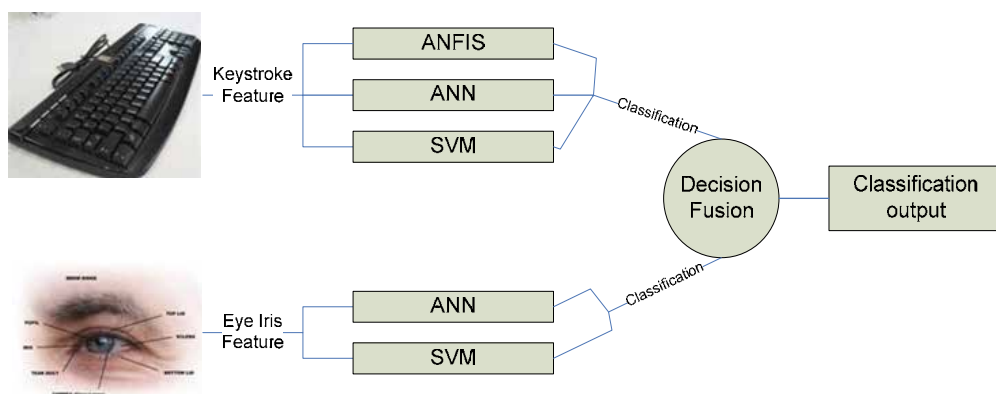


Fig. 2. Decision fusion for eye iris and keystroke classifiers.

## 2. Pressure based typing biometric authentication system

Keystroke authentication systems available in the market are mostly software-based. This is due to the ease of use as well as the low cost of the mechanism. Any new keystroke



authentication system has to consider these factors in the design. Likewise, the system designed for PBAS uses simplified hardware which minimizes the cost of production. The system is designed to be compatible with any type of PC. Moreover, it does not require any external power supply. In general, the system components are low cost and commonly available in the market.

## 2.1 System hardware components

The hardware layout of the system is unique. It consists of a desktop computer equipped with a data acquisition card, and a specially designed alphanumeric keyboard embedded with pressure sensors (Eltahir et al., 2003).

As illustrated in Figure 3, the main hardware components of PBAS are as follows:

1. Alphanumeric keyboard (biokeyboard) embedded with force sensors to measure the keystroke pressure while typing.
2. Data Acquisition System consisting of the following components:
  - a. Analogue Interface Box (filtering and amplification of signal).
  - b. DAQ PCI card fitted onto the PC.
3. PC/central processing unit (CPU) for running the PBAS program using Windows XP operating system.

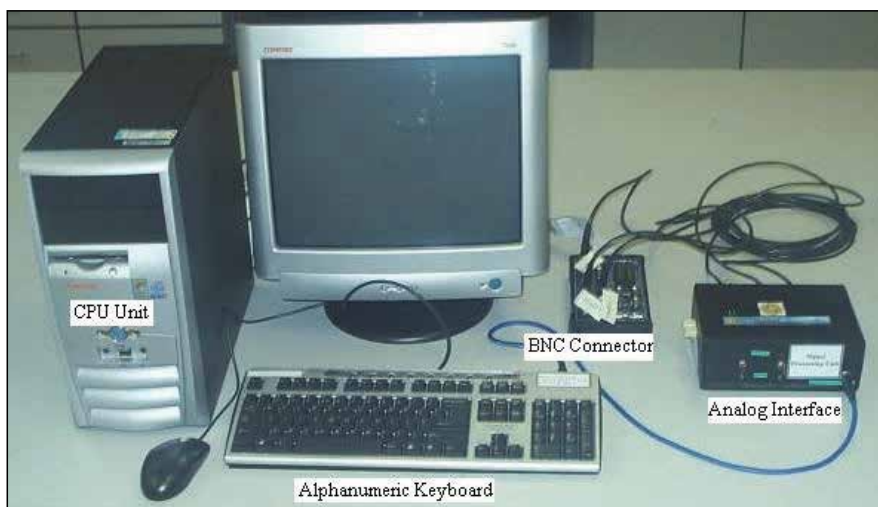


Fig. 3. PBAS system components.

A special keyboard has been manufactured to acquire the alphanumeric password and the keystroke pressure template of the user. The biokeyboard layout is identical to normal commercial keyboard. This is crucial to maintain an intrinsic system that does not alter user typing habits. Figure 4 shows the biokeyboard front, back and side views.

To measure the keystroke pressure, ultra-thin flexible force sensors are fixed below each keyboard key. A plastic spring is fixed between the key and the sensing area to ensure that it does not get dislodged. This is necessary to avoid erroneous readings.

The keyboard operates just as a normal alphanumeric keyboard in addition to measuring keystroke pressure. Thus, the users of this system would not find any differences between this keyboard and the commercial ones.



Fig. 4. Pressure sensitive alphanumeric keyboard (biokeyboard).

## 2.2 Software layout

The system starts by prompting user to enter his/her user ID and password. The alphanumeric keyboard (biokeyboard) extracts the pressure template for the password entered. At the same time, the system calculates the latency pairs for the entered password and accompanies it with pressure template in a single data file. This data file is transferred to the system's database.

In the learning mode, the user is required to repeatedly key in the password for several times to stabilize his/her keystroke template.

In the authentication mode, the user is requested to enter his/her ID and password. The resulting pressure template and latency vector are compared with those modelled in the database using data classifiers. Depending on the results of this comparison, the user will be either granted or denied access to the system. Figure 5 illustrates the operation of PBAS.

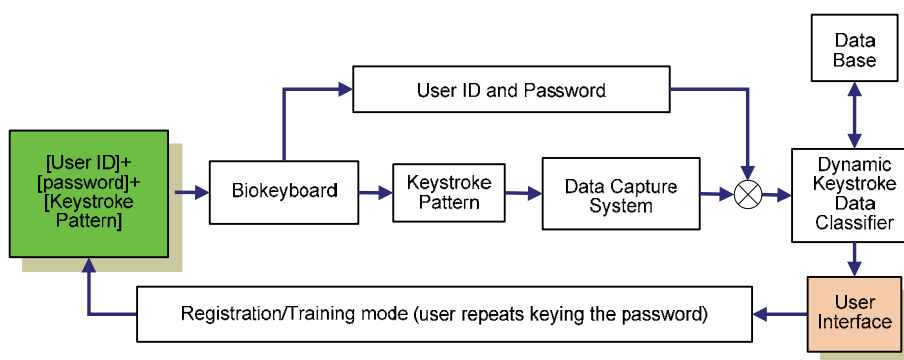


Fig. 5. PBAS block diagram.

### 2.3 Feature extraction

After acquiring the keystroke pressure template and latency pairs, PBAS applies feature extraction to prepare the pressure template for the specific classifier. The classifiers used in PBAS require two types of pressure templates, while the latency is the same for all classifiers.

A complete template is required for the AR classifier, while ANN, ANFIS, and SVM require a vector of highest pressure score for each keystroke. Figure 6 shows a complete keystroke template for single password entered by the user; this template is saved in the database suitable for the AR classifiers which is based on the signal modelling. Maximum pressure points are recorded and saved separately in the database for ANN, ANFIS and SVM classifiers.

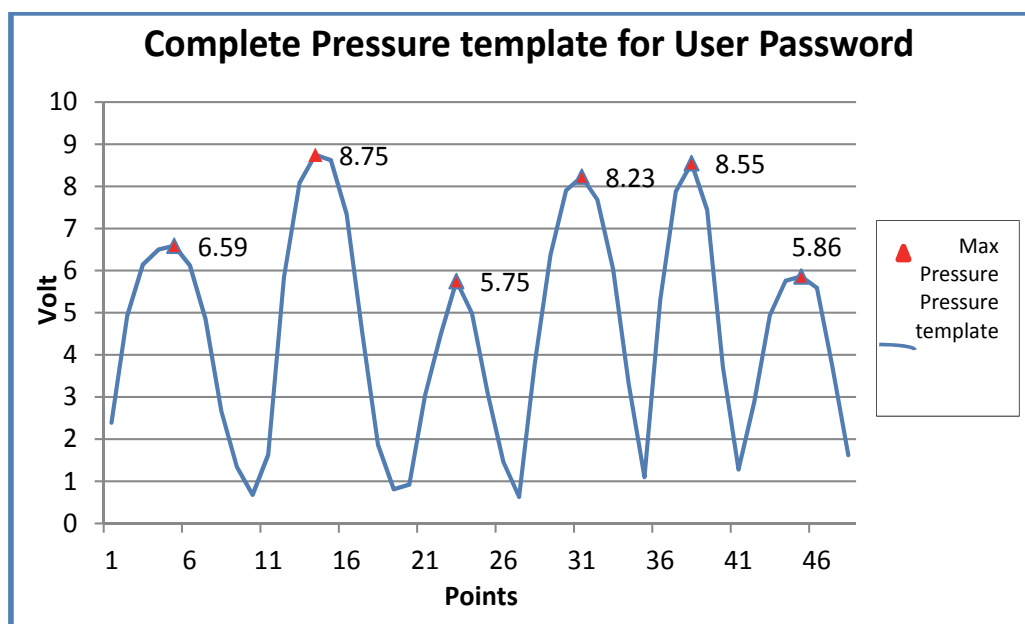


Fig. 6. Complete pressure template for single user password.

