



IntechOpen

Biometrics

Unique and Diverse Applications in Nature,
Science, and Technology

Edited by Midori Albert



BIOMETRICS - UNIQUE AND DIVERSE APPLICATIONS IN NATURE, SCIENCE, AND TECHNOLOGY

Edited by **Midori Albert**

Biometrics - Unique and Diverse Applications in Nature, Science, and Technology

<http://dx.doi.org/10.5772/645>

Edited by Midori Albert

Contributors

M. Angeles Hernández, Raúl Martín, Tomás Santamaría, Francisco Campos, Aparecido Nilceu Marana, João Paulo Papa, Michael Hofer, Denise Tostes Oliveira, Elizabeth Bonsaglia Barboza, Bon K. Sy, Arun P. Kumara Krishnan, Adam Ramirez, José Alberto Hernández, Jingyan Wang, Yongping Li, Yuefeng Huang, Ying Zhang, Midori Albert, Karl Ricanek, Amrutha Sethuram, Kensuke Baba, Ali Ismail Awad, Lik-Kwan Shark, Bogdan J. Matuszewski, Wei Quan, Huiqi Lu, Chris R. Chatwin, Rupert C.D. Young

© 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).

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 those 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

Biometrics - Unique and Diverse Applications in Nature, Science, and Technology

Edited by Midori Albert

p. cm.

ISBN 978-953-307-187-9

eBook (PDF) ISBN 978-953-51-5525-6

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

4,000+

Open access books available

116,000+

International authors and editors

120M+

Downloads

151

Countries delivered to

Our authors are among the
Top 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

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

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

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



Meet the editor



Midori Albert is a professor and forensic anthropologist in the Department of Anthropology at the University of North Carolina Wilmington (UNCW), where she joined the faculty in 1995. She received her doctorate degree in anthropology from the University of Colorado at Boulder (CU Boulder, 1995); her master's degree in anthropology and bachelor's degree in psychology are from the University of Florida (UF, 1993 and 1990). Dr. Albert's educational background involves human skeletal identification methods and techniques as well as an understanding and exploration of normal human skeletal variability. These specializations led her to link with colleagues in computer science, statistics, and mathematics, where her primary research focus since 2003 has been on adult age-related craniofacial morphological changes and their effects on computer automated face recognition technologies. Together with her colleagues, Dr. Albert collaborates on research within the university's newly established Institute for Interdisciplinary Studies in Identity Sciences (I2SIS). Students often actively engage in research learning opportunities with Dr. Albert as interests in human biometrics and identity recognition continue to evolve.

Contents

Preface XI

- Chapter 1 **Usefulness of Biometrics to Analyse Some Ecological Features of Birds 1**
M. Ángeles Hernández, Francisco Campos,
Raúl Martín and Tomás Santamaría
- Chapter 2 **Toward An Efficient Fingerprint Classification 23**
Ali Ismail Awad and Kensuke Baba
- Chapter 3 **Dental Biometrics for Human Identification 41**
Aparecido Nilceu Marana, Elizabeth B. Barboza,
João Paulo Papa, Michael Hofer and Denise Tostes Oliveira
- Chapter 4 **Facial Expression Recognition 57**
Bogdan J. Matuszewski, Wei Quan and Lik-Kwan Shark
- Chapter 5 **Implications of Adult Facial Aging on Biometrics 89**
Midori Albert, Amrutha Sethuram and Karl Ricanek
- Chapter 6 **Iris Recognition on Low Computational Power Mobile Devices 107**
Huiqi Lu, Chris R. Chatwin and Rupert C.D. Young
- Chapter 7 **Biometric Data Mining Applied to On-line Recognition Systems 129**
José Alberto Hernández-Aguilar, Crispin Zavala, Ocotlán Díaz,
Gennadiy Burlak, Alberto Ochoa and Julio César Ponce
- Chapter 8 **Parallel Secure Computation Scheme for Biometric Security and Privacy in Standard-Based BioAPI Framework 145**
Arun P. Kumara Krishan, Bon K. Sy and Adam Ramirez
- Chapter 9 **Implementing Multimodal Biometric Solutions in Embedded Systems 173**
Jingyan Wang, Yongping Li, Ying Zhang and Yuefeng Huang

Preface

From time immemorial, we as humans have been intrigued by, perplexed by, and entertained by observing and analyzing ourselves and the natural world around us. Science and technology have evolved to a point where we can empirically record a measure of a biological or behavioral feature and use it for recognizing patterns, trends, and or discrete phenomena, such as individuals—and this is what biometrics is all about. Understanding some of the ways in which we use biometrics and for what specific purposes is what this book is all about.

Throughout the nine chapters of this book, an international and interdisciplinary team of researchers will enable you to become familiar with the birth and growth of biometrics in ecology (Chapter 1) and how it has reproduced, in a sense, offspring that are quite diverse—from applications in individual human identification (Chapters 2 through 6) to technologic improvements in obtaining information or securing privacy in large bodies of data (Chapters 7 through 9). Whereas each chapter focuses on a definitive aspect of biometrics; the book as a whole is an amalgamation of examples of state of the art research within the biometrics paradigm.

In Chapter 1 we discover what the first biometrics studies were, and how biometrics works in nature—how we gather information on biological species, such as ecology, sex differences, seasonality, reproduction and more.