### 2.4 System algorithm and program structure

With the integration of software and hardware, the PBAS algorithm has been designed to have two main operational modes:

1. Training users and creating biometric template profiles; at this stage the user is requested to key-in his/her ID and the user trains his/her password.
2. Authenticating existing users based on their claimed identity; users provide ID and password which are compared with the biometric profiles of the users in the database.

Figure 7 shows the flow graph for the overall PBAS training-authentication processes. The authentication mode consists of two phases:

1. Normal authentication, which involves the password combination and its compliance with the one saved in the database.
2. Biometric authentication, which is done by combining latency with the keystroke pressure classifiers.

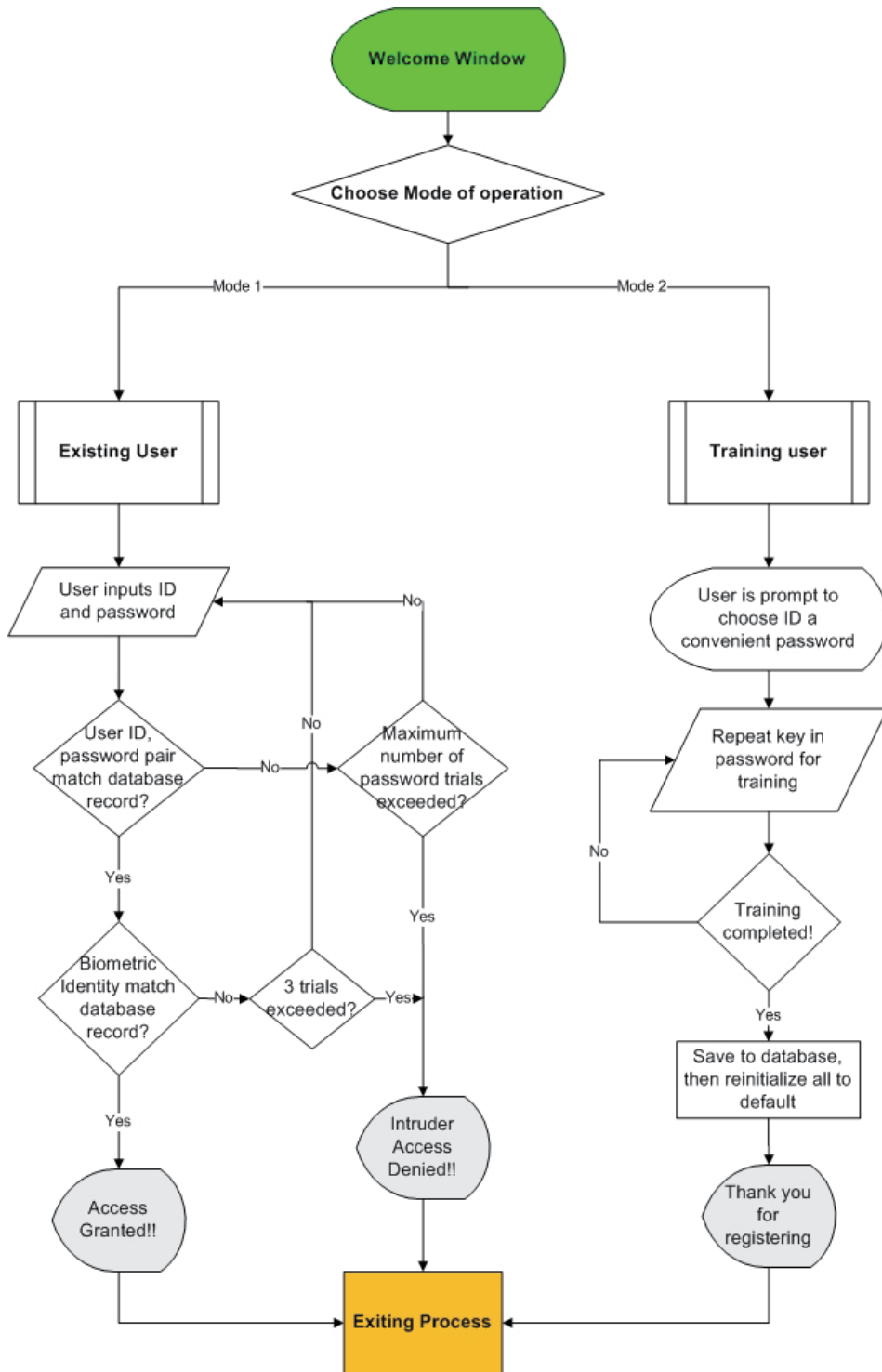


Fig. 7. PBAS main algorithm flowchart.

Firstly, the user will select the mode of operation. In the training mode, the access-control system requests the user to type in the login ID and a new password. The system then asks the user to re-enter the user ID and password to train the new password. The resulting latency and pressure keystroke templates are saved in the database. During training, if the user mistypes the password the system prompts the user to re-enter the password from the beginning. The use of backspace key is not allowed as it can disrupt the biometric pattern. When registration is done, the system administrator uses these training samples to model the user keystroke profile. The design of the user profile is done offline. Finally, the system administrator saves the user's keystroke template model along with the associated user ID and password in the access-control database.

In the authentication mode, the access-control system requests the user to type in the login ID and a password. Upon entering this information the system compares the alphanumeric password combination with the information in the database. If the password does not match, the system will reject the user instantly and without authenticating his/her keystroke pattern. However, if the password matches then the user keystroke template will be calculated and verified with the information saved in the database. If the keystroke template matches the template saved in the database, the user is granted access.

If the user ID and alphanumeric password are correct, but the new typing template does not match the reference template, the security system has several options which can be revised occasionally. A typical scenario could be that PBAS advises a security or network administrator that the typing pattern for a user ID and password is not authentic and that a security breach might be in progress. The security administrator can then closely monitor the session to ensure that the user does nothing un-authorized or illegal.

Another practical situation applies to the automatic teller machine "ATM" system. If the user's password is correct but the keystroke pattern does not match, the system can restrict the amount of cash withdrawn on that occasion to minimize the possibility (and penalty) of theft.

### 3. Dynamic data classifiers

#### 3.1 Latency classifier

Identity authentication using keystroke latency timing for password digraphs has been investigated thoroughly by many researchers (Joyce & Gupta, 1990; Monroe & Ruben, 1997; De Ru & Eloff, 1997; Ord & Furnell, 2000; Araujo et al, 2005). Many useful methodologies have been presented and are in use with the current latency keystroke authentication systems available in the market.

Joyce and Gupta discussed the design of an identity verifier based on four input strings (login name, password, first name, last name). The verification is done by comparing the mean reference signature  $M$  with a test signature  $T$ . The norm  $\|M - T\|$  is computed and if it is less than the threshold for the user, the attempt is accepted; otherwise it is flagged as an imposter attempt.

Though this approach produces some satisfactory results, it requires a relatively lengthy input string (Joyce & Gupta, 1990). A modified approach has been devised here for PBAS latency authentication. PBAS uses only the password string for latency verification.

##### 3.1.1 Creating mean reference latency vector

1. Registering users are prompt to re-enter their password several times, latency vector for each trial is saved in an individual data file resulting in ( $n$ ) number of files in the database, where  $n$  is the number of trials.

2. Data treatment is applied on the data files to remove outliers and erroneous values.
3. An average latency vector is calculated using the user trial sample. This results in a single file containing the mean latency vector ( $R$ ) for  $n$  password trials. This file is used as reference for latency authentication.

### 3.1.2 Calculating suitable threshold

Thresholding is used to decide an acceptable difference margin between the reference latency vector ( $R$ ) and the latency vector provided by the user upon verification ( $V$ ). The threshold is computed based on the data files saved in the database. A threshold is set for each user based on the variability of his/her latency signatures. A user that has little variability in his/her latencies would have a small threshold, while another with high variability should have larger threshold. Standard deviation is the variability measure used. The standard deviation between the mean ( $R$ ) latency vector and the user sample is computed. A threshold based on the standard deviation is calculated by the following equation:

$$\sum_{k=1}^{m-1} |R_k - V_k| \leq c * d \quad (1)$$

where  $m$  is the password length,  $R_k$  is the  $k$ th latency value in the reference latency vector,  $V_k$  is the  $k$ th latency value in the user-inputted latency vector,  $c$  is an access threshold that depends on the variability of the user latency vector,  $d$  is the distance in standard deviation units between the reference and sample latency vectors.

In order to classify user attempt, the latency score  $S_L$  for the user attempt is defined as

$$S_L = \frac{\sum_{k=1}^{m-1} |R_k - V_k|}{c * d} \quad (2)$$

Therefore, depending on the value of  $S_L$  the classifier output can be expressed as

$$S_L \begin{cases} \leq 1, & \text{accept template} \\ > 1, & \text{reject template} \end{cases} \quad (3)$$

Table 1 shows the reference latency vector for user "MJE1" which was calculated by the above mentioned method for a sample of 10 trials. Five latency vectors are used to test the threshold  $c$  for this reference profile (Table 1). The standard deviation was calculated to be  $S_y=46.5957$  m.sec and a threshold of two standard deviations above the mean ( $c=2.0$ ) resulted in the following variation interval  $253.9748 \geq R-V \geq 67.83189$ . This threshold takes in all 5 trials of the user. However, this is a relatively high threshold value and in many practical situations such values would only be recommended for unprofessional users who are usually not very keen typists. The user here is a moderate typist. This is evident by his/her relatively high standard deviation. High standard deviation is also a measure of high variability in the users' typing pattern; this usually indicates that the user template has not yet stabilized, perhaps due to insufficient training.

Table 2 shows the variation of threshold values  $c$  (from 0.5 to 2.0) and its effect on accepting the user trials. For this user, a threshold value that is based on standard deviation of 2.0 provides an acceptance rate of 100% (after eliminating outliers). However, a high threshold value would obviously increase the imposter pass rate. Therefore for normal typists, the threshold values should only be within the range of 0.5 to 1.5.

An experiment was conducted to assess the effect of varying the latency threshold value on the false acceptance rate (FAR) and the false rejection rate (FRR). In this experiment an ensemble for 23 authentic users and around 50 intruders were selected randomly to produce authentic and intruder access trials. Authentic users were given 10 trials each while intruders were given 3 trials per account. All trials were used for the calculations and no outliers were removed.

| Password (asd123) | Reference Vector | T 1 | T 2 | T 3 | T 4 | T 5 |
|-------------------|------------------|-----|-----|-----|-----|-----|
| <b>a-s</b>        | 201.9231         | 203 | 187 | 172 | 203 | 235 |
| <b>s-d</b>        | 162.8571         | 156 | 188 | 171 | 156 | 93  |
| <b>d-1</b>        | 316.2308         | 235 | 187 | 219 | 188 | 250 |
| <b>1-2</b>        | 185.2000         | 187 | 203 | 203 | 187 | 203 |
| <b>2-3</b>        | 108.5714         | 110 | 110 | 110 | 94  | 79  |

Table 1. Reference latency tested against 5 authentic user trials.

| Standard Deviation | Upper Limit | Lower Limit | Acceptance Percentage |
|--------------------|-------------|-------------|-----------------------|
| <b>0.5</b>         | 184.1712    | 137.6355    | 40                    |
| <b>1.0</b>         | 207.439     | 114.3676    | 60                    |
| <b>1.5</b>         | 230.7069    | 91.09975    | 80                    |
| <b>2.0</b>         | 253.9748    | 67.83189    | 100                   |

Table 2. Effect of threshold value on user acceptance rate.

Figure 8 shows that the equal error rate (EER) for the FAR and the FRR rates was 24% and it occurred at a threshold value of 2.25. This relatively high FRR rate is expected since the password strings used were mainly short in length and weak in strength.

### 3.2 AR-burg classifier

The AR algorithm uses the notion of signal analysis to reproduce the users' keystroke pressure template which is then compared with the keystroke template produced by the alleged intruders. Based on this comparison an authentication decision is made.

A signal model approach is advocated here since the pressure template points are interrelated across time. The AR signal model is expressed as

$$y(n) + a(1)y(n-1) + a(2)y(n-2) + \dots + a(p)y(n-p) = x(n) \quad (4)$$

where  $n$  is the time index,  $y(n)$  is the output,  $x(n)$  is the input, and  $p$  is the model order.

For signal modelling  $y(n)$  becomes the signal to be modelled and the  $a(i)$  coefficients need to be estimated.

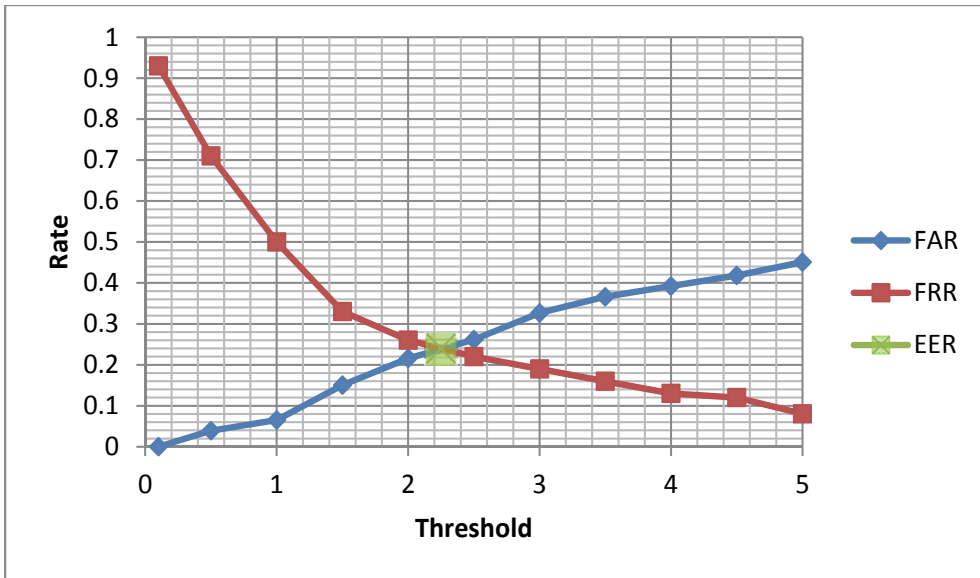


Fig. 8. Latency threshold versus FAR and FRR rates.

Equation (4) can be used to predict future values of the signal  $\hat{y}(n)$  where

$$\hat{y}(n) = -a(1)y(n-1) - a(2)y(n-2) - \dots - a(p)y(n-p) \tag{5}$$

Denoting the model error as  $e(n)$  then

$$y(n) + a(1)y(n-1) + a(2)y(n-2) + \dots + a(p)y(n-p) = e(n) \tag{6}$$

The Total Squared Error (*TSE*) for the predicted signal is

$$TSE = \sum_{n=1}^{N-1} e_n^2 \tag{7}$$

The AR model is used most often because the computation of its parameters is simpler and more developed than those of moving average (*MA*) and autoregressive moving average (*ARMA*) models (Shiavi, 1991; Manson, 1996).

Burg method has been chosen for this application because it utilizes both forward and backward prediction errors for finding model coefficients. It produces models of lower variance ( $S_p^2$ ) as compared to other techniques (Shiavi, 1991).

Authentication is done by comparing the *TSE* percentage saved in the users' database with that generated by the linear prediction model. Previous experiments proved that authentic users can achieve *TSE* margin of less than 10% (Eltahir et al., 2004).

### 3.2.1 Identifying optimum pressure template for AR model

An algorithm has been developed in MATLAB to identify the best pressure template among the user sample. This pattern is used for estimating the AR model parameters for the user keystroke pressure pattern. The algorithm uses the correlation technique to calculate the



accumulative correlation index "ACI", which is the accumulation of the correlation between each pressure pattern and the whole sample. The pattern with the highest ACI is selected for the model.

### 3.2.2 Identifying the optimum TSE acceptance margin

The TSE relative prediction error  $RPE$  is calculated according to

$$\text{Relative Prediction Error} = \left| \frac{TSE_m - TSE_s}{TSE_m} \right| \quad (8)$$

where  $TSE_m$  is the  $TSE$  calculated for the user's AR-Burg model in database.  $TSE_s$  is the  $TSE$  for the pressure pattern of the user.

Classification of user attempt is done by comparing  $RPE$  to threshold  $T$  according to the following equation:

$$RPE \begin{cases} \leq T, & \text{accept template,} \\ > T, & \text{reject template,} \end{cases} \quad (9)$$

where  $0 < T \leq 1$

Based on previous research experiments it has been reported that authentic users can achieve up to 0.1  $RPE$  while intruders exhibiting unbounded fluctuation can have  $RPE$  above 3 (Eltahir et al., 2004).

An experiment was conducted to assess the effect of varying the TSE threshold value on the FAR and FRR rates. In the experiment an ensemble of 23 authentic users and around 50 intruders were selected randomly to produce authentic and intruder access trials. Authentic users were given 10 trials each while intruders were given 3 trials per account. All trials were used for the calculation of results with no outliers removed. Figure 9 shows the variation of FAR and the FRR against the TSE threshold values. The EER was 25% and it was recorded at TSE of 37.5%. When compared to latency, TSE has lower FRR spread out as the threshold is increased.

The AR modelling algorithm has been implemented in the following order:

1. The user is prompted to enter the password several times.
2. The optimum pattern for modelling the user pressure template is identified using the ACI values obtained from the sample.
3. The best AR model order is determined based on the final prediction error (FPE) and the Akaike's information criteria (AIC) (Akaike, 1970).
4. The AR model is constructed and the model coefficients are saved for user verification.
5. When a user enters a password, the linear prediction model is constructed to create a pressure template from the users' keystroke.  $TSE_s$  is calculated for this template.
6. From the system database,  $TSE_m$  is used to calculate the  $RPE$  score.
7. If  $RPE \leq T$  user is authentic, whereas if  $RPE > T$  then user is an intruder.