Shifting the focus of biometrics towards humans, particularly human identification, Chapter 2 provides some background on fingerprint classification and explains a robust fingerprint classification algorithm—how patterns are determined and classified. The authors share with us their results of a performance evaluation as well.

Chapter 3 explores human identification via the dentition. After a brief history of the use of dental features in human identification, we can see the value of a computer automated approach to dental recognition. Imagine a networked database of all people's dental records. Imagine you can query this enormous database with information about an individual's teeth. Imagine the computer finds matches of a reasonable number that can then be analyzed and your individual is recognized or identified.

Further on the topic of human identification is Chapter 4 where we learn about research on computer automated facial expression recognition. Given that facial expressions derive from emotions and cognition and manifest our affective states, we know that we can often understand how other people feel by observing these facial expressions. However, what if we could not be there to monitor someone at critical times—someone

who is ill, unable to speak, and in pain, for example, as the authors suggest. This chapter describes existing methods and presents unique work on the use of 3D facial data for the automatic recognition of facial expressions.

More generally, the issue of computer automated facial recognition technologies for forensic purposes is raised in Chapter 5. Herein we learn about the challenges normal aging presents to computer face recognition. As faces age, they change. How much do they change? How does this affect computer face recognition? We see that after many years a person's face may be amazingly different. However, what if our faces change slightly on a daily basis? Would this subtle change affect computer face recognition? Chapter 5 explores face aging in face recognition, and introduces an experiment on face changes in one person in one day.

Aside from recognizing entire faces, much work in biometrics may be found in exploring identity markers based on single features, such as the iris of the eye. Chapter 6 reviews the latest developments in iris recognition used on handheld iris recognition devices, both for government or private sector endeavors. Through a mobile biometric identification system (MBI system) case study, we learn about hardware specifics, iris recognition algorithms, and system performance. Current solutions and the step-by-step format of this chapter are sure to captivate interest.

From ecology to human identification, it can be seen that biometrics has both breadth and depth of utility. And with all the biometric data collected in large databases, one issue that has been raised with regard to the use of these data is the issue of privacy.

Chapter 7 addresses the issue of privacy through an explanation of biometric data-mining. Biometrics systems recognize us in two ways—physically (e.g., fingerprints) or behaviorally (e.g., voice); and biometric data-mining merges these aspects of recognition such that we may be identified by how we use computers, for example keystroke patterns, mouse movements, and online behaviors. Detailed examples and intriguing descriptions of biometric data-mining and its implications are presented in this chapter.

Continuing in the theme of privacy issues, Chapter 8 introduces us to the BioAPI 2.0—a new industry standard in biometric systems that allows for interoperability while maintaining security and privacy. If one biometric can serve as a security measure (for example, the iris of the eye is read rather than a key being used to unlock a door), then security may be increased if more than one biometric system may be used. However, because biometric systems are composed of various segments and those segments do various things that are often isolated—they are non-interchangeable between systems. Vendors of biometric systems are therefore limited; and interfacing is compromised. The BioAPI 2.0 is explained in this chapter as a means to creating an interface that allows different biometric systems to work together. The authors provide an excellent background and detailed information on the BioAPI 2.0.

Also working on improving the technology and usability of biometric systems are the authors of Chapter 9 who research multi-modal biometrics solutions for embedded systems. Embedded systems that collect, store, modify, and retrieve data, such as personal information, are often at risk. In this chapter, the researchers discuss the development of multi-modal biometric systems as opposed to less robust uni-modal

systems; and, they tell us how to design a high performance embedded multimodal biometrics system—one solution to the privacy issue.

As can be seen, “Biometrics: Unique and Diverse Applications in Nature, Science, and Technology” provides a unique sampling of the diverse ways in which biometrics is integrated into our lives and our technology. I hope you will enjoy learning or reviewing the biometric applications presented in this collection of research studies, a collection that at this moment is leading a new frontier.

February 18, 2011

Midori Albert
Wilmington,
North Carolina

Usefulness of Biometrics to Analyse Some Ecological Features of Birds

M. Ángeles Hernández¹, Francisco Campos²,
Raúl Martín³ and Tomás Santamaría⁴

¹University of Navarra

²European University Miguel de Cervantes

³University of Castilla – La Mancha

⁴Catholic University of Avila
Spain

1. Introduction

Morphometric measurements of birds are the first data to be really considered as biometric in this discipline. Baldwin et al. (1931) depicted and explained in detail the external measurements used in ornithology. Currently, many of these measurements have been forgotten or are rarely used both in books dedicated to bird taxonomy (Cramp & Simmons, 1977) and in field guides on different geographical areas or on large bird groups such as shorebirds, raptors, passerines, etc. (Svensson, 1992; Baker, 1993).

Old biometric analyses used measurements performed on birds preserved in natural history museums. An appropriate representation of specimens is generally found in these museums, both in numbers (which allows for a large sample size) and in geographic origin (which enables the establishment of comparisons between birds of different areas) (Jenni & Winkler, 1989; Winker, 1993, 1996).

Body mass was another one of the data used in the initial biometric analyses. Its objective was to determine the presence of daily or seasonal variations, or variations linked to other specific periods: breeding, rearing and migration.