### 3.3 Combined latency and AR classifiers

The receiver operating characteristic curve (ROC) is used to assess the effect of the threshold value on the FAR and FRR rates. ROC curve assesses the trade-off between low intruder

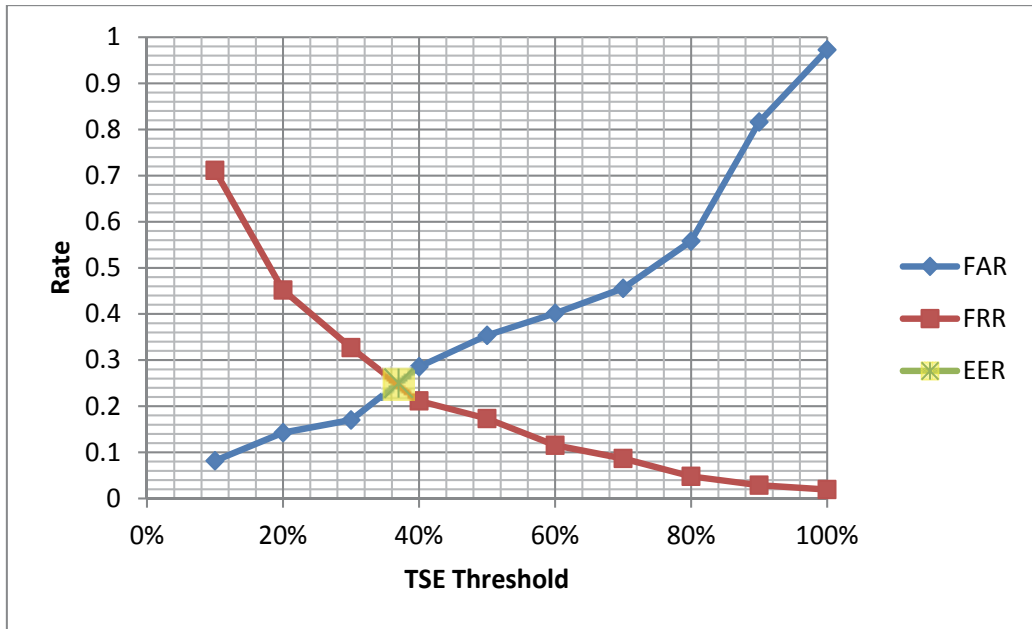


Fig. 9. TSE threshold versus FAR and FRR rates.

acceptance rate (IAR) and high authentic acceptance rate (AAR) as the decision threshold value varies. Figure 10 shows that the latency classifier has slightly better separation than the AR classifier. In addition, the latency classifier has better intruder rejection rate whereas AR classifier has a higher true pass rate. This graph also shows that the performance of both

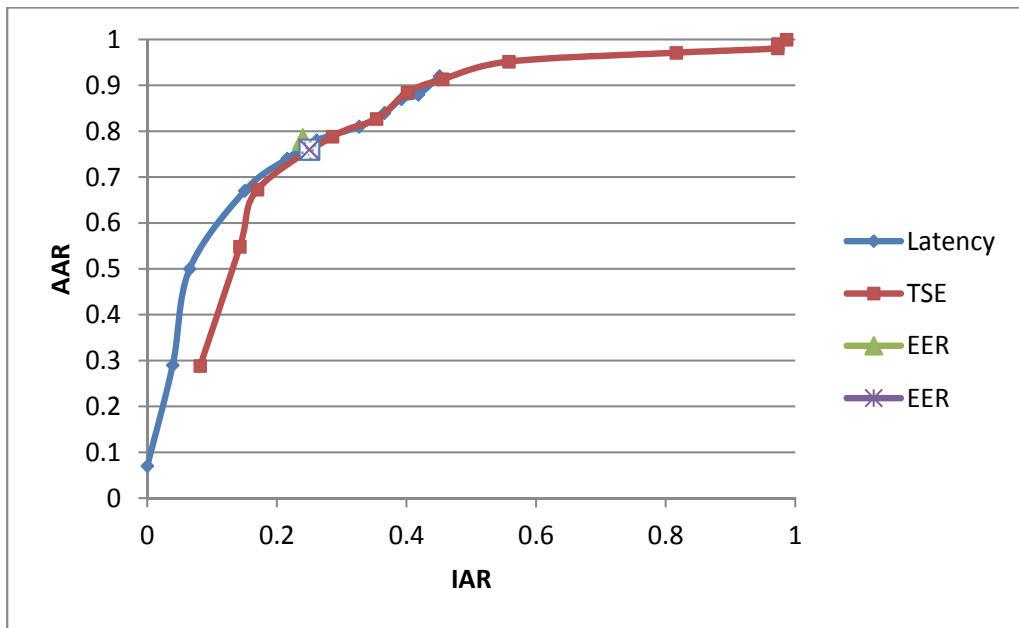


Fig. 10. ROC showing performance of latency and TSE.

classifiers at the EER points is very similar; therefore it is expected that by combining both algorithms the overall system performance would be improved.

The threshold used for the TSE classifier is  $T = 0.4$  based on the calculated EER. As for the latency threshold  $c$ , it is recommended to use a threshold value of 2.0 to 2.25 for unprofessional typists and 1.0 to 1.5 for professional typists.

An experiment has been conducted with the combined latency and AR classifiers, whereby 23 users participated in the data collection. All results were calculated online. The results showed FRR of 3.04 and FAR of 3.75 (Eltahir et al. 2008).

### 3.4 Artificial Neural Network (ANN)

In general, an ANN is characterized by its architecture, learning algorithm, and activation function. The architecture describes the connections between the neurons and consists of an input layer, an output layer and at least one hidden layer.

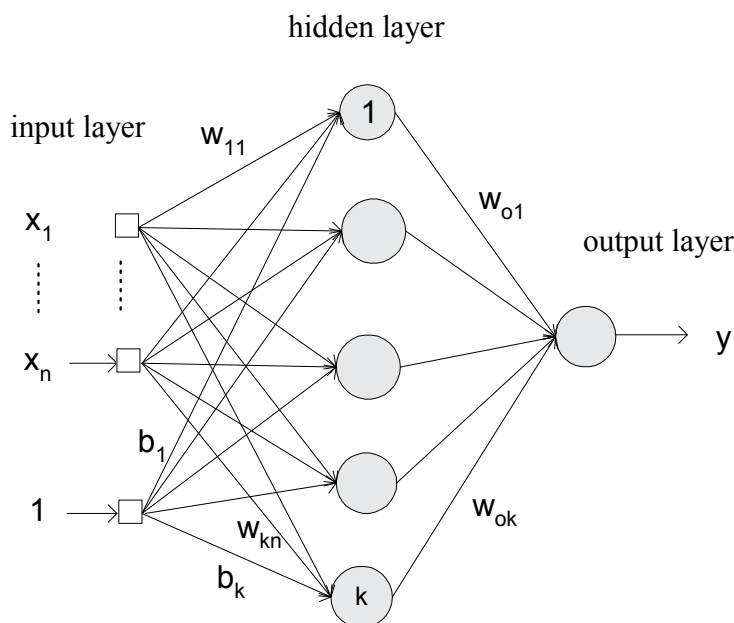


Fig. 11. A Typical multilayer feedforward neural network structure.

Among the available architectures of the ANN, Multilayer Feedforward Network (MFN) is a very popular network. The *learning algorithm* or *adaptation rule* refers to the way the weights are changed in order to achieve a desired mapping between input and output. The backpropagation (BP) learning algorithm is one of the most important historical developments in neural network (Lin & George, 1995). It is an iterative gradient descent algorithm designed to minimize the *mean square error* (MSE) between the actual output of a backpropagation neural network (BPNN) and the desired output (Jang et al., 1996). MFN trained with backpropagation algorithm has been widely used in many engineering applications due to its capability as a general approximation. Figure 11 shows the MFN structure with input  $x_1, x_2, \dots, x_n$  and an output  $y$ . The connections between the neurons of the different layers are called weight and bias.

### 3.4.1 ANN keystroke model

Using multilayer feedforward network and the backpropagation algorithm, the process of developing keystroke models based on the typing biometric data can be summarized in the following steps:

1. The weight of the network is initialized with random values and the number of iteration ( $n$ ) is set to zero.
2. Enter the input vector  $p$  from the training dataset to the network for the first time.
3. Send the input vector  $p$  through the network to get an output

$$o_{pk} = f \left( \sum w_{jk} f \left( \sum_m w_{mj} f \left( \dots f \sum_i w_{il} x_i \right) \right) \right) \quad (10)$$

4. Evaluate the error value by computing the difference between the real output and the desired output: calculate also the sum of the squared error and increment  $n$ .
5. Carry out the back-propagation of output layer error to all elements of network.
6. For every unit of output  $k$ , calculate,

$$\delta_k = (O_k - y_k) f'(net_k) \quad (11)$$

7. For every hidden unit  $j$ , calculate,

$$\delta_j = f'(net_j) \sum_k \delta_k w_{jk} \quad (12)$$

8. Update the weights of all elements between output and hidden layer and then between all hidden layers moving toward the input layer. Changes of the weights can be obtained as

$$\Delta w_{ij}^{(n+1)} = \eta \delta_j o_i + \alpha \Delta w_{ij}^{(n)} \quad (13)$$

where  $\alpha$  is a constant called momentum which serves to improve the learning process.

9. Repeat the step 2-8 with the next input vector until the error becomes small or maximum number of iterations is reached.

### 3.4.2 Training and evaluating the ANN based model

To evaluate the effectiveness of the proposed classifier, a group of five (5) participants was used; the collected data size was 1000. Three (3) of the participants were considered as authorized persons while the other two (2) were considered as imposters. Each person was required to type a password of six characters of his/her own choice for 200 times, 100 data were used for training whereas the rest were used for data testing. Each typed character of the password had maximum pressure and latency features. For example, password "asd123" consists of five latency and six maximum force features, resulting in eleven features for each trial. The performance of the classifier was assessed by FRR and FAR rates.

Table 3 shows the result of training the user data. User 1 and user 3 have single hidden layer with fifteen and six neurons respectively. Meanwhile, user 2 has two hidden layers consisting of seven (7) and six (6) neurons for the first and the second hidden layers respectively. The combined features (maximum pressure and latency) gave perfect classification rates for all users. Furthermore, the MFN model can be trained in less than one second.

| Model   | Authorized User | Training Time(sec) | Classification Rate (%) |
|---------|-----------------|--------------------|-------------------------|
| [15 1]  | User 1          | 0.9062             | 100                     |
| [7 6 1] | User 2          | 0.8125             | 100                     |
| [6 1]   | User 3          | 0.5312             | 100                     |
| Average |                 | 0.7499             | 100                     |

Table 3. Training performance of ANN user model.

Table 4 shows the user verification results for the user model. The proposed system gives good average FRR which is about 0.67.

| Model   | Authorized User | FRR (%) | FAR (%)<br>Close Set | FAR (%)<br>Open Set |
|---------|-----------------|---------|----------------------|---------------------|
| [15 1]  | User 1          | 1       | 0                    | 5                   |
| [7 6 1] | User 2          | 1       | 0                    | 38                  |
| [6 1]   | User 3          | 0       | 0                    | 0.5                 |
| Average |                 | 0.67    | 0                    | 14.5                |

Table 4. Verification performance of ANN user model.

### 3.5 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is an architecture which is functionally equivalent to a Sugeno type fuzzy rule base (Jang et al., 1996). It is a method for tuning an existing rule base with a learning algorithm based on a collection of training data.

ANFIS is an approach that integrates the interpretability of a fuzzy inference system with the adaptability of a neural network. Moreover, because ANFIS has much less tuneable parameters than traditional neural networks, it has the advantage of being significantly faster and more accurate than many ANN-based methods.

An ANFIS architecture corresponding to a Sugeno fuzzy model for two inputs and single output is shown in Figure 12. It is a feedforward network consisting of 5 layers. To simplify

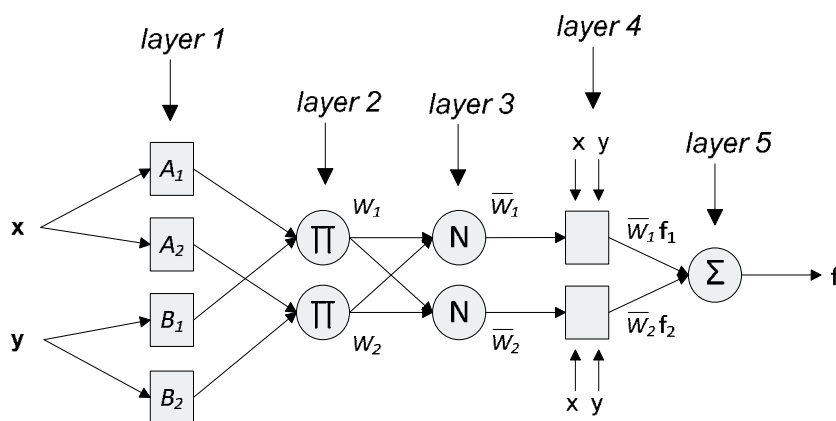


Fig. 12. ANFIS architecture (Jang et al., 1997).

the notation, consider the two inputs,  $x$  and  $y$  and one output  $f$ . A rule in the first order Sugeno fuzzy inference system has the form:

**Rule 1:** If  $x$  is  $A_1$  and  $y$  is  $B_1$ , then

$$f_1 = p_1x + q_1y + r_1$$

**Rule 2:** If  $x$  is  $A_2$  and  $y$  is  $B_2$ , then

$$f_2 = p_2x + q_2y + r_2$$

The parameters  $p_1, p_2, q_1, q_2, r_1$  and  $r_2$  are linear, whereas  $A_1, A_2, B_1$  and  $B_2$  are nonlinear.

The five layers in ANFIS are fuzzification, production, normalization, defuzzification, and aggregation layer arranged in this order. The following are the input and output relationships for each layer:

**Layer 1 (Fuzzification layer):**  $A_1, A_2, B_1$  and  $B_2$  are the linguistic expressions which are used to distinguish the membership functions (MFs). The relationships between the input-output and MFs are:

$$O_{1,i} = \mu_{A_i}(x), \text{ for } i = 1, 2 \quad (14)$$

$$O_{1,j} = \mu_{B_j}(y), \text{ for } j = 1, 2, \quad (15)$$

where  $O_{1,i}$  and  $O_{1,j}$  denote the output function of the first layer,  $\mu_{A_i}(x)$  and  $\mu_{B_j}(y)$  denote MFs so that

$$\mu_{A_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad (16)$$

where  $\{a_i, b_i, c_i\}$  is the parameter set. As the values of these parameters change, the bell-shaped functions vary accordingly. In fact, any continuous and piecewise differentiable functions such as trapezoidal or triangular-shaped functions are also qualified for use as MF. Parameters in this layer are referred to as *premise parameters*.

**Layer 2 (Production layer):** Every node in this layer is marked as symbol  $\Pi$  the outputs are  $w_1$  and  $w_2$ , the weight functions of the next layer are

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \text{ } i=1, 2. \quad (17)$$

where  $O_{2,i}$  is the output function of the second layer.

**Layer 3 (Normalization layer):** The node is marked as  $\mathbf{N}$ , and it is used to normalize the weight functions, that is

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \text{ } i=1, 2. \quad (18)$$

where  $O_{3,i}$  is the output function of the third layer. For convenience, outputs of this layer are called normalized firing strengths.

**Layer 4 (Defuzzification layer):** Being an adaptive node,  $\overline{w}_i$  is the output and  $\{p_i, q_i, r_i\}$  is the parameter set in this layer. The relationship between input and output is

$$O_{4,i} = \overline{w_i f_i} = \overline{w_i}(p_i x + q_i y + r_i) \quad (19)$$

where  $O_{4,i}$  is the output function of the fourth layer. Parameters in this layer are referred to as consequent parameters.

**Layer 5(Total output layer):** The single node in this layer is an output corresponding to the aggregate of all inputs so that the overall output is

$$O_{5,1} = \sum_i \overline{w_i f_i} = \frac{\sum_i \overline{w_i f_i}}{\sum_i \overline{w_i}} \quad (20)$$

where  $O_{5,i}$  is the output function  $f$  of the fifth layer.

### 3.5.1 ANFIS keystroke model

The steps for creating an ANFIS model for user typing pattern are as follows:

1. User retypes password until sufficient data is acquired for feature extraction.
2. Determination of the premise parameters.
3. Training of the ANFIS using the input pattern and the desired output to obtain the consequent parameters.
4. Validation of the trained ANFIS system.

### 3.5.2 Training and evaluating the ANFIS based model

In order to evaluate the user ANFIS typing model, an experiment was carried out with a group of seven (7) persons with data sample size of 1400. Five (5) persons were considered as authorized users, while the other two (2) persons were assumed as impostors. Each person was required to type a password of six characters of his/her own choice for 200 times. 100 samples were used for training, whereas the rest were used for data testing. Both impostors had 200 trials to break into the system. Table 5 shows the training results using a cluster radius of 0.25. It is noted that the training time was relatively long (about 45 seconds).

| Model   | Authorized User | Training Time(sec) | Classification Rate (%) |
|---------|-----------------|--------------------|-------------------------|
| ANFIS 1 | User 1          | 47.4688            | 100                     |
|         | User 2          | 45.9686            | 100                     |
|         | User 3          | 45.6406            | 100                     |
|         | User 4          | 43.5781            | 100                     |
|         | User 5          | 43.2344            | 100                     |
|         | Average         | 45.1781            | 100                     |

Table 5. Training performance of ANFIS of radius =0.25.

By increasing the cluster radius to 0.5, the training time was reduced dramatically to 1.75 seconds. Furthermore, classification rate was maintained at 100% except for user 4 which dropped slightly to 98%.

Furthermore, increasing the radius cluster to 1, training time was further reduced to 1.6 seconds while maintaining the average classification rate at 99.6% as shown in Table 6.

| Model         | Authorized User | Training Time(sec) | Classification Rate (%) |
|---------------|-----------------|--------------------|-------------------------|
| ANFIS 3 (R=1) | User 1          | 1.5469             | 100                     |
|               | User 2          | 1.5625             | 99                      |
|               | User 3          | 1.5625             | 100                     |
|               | User 4          | 1.5781             | 99                      |
|               | User 5          | 1.5469             | 100                     |
|               | Average         | 1.5594             | 99.6                    |

Table 6. Training performance of ANFIS of radius =1.

It is observed that the training time varies significantly for different cluster radius. A smaller radius leads to a longer training time due to the formation of smaller clusters in the data, this creates more rules and a larger number of consequent parameters to be tuned and optimized in ANFIS system. Hence it can be concluded that a larger radius is preferable to achieve shorter training time. However, larger radius tends to give lower accuracy as compared to smaller radius.