The next step was the establishment of a link between metric differences and the sex of birds. In some species, these differences were very visible and therefore statistical analyses were not required to support the distinction between males and females as in some raptors such as the Merlin *Falco columbarius* (Newton, 1979; Wiklund, 1990), owls and skuas (Andersson & Norberg, 1981). Similarly, marked biometric differences between bird populations of the same species found in different geographical areas were recorded (Svensson, 1992). This resulted in the identification of subspecies when these populations were geographically isolated, not sharing potential hybridization areas. Thus, for example, 10 subspecies of the Bluethroat *Luscinia svecica* have been identified throughout Europe, Asia and Alaska (Collar, 2005), a further 10 subspecies of Southern grey shrike *Lanius meridionalis* have been identified (Lefranc & Worfolk, 1997; Klassert et al., 2007), etc.

Substantial databases were created as a result of the routine collection of a minimum number of measurements when a bird was captured, this information being used for specific purposes. Possibly, the existence of these data and the ability of observation lead researchers

to conduct comparisons of measurements taken in each species, taking into account different parameters, geographical situations, habitats, etc. Thus, the first biometric studies were initiated and now days they add up to many studies already published.

Scientific articles which include morphometric measurements help to provide answers to theoretical and applied bird ecology issues (Morgan, 2004). In this chapter, we discuss how some of these issues may be analysed through biometry and which precautions need to be taken in order to avoid wrong conclusions.

2. What measures should be taken into account

Usually, wing length (maximum chord), third primary (counted in ascending order, or eighth primary counting in a descending order), tail, beak and tarsus are measured in every specimen of passerines. Most researchers follow Svensson's criteria (1992) when taking these measurements. In raptors, the measurements must be taken with a co-worker, and the following measurements are required to establish body size: a) four measurements of different parts of the right leg: tarsus-metatarsus, tibia-tarsus, middle toe and foot span, all to the nearest ± 0.1 mm; b) three measurements which include areas covered with feathers, hand-wing, the length of the right wing and the tail; c) body length (from tip of central tail feathers to crown of the bird lying relaxed on the ruler), with an accuracy of ± 1 mm. Birds moulting their longest primaries and/or the tail central feathers are excluded from the studies. Similarly, yearlings are excluded because their feathers are shorter than the adults' (Wiklund, 1996).

Body mass is usually measured with an accuracy of ± 0.1 g in small birds and ± 1 g in large birds. Two main methods are used: a) if the bird is captured and handled directly, the bird is bagged and weighed on a scale. This method can stress the bird and cause body mass reduction in a short time (Rands & Cuthill, 2001), a fact that should be taken into account when analysing data. b) When we do not wish to capture the bird, attracting it to a place situated on a scale which automatically records mass variation will suffice. This procedure has been employed, for example, to analyze body mass variation in breeding birds when they regularly visit the nest (Moreno, 1989; Szép et al., 1995), and amount of food brought to the nestlings (Reid et al., 1999), etc.

Two important issues should be taken into account in the data analysis: body mass variation with time of day and pseudoreplication of data which could distort the conclusions obtained (see review by Rands et al., 2006). Body mass shows circadian fluctuations depending on variables such as, for example, time elapsed between the time of feeding and the activity. Furthermore, there is seasonal variability depending on sex, which is more pronounced in females during the breeding period, especially in raptors (Newton, 1986).

3. Some problems with data

The quality of the measurements is essential in any scientific field but it becomes especially interesting in the study of birds. The data obtained from handling specimens (e.g., during ringing) will be subsequently analysed by other researchers, and therefore mistakes made during data collection may invalidate the rest of the work. Ensuring the quality of the measurement procedures is an essential aspect of the research.

Mentioning the quality of the measurements is equivalent to mentioning the extent of the errors. In general, the errors that can be made in this type of studies are of two types: systematic errors (bias) and random error (sampling error).

Morgan (2004) listed seven potential errors that affect the correct collection of measurements: 1) systematic *vs* random error, caused by the person taking the measurements or by the tools used; 2) errors in practice, caused by fortuitous agents at the time of taking the measurement, such as instability when measuring weight caused by the effect of the wind; 3) management error, when the measurements used for a study come from other researchers without having previously standardized the measuring protocols; 4) error from measuring devices, inaccuracy when reading the measurements of non digital equipments, as they don't always reach exactly the marks on the scales and an estimation has to be made, and each researcher can do it differently; 5) error in continuous variables, generated when the values of a continuous variable are rounded off; it must be done in accordance with the unit of measurement, as an error of 0.5 mm is not the same when measuring a passerine wing than a raptor wing; 6) errors arising from rounding off, both in continuous variables and in statistical tests in which decimal values are often rounded up or down; 7) error compounding in indices, occurring when ratios, indices, etc., are calculated by multiplying or dividing the original measurements.

The equipment used for data collection must be appropriate and must have been designed for that purpose, and the person collecting data needs to have a basic knowledge of statistical processing.

Some aspects that must be taken into account regarding the individuals who take the measurements, the repeatability of measurements taken on museum skins and on live birds, and the shrinkage effect of museum skins are discussed below.

- a. The observers must be qualified for the collection of measurements as they are not the same in museum birds than in live birds and in both cases, experience and practice are required. For measurements taken on museum specimens, data to the nearest ± 0.01 mm are commonly found. For live birds, on the contrary, measurements with that accuracy are difficult to replicate, and it is therefore preferable to take measurements with an accuracy of ± 0.1 mm. To verify the error, a small sample (e.g., 10 individuals) may be taken and measurements may be repeated until appropriate handling with a minimum error is achieved.