Table 7 shows performance results for a radius of 0.25. The FRR is about 9%, FAR for closet set was 3.2% which is relatively low. However, FAR for open set is relatively very high for user 2 and user 5.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 0       | 1.5               | 1                |
| User 2          | 8       | 1.75              | 39.5             |
| User 3          | 17      | 0.5               | 0.5              |
| User 4          | 16      | 10                | 14               |
| User 5          | 4       | 2.25              | 47.5             |
| Average         | 9       | 3.2               | 20.5             |

Table 7. Testing performance of ANFIS of radius = 0.25.

Table 8 shows the results obtained when the radius was increased to 0.5 in which the FRR results were significantly improved. The average value for FRR dropped to 3.2 as compared to 9 in previous results. The FAR for closet set is greatly improved with an average value of 0.45. However, FAR for open set shows only slight improvement.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 1       | 0                 | 1                |
| User 2          | 12      | 0.25              | 31.5             |
| User 3          | 1       | 0.5               | 3                |
| User 4          | 1       | 0.5               | 0                |
| User 5          | 1       | 1                 | 48               |
| Average         | 3.2     | 0.45              | 16.7             |

Table 8. Testing performance of ANFIS radius = 0.5.



By selecting a cluster radius of 1, the performance of the ANFIS model improves further as shown in Table 9. The average FRR dropped to 1.2, while FAR for the close set dropped to an average value of 0.15. However, FAR of the open set is still relatively high with an average of about 15.7. Therefore, further improvement is needed to enable the proposed system to produce a small FAR (of the open set), that is smaller than 1.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 0       | 0                 | 1                |
| User 2          | 5       | 0                 | 29               |
| User 3          | 0       | 0.5               | 0                |
| User 4          | 0       | 0                 | 0                |
| User 5          | 1       | 0.25              | 48               |
| <i>Average</i>  | 1.2     | 0.15              | 15.7             |

Table 9. Testing performance of ANFIS radius = 1.

In terms of FRR and FAR, it is observed that a radius of 1 produced the best results here.

### 3.6 Support Vector Machines based keystroke model

Support Vector Machines (SVMs) is a relatively new learning machine technique, which is based on the principle of structural risk minimization (minimizing classification error). SVM is a binary classifier that optimally separates two classes of data (Burges, 1998). Figure 13 shows the flowchart of the SVM algorithm.

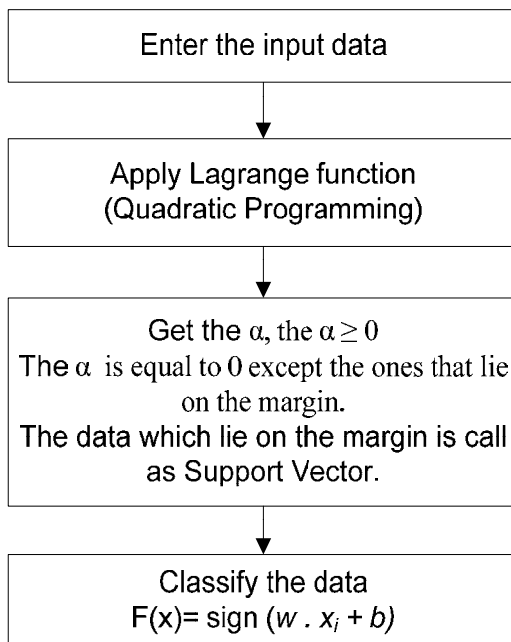


Fig. 13. Flowchart of SVM algorithm.

### 3.6.1 Training and evaluating the SVM based model

In order to evaluate the user SVM typing model, an experiment was carried out with a group of seven (7) persons. The sample size was 1400. Five (5) of the participants were considered as authorized users, while the other two (2) participants were considered as impostors. Each user was required to type a password of six characters of his/her own choice for 200 times, 100 trials were used for training whereas the rest were used for data testing. Both impostors had two hundred (200) trials to break into the system.

SVM with fifth-order polynomial kernel function was used for developing keystroke typing biometric models. Each authorized user had his/her individual SVM-based model characterized by a set of support vectors. The support vectors are obtained by developing quadratic algorithm for use in the MATLAB toolbox. A penalty term  $C$  of 100 is used to anticipate misclassified data.

Table 10 shows the training time and the classification results obtained, it is observed that the entire SVM-based models of authorized users gave perfect classification of 100%, i.e. there were no errors in identifying the authorized users. Furthermore, all the SVM-based models of authorized users can be trained in a relatively short time of less than 0.5 second on average. Therefore, it can be concluded that the SVM models produce promising results when used in the proposed system.

| Authorized User | Training Time(sec) | Classification Rate (%) |
|-----------------|--------------------|-------------------------|
| User 1          | 0.0625             | 100                     |
| User 2          | 0.6563             | 100                     |
| User 3          | 0.4687             | 100                     |
| User 4          | 0.6250             | 100                     |
| User 5          | 0.2815             | 100                     |
| <i>Average</i>  | 0.4188             | 100                     |

Table 10. Training performance of SVM.

By examining the results shown in Table 11 it is observed that the average FRR rate is 5.6% which is acceptable as compared to the previous results. The average FAR for close set was 1.75% which is relatively good. However, the FAR for open set is high, user 5 has FAR of 48% which suggests that his password typing pattern is weak and easy to imitate, whereas user 4 has a strong password pattern that is hard to imitate. Further improvements are needed as the average value for the FAR is 14.7.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 2       | 0                 | 10               |
| User 2          | 7       | 1.75              | 14               |
| User 3          | 19      | 0.5               | 1.5              |
| User 4          | 0       | 0.75              | 0                |
| User 5          | 0       | 1.75              | 48               |
| <i>Average</i>  | 5.6     | 0.95              | 14.7             |

Table 11. Testing performance of SVM.

Many experiments have been conducted which indicate that the proposed system, involving the combined maximum pressure and latency features could produce low FRR and FAR values.

#### 4. IRIS recognition

The number of features in human iris is large. Its complex pattern can contain many distinctive features such as arching, ligaments, furrows, ridges, crypts, rings, corona, freckles and zigzag collarette. Iris pattern can have up to 249 independent degrees of freedom; therefore it is very difficult to falsify an iris pattern. As a result of this, iris pattern is a very attractive feature for user verification (Daugman, 2003; Masek, 2003; Wildes, 1997).

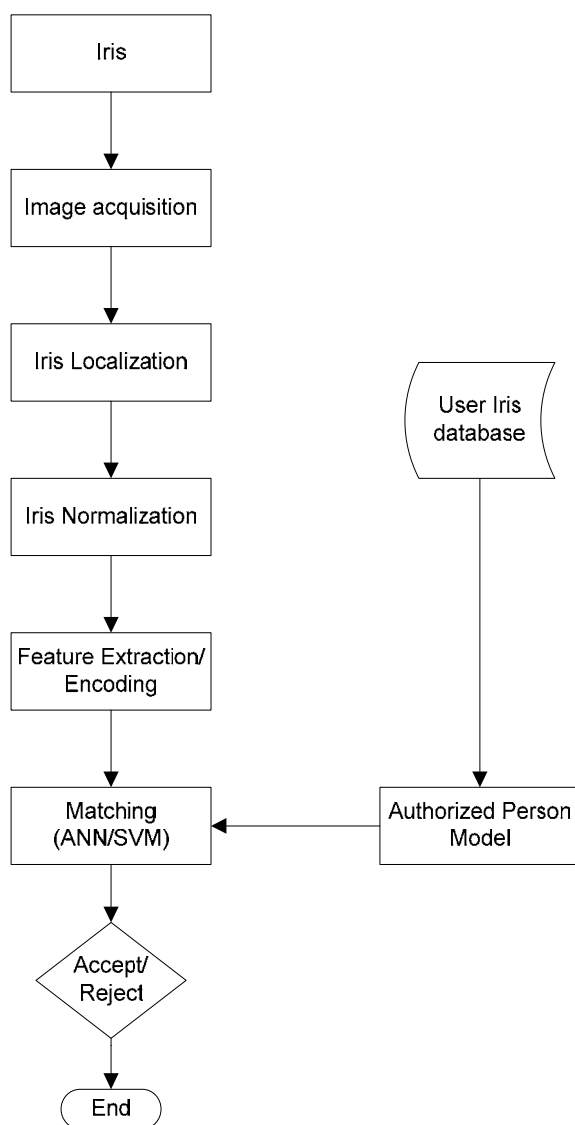


Fig. 14. Basic structure of iris-based verification system.

The first stage of the iris recognition process is to isolate the actual iris region in a digital image so as to localize an acquired image that corresponds to the iris.

Figure 14 shows the complete diagram of iris authentication system. ANN and SVM techniques are used for pattern matching and classification in order to recognize the authorized user.

#### **4.1 ANN based IRIS model**

Multilayer Feedforward Network (MFN) architecture is also used in developing the iris - based model. The process of developing iris model through extracted features using MFN is as follows:

1. Acquisition of iris image which is then analysed to obtain the extracted features.
2. Determination of the structure of the ANN based on the inputs, hidden layers and activation function.
3. Training of the ANN using the input pattern and desired output.
4. Validation of the iris model.