Whenever possible, live bird measurements should be taken by several people in order to obtain a certain range of diversification. However, an objection to this practice is the stress caused to a bird when it is handled by two or more people. On the other hand, a measurement team system (3-4 people) allows for a greater precision. This way, 1) measurements are validated when the differences obtained by each person are verified and these differences are maintained; 2) turns are taken to make the measurements so that each bird is measured by a single person but every person measures a similar number of birds; 3) if the measurements taken by each person are taken separately, the differences between them could be calculated and taken into account at the time of data analysis. In other cases, the data used for studies may come from databases from ornithological organizations, in which the data have been taken by different people but following the same measuring protocol.

- b. Repeatability (known as intra-class correlation coefficient) is a statistical measurement which shows data consistency between repeated measurements of the same characteristic in a single individual. The value of the repeatability r is calculated using the formula $r = s^2_A / (s^2 + s^2_A)$, in which s^2_A is the value of the inter-group variance and s^2 is the value of the intra-group variance (Sokal & Rohlf, 1981). The Measurement Error (ME) which is the opposite value of repeatability and is defined as the phenotypic

proportion of a characteristic attributable to the error that may be made must also be taken into account. The value $r = 1$ is the maximum possible and shows that the measurement is completely consistent and repeatable. Measurements showing an r value below 0.70 may be considered as repeatable although, to be considered as reliable, values above 0.90 should be obtained (Harper, 1994). This calculation is important in order to ensure the accuracy of the conclusions when researchers are dependent on the collection of measurements as is the case here. Lessells & Boag (1987) indicated the existence of published and unpublished works in which repeatability had been wrongly calculated mainly as a result of applying, in the formula, the least squares values instead of the inter and intra-group variance component.

Kuczynski et al. (2003) conducted a study on Northern grey shrike *Lanius excubitor* which provided information on measurement repeatability between observers and on the differences between the measurements carried out on live birds and on birds kept in museums. Four measurements were taken (wing length, tarsus length, beak length and tail length) on 50 live specimens, their skins were prepared subsequently and the same measurements as those taken on the live birds were taken except for the tarsus because the fingers of a dissected bird's leg cannot be opened. Repeatability was calculated as intra-class correlation coefficients, and a difference in the repeatability of beak measurement was obtained. Similarly, Szulc (1964) studied three passerines (Siskin *Carduelis spinus*, Robin *Erithacus rubecula*, and Blue tit *Cyanistes caeruleus*) and found a greater variation between observers in bone measurements (beak and tarsus) than in feathers (wing and tail). Repeatability of beak length appears not to be very consistent, even when the observers are specialists in collecting these measurements, which may be due to the fact that the points from which to obtain the measurements are not well defined.

In order to avoid differences between observers, Kuczynski et al. (2003) suggested that the data should be taken by one person or by a specially trained team (see also Busse, 1983; Gosler et al., 1998). At the end of the study, Kuczynski et al. (2003) suggested the following recommendations for the collection of measurements from museum skins: 1) To define exactly the method to take the measurements required; 2) within a single study, the measurements should be taken by the same person; 3) the age of the specimen measured should be known and this datum should be included in the analysis as a covariant in order to avoid bias resulting from shrinkage.

On the other hand, Berthold & Friedrich (1979) compared two ways of measuring wing length, one based on the length of the maximum chord (Svensson, 1992) and the other based on the length of the third primary. The latter was obtained by inserting a pin mounted on a ruler between the second and third primaries near their bases, flattening the third primary and measuring it on ruler (see Bertold & Friedrich, 1979; Jenni & Winkler, 1989; Svensson, 1992). For this, the data from experienced and unexperienced observers obtained from the same 23 Tree sparrows *Passer montanus* were obtained. The mean values of the measurements taken for each one of the two groups of observers were significantly different, more so for tail length than for the length of the third primary. However, repeatability of wing length was lower than that of the third primary and therefore wing length appears to be more affected by experience or training. This is why standard ringing procedures have been regulated in England and Ireland for decades and strict training is required. As a result of this study, length of the third primary has been proposed in several countries as a measure, in passerines, which

is better than wing length to reflect body size. However, sexual dimorphism in the length of the third primary is perhaps less marked than in wing length and may therefore be a poorer measurement as sex discriminant (Gosler et al., 1995).

- c. Specimen museums shrink and are dry, and therefore the length of primary feathers (and of the wing in general) is affected (Jenni & Winkler, 1989). The study of Kuczynski et al. (2003) enabled to establish the potential error resulting from shrinkage. The mean shrinkage rate between observers was different for all the measurements except for the tarsus, reaching in some cases as much as 5%. This value is above the 1 - 4 % obtained in waders and passerines (Vepsäläinen, 1968; Knox, 1980; Bjordal, 1983), although the data from these authors were obtained from a small sample size and over a short period of time following skin preparation.

On the other hand, it is common for the development of bilateral traits (e.g. wing or tarsus lengths) not to be symmetrical which causes the issue of bilateral asymmetry, widely discussed for decades (Palmer & Strobeck, 1986). When this occurs, the issue to be resolved is which of the two tarsi or wings should be considered. In addition, it is possible that the way in which measurements are obtained by the researchers influences the values of the bilateral traits (Helm & Albretch, 2000).

4. Some applications of biometry in the study of birds

Biometry has been used to study many aspects of birds but in the present work, only four aspects will be discussed: 1) sex determination, 2) differences in size among populations, 3) wing morphology, d) body mass - body size relationship.

These sections are detailed below, including: a) the more appropriate statistical analysis in each case, b) what type of issues has the application of biometry intended to clarify, c) some specific examples of these applications.

4.1 Sex determination

Many bird species are monomorphic in their plumage and therefore sex cannot be determined through colour traits, etc. Others, on the contrary, show size differences, either of a certain trait, or of a set of traits. Thus, by determining which trait is different between sexes, it is possible to separate males from females. However, even when there are statistically significant differences in the mean values of each measurement, there is often an overlap in the measurement which renders this trait not valid as a sex differentiator (Ellrich et al., 2010).

Biometric characteristics have been used to determine the sex of birds as different as seabirds (Hansen et al., 2009), raptors (Bavoux et al., 2006), passerines (Svensson, 1992), etc. However, currently, sex can be determined using molecular techniques (Griffiths et al., 1998; Bantock et al., 2008) which are often more accurate than biometric calculations. Molecular techniques show certain disadvantages with regard to biometric techniques, among which: a) they require more time to obtain accurate results, b) they are more expensive as a well equipped laboratory is required and expensive chemical compounds are needed, c) these are invasive techniques that often require blood or feathers from live birds, although sometimes a small portion of the rachis of a feather is enough (Wang et al., 2006).

Molecular techniques have enabled to verify the validity of the biometric criteria previously used to determine sex. In general, a high level of accuracy is obtained (up to 99 %) in sex determination through biometric characteristics. However, there are also many occasions in which the error in the determination is greater than 10 % which can render the results as not

Species	Order	Sample size	Statistical analysis	Accuracy (%)	Source
Red-necked grebe <i>Podiceps grisegena</i>	Podicipediformes	76	DFA	79-80	Kloskowski et al. (2006)
Great cormorant <i>Phalacrocorax carbo</i>	Pelecaniformes	81	DFA	92-95	Liordos & Goutner (2008)
Imperial shag <i>Phalacrocorax atriceps</i>	Pelecaniformes	291	DFA	94-97	Svagej & Quintana (2007)
Australasian gannet <i>Morus serrator</i>	Pelecaniformes	201	Two-tailed binomial test	99.5	Daniel et al. (2007)
Great egret <i>Ardea alba</i>	Ciconiiformes	79	DFA	81-88	Herring et al. (2008)
Lesser flamingo <i>Phoenicopterus minor</i>	Phoenicopteriformes	154	DFA	93-98	Childress et al. (2005)
Griffon vulture <i>Gyps fulvus</i>	Falconiformes	97	DFA	94.1	Xirouchakis & Poulakakisi (2008)
Peregrine falcon <i>Falco peregrinus</i>	Falconiformes	131 nestlings	DFA	96.2	Hurley et al. (2007)
Red-tailed hawk <i>Buteo jamaicensis</i>	Falconiformes	69	DFA	91.3	Pitzer et al. (2008)
White-tailed Eagle <i>Haliaetus albicilla</i>	Falconiformes	211 nestlings	DFA	15-98 ^a	Helander et al. (2007)
Yellow-legged gull <i>Larus michahellis</i>	Charadriiformes	155	DFA	89.5	Arizaga et al. (2008)
Black-tailed godwits <i>Limosa limosa</i>	Charadriiformes	42	DFA	95.2	Gunnarsson et al. (2006)
Black tern <i>Chlidonias niger</i>	Charadriiformes	449	DFA	81.0	Shealer & Cleary (2007)
Redshank <i>Tringa totanus</i>	Charadriiformes	157	LDF	81	Ottwall & Gunnarsson (2007)
Blue-fronted Amazon <i>Amazona aestiva</i>	Psittaciformes	202	DFA	85	Berkunsky et al. (2009)
White-throated dipper <i>Cinclus cinclus</i>	Passeriformes	231	Logistic regression	98.7	Campos et al. (2005a)

Species	Order	Sample size	Statistical analysis	Accuracy (%)	Source
Dupont's lark <i>Chersophilus duponti</i>	Passeriformes	317	DFA	99.0	Vogeli et al. (2007)
Reed bunting <i>Emberiza schoeniclus</i>	Passeriformes	99	DFA	95	Belda et al. (2009)
Corn bunting <i>Miliaria calandra</i>	Passeriformes	103	LDF	96.1	Campos et al. (2005b)
Northern great shrike <i>Lanius excubitor</i>	Passeriformes	50	LDF	85.7	Brady et al. (2009)

Table 1. Accuracy obtained in sex determination through biometric characteristics in different species, based on a sample of 20 studies published since 2005. In all cases, sex determination was also performed through molecular techniques. The type of statistical method used is detailed. LDF: Linear Discriminant Function. DFA: Discriminant Function Analysis. ^a It varied according to sex and sampling zone.

statistically valid. In a sample of 20 published studies between 2005 and 2010 (Table 1) dealing with orders as diverse as Pelecaniformes and Passeriformes, eight (40.0 %) showed an accuracy below 90 %.