Each authorized user has a unique ANN-based model. The initial weights and biases of the MFN are randomly selected. The target of the MSE is set at first as  $10^{-7}$  and log-sigmoid is used as the activation function for all layers. The ANN is trained using backpropagation learning algorithm with a learning constant of 0.05.

#### **4.2 SVM based iris model**

In developing user model based on iris code, a SVM with polynomial kernel function of order eight (8) is used. Each authorized user has his/her individual SVM-based model characterized by a set of support vectors. By using quadratic programming in the MATLAB environment, appropriate support vectors are determined. The penalty term C was set at  $10^{15}$  to anticipate misclassified data.

#### **4.3 Iris based authentication system**

Figure 15 shows the flow chart of the proposed system. When trying to access the system the user is required to enter his/her ID to associate it with his/her iris image being acquired. The captured iris image is then processed and compared with the person model to verify his/her identity claim. The iris testing phase has a decision process in which the system decides whether the extracted features from the given iris image matches with the model of the claimed identity.

In order to give access to a user or reject an intruder, a threshold is set. If the degree of similarity between a given iris image and its model is greater than a given threshold, then the user will access the system, otherwise the user will be rejected.

### **5. Integrated biometric system**

In order to reinforce the user authentication system, PBAS has been integrated with iris (keystroke and iris features) to improve the FAR and FRR rates. Figure 16 shows how the system is integrated.

The algorithm for the system is similar to what was discussed earlier. In the validation mode, a user is given a prompt to enter his/her ID together with password and iris image to validate whether there is an existing profile in the database or not. After the user enters the

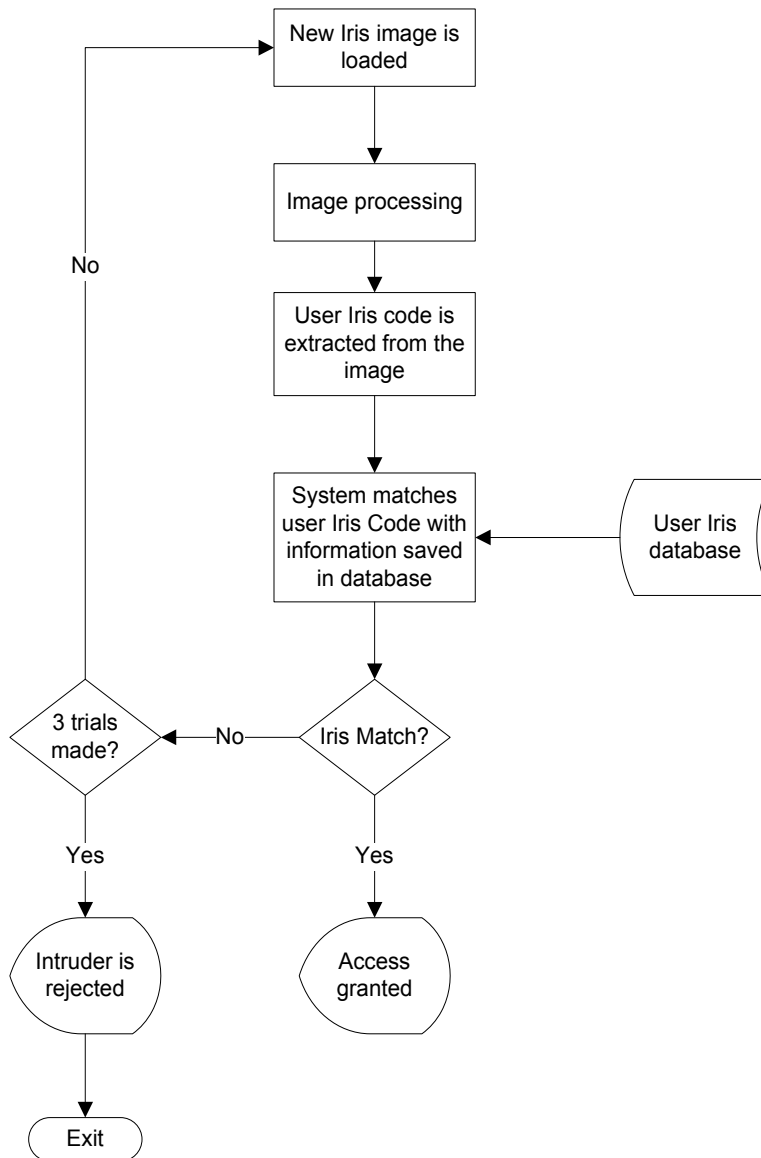


Fig. 15. Flowchart of iris-based authentication system.

password, maximum pressure and latency would be calculated and transferred to the classifiers (ANN, ANFIS and SVM). At the same time, an iris code is extracted from acquired image and transferred to the classifiers (ANN and SVM). The classifiers use MATLAB based algorithms to process the maximum pressure, latency and iris code and verify its compatibility (or level of matching) with the model in the database. If both keystroke template and iris of the user are nearly identical with the classifier models, then the system grants access to the user, otherwise access is denied after three (3) trials.

An experiment has been conducted to evaluate the effectiveness of the proposed system. The databases containing dynamic keystrokes and iris images were used in this experiment.

The dynamic keystroke data were collected from PBAS and stored in database 1. On the other hand, database 2 contains the iris images acquired from the Chinese Academy of Sciences–Institute of Automation (CASIA) eye image database.

A group of seven (7) persons was used for this experiment. Five (5) persons were considered as authorized users while the other two users were considered to be impostors.

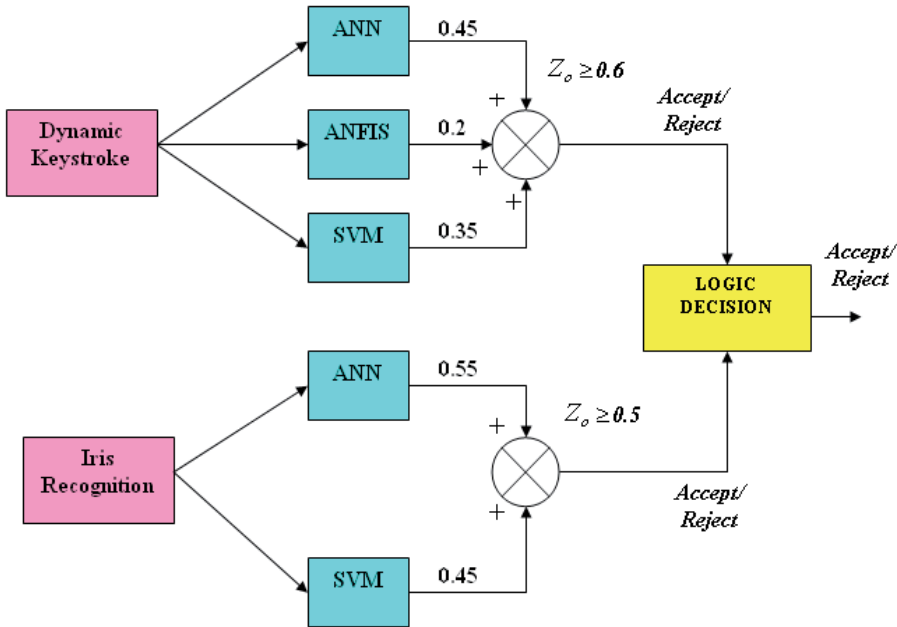


Fig. 16. Architecture proposed on integrating biometric system: first trial.

### 5.1 Training and testing performances of the integrated biometric system

In the first trial experiment, shown in Figure 16, the weight coefficients are assigned to each classifier experimentally by applying nonlinear programming. The ANN, ANFIS and SVM classifiers operate in such a way that every weight coefficient will multiply the classifier output to produce the best single classifier output which can be expressed as

$$\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \geq Z_o \quad (21)$$

where

$$\sum_{i=1}^3 \alpha_i = 1$$

The outputs of the classifiers are compared with their respective threshold values. If the result is greater than the threshold value, then the decision unit would accept the user, otherwise the user would be denied access. A logic AND operator is used for the final decision, thus, the system only gives access to the user when “accept” condition for both keystroke and iris recognition are met.

Table 12 and Table 13 illustrate the training and verification results for the users. The weight coefficients and thresholds for the system are shown in Figure 16. Results show that all users

had perfect classification rates and the average training time for the system was less than 15 seconds.

Table 13 shows excellent FAR results for both close and open set. From security point of view, one can say that the system is fully protected from impostor attacks. However, the FRR values are relatively high with an average of 39.6 which would not be suitable for robust systems.

In order to reduce the FRR values, a second experiment was conducted by changing the values of  $\alpha_1$  and  $\alpha_3$  to 0.4 and 0.4 respectively and the threshold value for dynamics keystroke was changed to 0.7. All other values remain unchanged as shown in Figure 16.