It has been suggested that, whenever possible, in some species it is more advantageous to sex the two members of a breeding pair through biometric characteristics (Fletcher & Hamer, 2003). However, in passerines, this is difficult given that the overlap of the measurements is high (Gutiérrez-Corchero et al., 2007a) and, at least in Southern Europe, there are few species showing sexual dimorphism in size such as Cetti's warbler *Cettia cetti* (Bibby & Thomas, 1984) and Corn bunting (Campos et al., 2005b).

Different multivariate statistical methods are used for the classification of birds by categories. One of the most rudimentary ways of doing this is by differentiating sex based on the study of morphological traits studied separately, using bimodal distributions for their classification (Catry et al., 2005). In practice, a single variable does not provide satisfactory results, the classification being improved by the combination of more variables. In addition, it is possible for differences between groups not to be found in any of the separate variables but in their combination. On the other hand, type I error increases when conducting repeated comparisons.

The most widely used method for the determination of sex is the Discriminant Analysis. With this method, classification functions are obtained which allow to assign sex and to evaluate the quality of the results. The classifications functions are linear functions of the morphological variables considered.

In order to validate the functions, the general way of proceeding is by dividing the sample in two groups: a) the training sample, made up of data for which the sex is unmistakably known, and b) the test sample, made up of the remaining observations. When the total sample is small, the Jackknife method is frequently used. This method is part of the so-called re-sampling methods which are characterized by the fact that they hardly require assumptions on the population model from which the sample is obtained. The idea of the method, developed in various steps, consists of leaving out one datum from the observers in

each step and in calculating the classification functions using the remaining data. Once obtained, the excluded observation is classified. An analogous procedure is followed by excluding a different observation in each step.

This technique has been used in some studies (Hermosell et al., 2007). When the conditions for the application of the discriminant analysis are not met (normal distribution and identical variances) the Logistic Discriminant is used (Ellrich et al., 2010) which is based on the logistic regression. In this analysis, sex probability is estimated through a combination of explanatory variables through a logistic response model.

An issue raised recently is the variation of sexual dimorphism within and between years (Van de Pol et al., 2009), at least for some species. These authors showed that in the Eurasian oystercatcher *Haematopus ostralegus* some biometric traits used for sex determination varied through time, thus invalidating the determination of sex through biometrics. A possible solution to this problem is to calibrate these traits by month, year and area, something which seems complicated for many species.

The knowledge of the sex of each specimen favours management techniques and species conservation (McGregor & Peake, 1998). On the other hand, the knowledge of the sex of the birds studied is often essential given that individual discrimination is required in order to analyze their behaviour, etc. Spatial sexual segregation has been analysed in many bird species, mainly during the breeding season (see review of Catry et al., 2005), but also at other seasons (Campos & Martín, 2010). This raises the issue of which sex is the dominant one in each species and which habitat requirements has each sex through the annual cycle.

Another important issue which requires the prior knowledge of the sex is differential migration, understood as the variation in the distance covered and in the wintering areas according to bird categories, mainly sex and age (Ketterson & Nolan, 1983). Sex differentiation through biometric traits is very useful in this field, as during the migratory route, researchers have to handle a large number of birds in a short time.

Finally, biometry applied to sex determination enables the determination of the sex ratio in adult birds, another field which remains poorly known in spite of having been analysed for several decades (Mayr, 1939). In wild populations, there is often a bias in the proportion of sexes (see review of Donald, 2007), often in favour of males, perhaps as a result of high female mortality. Obviously, this influences population processes and, therefore, conservation of bird species.

4.2 Differences in size among populations

It is common, within a single species, for the size of the populations to vary gradually throughout their geographical distribution. The analysis of biometric differences between populations enables to relate them to environmental parameters and infer possible causes that may explain them. The study of significant differences between populations is carried out through the analysis of variance (ANOVA) on the residuals obtained from the covariance analysis models (ANCOVA) adjusted for the variables of interest in each species, including location as a factor.

Body size variation in endothermic animals has been the subject of many studies. A hypothesis put forward to explain this variation is Bergmann's rule that establishes that body size varies inversely with ambient temperature, so that body size increases with latitude, and this has been supported by some studies (Yom-Tov, 1993; Ashton, 2002; Meiri & Dayan, 2003), but not by others (Yom-Tov & Yom-Tov, 2005; Rodríguez et al., 2008; etc.).

The global warming experienced over the last decades may influence the variation in body size of birds through changes in factors such as environmental variability (Jakober & Stauber, 2000). However, there are also studies that show the difficulty of finding a relationship between global warming and body size variation (Guillemain et al., 2005; Moreno-Rueda & Rivas, 2007).

On the other hand, body size seems to be influenced by other factors apart from climatic factors such as feeding. Thus, in Blackbird *Turdus merula*, availability of food has been linked to body size increase (Yom-Tov et al., 2006) and in some passerines early nutritional stress negatively affects skeletal size that carries over into adulthood (Searcy et al., 2004).

Sometimes, biometrics also help in the taxonomy of birds as it enables subspecies differentiation. Among the various examples that could be mentioned, those of the Bluethroat, in which the subspecies *Luscinia svecica namnetum* found in France differs by its small size from others which are geographically nearby (*L. s. cyanecula* and *L. s. azuricollis*, Eybert et al., 1999), and that of the Red knot *Calidris canutus* which shows size differences between the African subspecies (*C. c. canutus*) and the subspecies from Northern Europe (*C. c. islandica*, Summers et al., 2010) are particularly clear.