The training performance based on the second trial is represented in Table 14. An average training time of less than 12 seconds is obtained in this experiment. Furthermore, there is no error in classifying the users as the results exhibit perfect classification for all people.

| Authorized User | Training Time(sec) | Classification Rate (%) |
|-----------------|--------------------|-------------------------|
| User 1          | 9.250              | 100                     |
| User 2          | 7.781              | 100                     |
| User 3          | 13.07              | 100                     |
| User 4          | 17.00              | 100                     |
| User 5          | 22.48              | 100                     |
| <i>Average</i>  | 13.92              | 100                     |

Table 12. Training performance based on the first trial.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 33      | 0                 | 0                |
| User 2          | 66      | 0                 | 0                |
| User 3          | 66      | 0                 | 0                |
| User 4          | 33      | 0                 | 0                |
| User 5          | 0       | 0                 | 0                |
| <i>Average</i>  | 39.6    | 0                 | 0                |

Table 13. Testing performance based on the first trial.

| Authorized User | Training Time(sec) | Classification Rate (%) |
|-----------------|--------------------|-------------------------|
| User 1          | 13.94              | 100                     |
| User 2          | 4.625              | 100                     |
| User 3          | 8.906              | 100                     |
| User 4          | 12.56              | 100                     |
| User 5          | 19.57              | 100                     |
| <i>Average</i>  | 11.92              | 100                     |

Table 14. Training performance based on the second trial.

Furthermore, Table 15 shows that FAR results are excellent for both close and open set. The FRR average value has been reduced to 19.8; nevertheless, it is still relatively high and should be reduced further.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 33      | 0                 | 0                |
| User 2          | 66      | 0                 | 0                |
| User 3          | 0       | 0                 | 0                |
| User 4          | 0       | 0                 | 0                |
| User 5          | 0       | 0                 | 0                |
| <i>Average</i>  | 19.8    | 0                 | 0                |

Table 15. Testing performance based on the second trial.

In the third trial all the weight coefficients were fixed, while the threshold value for iris recognition was changed to 0.45 whilst keeping other coefficients used in the second experiment constant.

Table 16 shows the perfect classification for all users in the third trial experiment. The proposed system produced no error in the classification of the individual users and its training time was about 17 seconds.

| Authorized User | Training Time(sec) | Classification Rate (%) |
|-----------------|--------------------|-------------------------|
| User 1          | 20.45              | 100                     |
| User 2          | 15.62              | 100                     |
| User 3          | 19.44              | 100                     |
| User 4          | 21.42              | 100                     |
| User 5          | 7.718              | 100                     |
| <i>Average</i>  | 16.93              | 100                     |

Table 16. Training performance based on third trial.

This experiment produces the best performance as illustrated in Table 17. The proposed integrated biometric system has been able to produce the best results for both FRR and FAR.

| Authorized User | FRR (%) | FAR (%) Close Set | FAR (%) Open Set |
|-----------------|---------|-------------------|------------------|
| User 1          | 0       | 0                 | 0                |
| User 2          | 0       | 0                 | 0                |
| User 3          | 0       | 0                 | 0                |
| User 4          | 0       | 0                 | 0                |
| User 5          | 0       | 0                 | 0                |
| <i>Average</i>  | 0       | 0                 | 0                |

Table 17. Testing performance based on third trial.



Consequently, one can conclude that with the proper choice of threshold values the performance of the system can be optimized to achieve the best FRR and FAR rates.

The performance of the proposed system, especially with larger number of participants and sample size is still under investigation.

## 6. Conclusion and recommendation

Several commercial keystroke biometric algorithms rely on the use of the time-information (latency) alone and thus utilize only one information aspect of the keystroke action. It neglects the force applied during the keystroke action. PBAS has successfully acquired both time-frame and applied force information to develop varied biometric systems. Experiments have proved that force is highly distinctive to users typing keystroke. While some users may have similar latency profiles, however their keystroke pressure templates are easily discriminated.

The combination of AR and latency classifiers allows for an increase in latency threshold value to decrease the FRR rates. This increase does not have great effect on the FAR rates as the AR classifier would reject intruders based on their pressure templates, but it will make the system more user friendly without compromising the security.

By utilizing maximum pressure for keystroke template it is possible to merge the force and latency as a single feature for users. Multiple classifiers consisting of ANN, ANFIS and SVM are employed to authenticate personal identity based on extracted features. The combined features (maximum pressure and latency) are considerably more effective in verifying the authorized users. Results of experiments have shown that the use of combined features are more effective as compared to that of individual feature such as maximum pressure or latency as it gives better FRR and FAR for both open and close set conditions.

A multimodal biometric system uses multiple applications to capture different types of biometrics. This allows the integration of two or more types of biometric recognition and verification systems in order to meet stringent performance requirements. Furthermore, multimodal biometric systems also provide anti-spoofing measures by making it difficult for impostor to spoof multiple biometrics simultaneously. In this chapter, an integrated biometric system, which integrates keystroke pressure-based typing biometric system and iris verification has been presented. The proposed integrated biometric system is designed to operate in parallel mode and the integration process is made at the decision level. Integrated biometric fusion at the decision-level has shown to have good flexibility, high level of robustness and a great potential for improved accuracy.

## 7. References

- Akaike, H. (1970). Statistical predictor identification, *Annals of the Institute of Statistical Mathematics*, Vol. 22, No. 2, (1970) pp. 203-217, ISSN 0020-3157.
- Araujo, L.C.F.; Sucupira, L.H.R.; Lizarraga, M.G.; Ling, L.L. & Yabu-Uti, J.B.T. (2005). User Authentication through Typing Biometrics Features. *IEEE Transactions on Signal Processing*. Vol. 53, No. 2, (February 2005), pp. 851 – 855, ISSN1053-587X.
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, (1998), pp. 121-167, ISSN 13845810.

- Daugman, J. (2003). The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, Vol. 3, No. 2, (February 2003), pp. 279-291, ISSN 00313203.
- De Ru, W. & Eloff, J. (1997). Enhanced Password Authentication through Fuzzy Logic, *Intelligent Systems and Their Applications*, Vol. 12, No. 6, (November 1997), pp.38-45, ISSN 0885-9000.
- Eltahir, W.S.; Salami, M.J.E.; Ismail A.F. & Lai, W.K. (2008). Design and Evaluation of a Pressure-Based Typing Biometric Authentication System, *EURASIP Journal on Information Security*, Vol. 2008, No. 1, (2008), doi:10.1155/2008/345047, Available from <http://www.hindawi.com/journals/is/2008/345047/cta/>
- Eltahir, W.S.; Salami, M.J.E.; Ismail A.F. & Lai, W.K. (2004). Dynamic Keystroke Analysis Using AR Model, *Proceedings of the International Conference on Industrial Technology (IEEE-ICIT04)*, pp. 1555 - 1560, ISBN 0-7803-8662-0, Hammamet, Tunisia, 08-10 December, 2004.
- Eltahir, W.E.; Lai, W.K.; Salami, M.J.E. & Ismail, A.F. (2003). Hardware Design, Development and Evaluation of a Pressure-based typing Biometric Authentication System. *Proceedings of the Eighth Australian and New Zealand Intelligent Information Systems Conference (ANZIIS)*, pp. 49-54, ISBN1-74107-043-0, Sydney, Australia, December 10-12, 2003.
- Jang, J. S. R., Sun, C. T. & Mizutani, E. (1996). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice Hall, ISBN 978-0132610667, USA.
- Joyce, R. & Gupta, G. (1990). Identity Authentication Based On Keystroke Latencies. *Communications of the ACM*, Vol. 33, No. 2, (1990), pp. 168-176, ISSN 0001-0782
- Lin, C. T. & George, C. S. (1995). *Neural Fuzzy System: A Neuro-Fuzzy Synergism to Intelligent Systems*. Prentice-Hall, ISBN 978-0132351690, USA.
- Manson, H. H. (1996). *Statistical Digital Signal Processing and Modelling*. John Wiley & Sons, ISBN 0-471-59431-8, USA.
- Masek, L. (2003). Recognition of Human Iris Patterns for Biometric Identification, In: *Measurement*, 01.11.2009, Available from <http://www.csse.uwa.edu.au/~pk/studentprojects/libor/index.html>.
- Monrose, F. & Rubin, A. (1997). Authentication via keystroke dynamics. *Proceedings of 4<sup>th</sup> ACM Conference on Computer and Communication security*, pp. 48-56, ISBN 0-89791-912-2, Zurich, Switzerland, April 01 - 04, 1997.
- Ord, T. & Furnell, S. (2000). User Authentication for Keypad-Based Devices Using Keystroke Analysis. *Proceedings of the Second International Network Conference (INC)*: pp. 263-272, ISBN 1 84102 066 4, Plymouth, UK, 03-06 July, 2000.
- Shiavi, R. (1991). *Introduction to Applied Statistical Signal Analysis*. Aksen associates incorporated publishers, ISBN0256088624, Homewood, Illinois, USA.
- Wildes, R.P. (1997). Iris Recognition: An Emerging Biometric Technology. *Proceedings of the IEEE*, Vol. 85, No. 9, (August 2002), pp. 1348-1363, ISSN 0018-9219.



*Edited by Zahid Riaz*

Biometric authentication has been widely used for access control and security systems over the past few years. The purpose of this book is to provide the readers with life cycle of different biometric authentication systems from their design and development to qualification and final application. The major systems discussed in this book include fingerprint identification, face recognition, iris segmentation and classification, signature verification and other miscellaneous systems which describe management policies of biometrics, reliability measures, pressure based typing and signature verification, bio-chemical systems and behavioral characteristics. In summary, this book provides the students and the researchers with different approaches to develop biometric authentication systems and at the same time includes state-of-the-art approaches in their design and development. The approaches have been thoroughly tested on standard databases and in real world applications.

Photo by Peshkova / iStock

**IntechOpen**