The conclusions reached by applying biometric characteristics are often confirmed through genetic analyses. Currently, a greater accuracy when defining different population taxonomic categories has been achieved through the analysis of genes present in mitochondrial and/or nuclear DNA. To continue with the example of the Bluethroat, molecular genetics have confirmed the validity of the subspecies *namnetum* and also of other subspecies which are biometrically similar between them (Johnsen et al., 2006). Similarly, in the Southern grey shrike, the biometric study suggested marked differences between the subspecies *meridionalis* from the Iberian Peninsula and the subspecies *koenigi* from the Canary islands (Gutiérrez-Corchero et al., 2007a,b). The same conclusion was reached through the analysis of mitochondrial DNA, both for the cytochrome b gene (Klassert et al., 2007) and for the tandem repeats of the Control Region (Hernández et al., 2010).

Size variation is seen more clearly in large geographical areas such as a continent like Europe (Dmitrenok et al., 2007). However, it is also possible to find, within a continent, biometric differences between populations of a single species in a more reduced geographical area such as, for example, the Iberian Peninsula and the British Isles (Wyllie & Newton, 1994). This is evidenced in the White-throated dipper. Throughout Europe, its size (measured by wing and tarsus length) increases towards Northern latitudes (Esteban et al., 2000), which is in agreement with Bergmann's rule mentioned previously. However, within the Iberian Peninsula, the White-throated dippers from the South are significantly greater than those from the North (Campos et al., 2005c), which contradicts Bergmann's rule and has been explained by the influence of local environmental conditions (Arizaga et al., 2009). Therefore, biometrics also help to raise new issues on bird ecology.

Through the statistical analysis of size differences in bird populations, other issues which affect threatened species requiring special attention may be resolved. This is the case of seabirds in Northern Europe affected by human activities and dying in fishing nets or oil spills (Barrett et al., 2008). For the Common guillemot *Uria aalge*, it has been possible to determine the area from which the affected specimens came from based on body measurements, whereas in other species, this method has shown little efficacy as a result of the lack of accuracy obtained in bird size differentiation between separate colonies.

4.3 Wing morphology

The study of wing shape has been conducted, mainly, in passerines who have ten primaries in each wing. The basic data that need to be obtained are the length of each one of these feathers (the so-called primary distances) although generally, the first primary is excluded because it is very short. Generally, the fourth and fifth primary are the longest (Fig. 1) and therefore, are the ones that will define whether total wing length is larger or smaller.

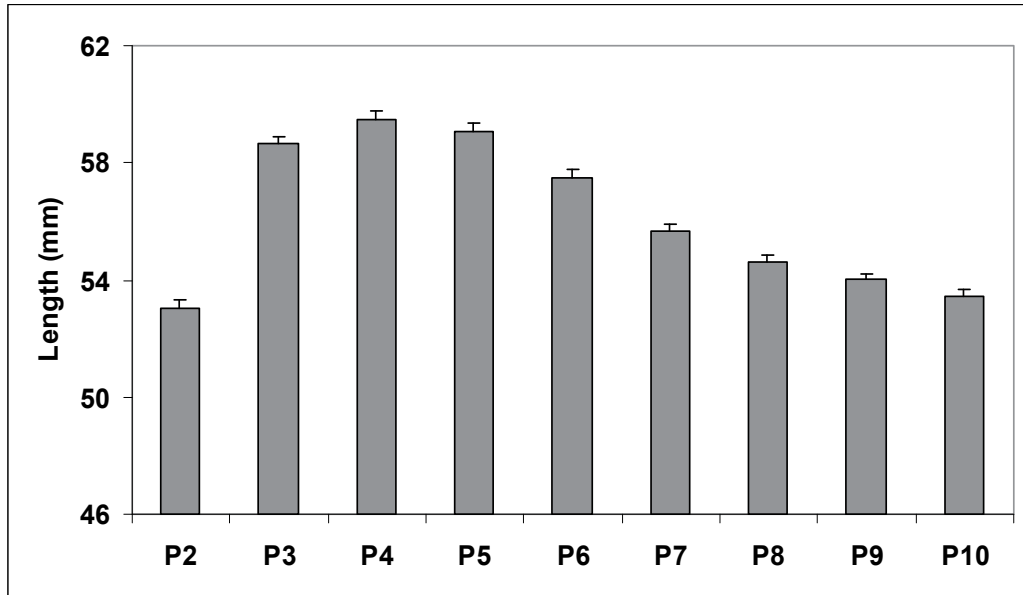


Fig. 1. Mean length (\pm SE) of primary feathers (P2-P10) in the wing of Bluethroat *Luscinia svecica azuricollis* in populations of central Spain.

In the majority of studies, wing morphology is characterised by the measurement of its pointedness and by its convexity, which are obtained from multivariate statistical methods. Thus, Principal Components Analysis (PCA) has been used in many studies to accurately describe the values of the primary distances using a smaller number of variables (Chandler & Mulvihill, 1988; Marchetti et al., 1995; Mönkkönen, 1995). Nevertheless, given the effect of size on wing shape, the direct application of PCA on primary distances would give wrong results. A first solution has been provided by Senar et al. (1994), who suggested a correction of the primary distances related to wing size and allometry. This method consists of multiplying the distance by a standard value of wing length divided by the specific value of bird length, raised to the power of the allometry coefficient of the distance that we wish to correct. PCA is applied on these corrected distances. The first component obtained is a good measure of wing pointedness. In spite of this correction, the results cannot be generalized either. Furthermore, this method presents statistical problems (Lockwood et al., 1998) and therefore a modification of the PCA was introduced providing a new valid method for the interpretation and characterization of the morphology within a single species and between different species (Lockwood et al., 1998). This new method is called Size-Constrained Component Analysis (SCCA). The first principal component (SCCA1) obtained through this method is a good index of wing pointedness.

Finally, the general linear model (MANCOVA) is used to study the presence of significant differences in morphological traits, controlling body size effect.

The design of bird wings is subject to various types of selective pressures. Generally, the wing is shorter and more rounded in juvenile birds than in adults (Pérez-Tris & Tellería, 2001). Longer and more pointed wings improve flight speed, whereas shorter and more rounded wings allow for better flight manoeuvrability.

Both aspects have important ecological consequences. The greater speed shortens the length of migratory journeys and therefore reduces energetic costs. Similarly, it also allows birds to reach stopover sites and wintering areas sooner, thus having an advantage over conspecifics in occupying the best sites (Bowlin, 2007, among others). On the other hand, short and rounded wings facilitate escape from predators as a result of enhanced manoeuvrability in flight, thus reducing mortality rate. Consequently, within a same species and also between species, there is a trade-off between both aspects of wing shape.

The length of primary feathers has also been analysed at the level of subspecies or migratory species populations that vary in the distance travelled in their migratory journeys. It is expected that populations travelling long distances will have longer primaries than those travelling shorter distances. This has been recorded in blackcaps (Fiedler, 2005) and bluethroats (Arizaga et al., 2006).

On the other hand, it has been detected in some non-migratory species, that some functional traits of the wings such as pointedness show covariation with weather conditions and the structure of the habitat they occupy (Vanhooydonck et al., 2009). This may be important to show the speed at which bird adaptations take place in changing local conditions.

All these questions require a knowledge of wing shape, for which biometrics are essential. Nevertheless, over the last years, it has become quite common to analyse the migratory behaviour of many bird species through stable hydrogen isotopes present in the feathers (Hobson, 2005). That way, the place of origin of the birds captured may be determined more accurately during their migratory flights or in the wintering areas. However, this method is laborious and expensive, and in addition it requires the extraction of one or several feathers from the bird. As for sex determination, when the handling of a large number of birds is required, the help of biometric analyses has shown to be important to resolve ecological issues related to migratory birds, given that it is simple, quick and its cost is low.

4.4 Body size – body mass relationship

Frequently, in birds, the greater the body size, the greater the body mass. The size of body mass may reflect the nutritional status of the bird (and therefore its fitness) and hence it is necessary to know its value.

Variation of birds' body condition is a subject of great interest in evolutionary ecology, and an accurate knowledge of it enables to confirm theories on bird adaptations to different environmental conditions. Thus, for example, the starvation-predation risk trade-off theory predicts that, in birds, body mass increases when starvation risk is greater and decreases when predation risk increases (McNamara & Houston, 1990; MacLeod et al., 2008). It is known that birds carry fewer fat reserves than the maximum possible (Witter & Cuthill, 1993), perhaps because body mass reduction favours greater flight manoeuvrability (Witter et al., 1994) and therefore, preys can escape more easily from predators, reducing thus predation risk (Lima, 1986; McNamara & Houston, 1990; Cresswell, 1998; MacLeod et al., 2005). For predatory birds, a greater manoeuvrability in flight may facilitate the capture of prey. On the other hand, body mass increase favours the resistance to adverse

environmental conditions and to food unpredictability, especially when birds must face a reduction in prey numbers.

There are many ways of analysing body condition in birds (refer to the review by Brown, 1996): size of subcutaneous fat reserves (Redfern et al., 2000), haematocrit (Cuervo et al., 2007), blood albumin level (Ardia, 2006), etc., but a simple one is the relationship between body size (generally expressed as wing or tarsus length) and body mass.

Body mass - body size relationship must be statistically analysed in order to ensure that the conclusions reached are accurate. Generally, a comparison of body mass in different groups is conducted, correcting the potential existing differences between them as a result of size that could affect the results. The statistical methods used for this are:

1. **Ratio Index.** It is the simplest and is calculated by dividing body mass by a measurement of size, for example tarsus or wing length, or by some power of it (Albrecht et al., 1993). This index has been criticized as a result of the problems it presents (Jacob et al., 1996). Atchely et al. (1976) showed that the ratio variables are skewed to the right, leptokurtic and that the non-normality is increased when the denominator coefficient is increased. Further, multivariate statistical procedures are affected when the analyses include ratios. And what is worst, it has been proved that in the scaling of data, ratios do not remove the effect of the scaling variables.
2. **Residual Index (RI).** This procedure is based on the least squares linear regression of body mass over size. Once the regression has been conducted, the residuals obtained are considered as a measure of body condition. In most studies, a single measure of body size is usually used to perform the regression. Given that the objective is to eliminate the effect of body size, a possibility for obtaining greater accuracy could be to perform Principal Component Analysis between different body measurements (e.g., tarsus or wing length) and conduct the regression with this new variable. In spite of being one of the methods which are used most frequently, the comparisons between RI values are not always valid. Furthermore, it has been shown that, often, the required hypotheses for the use of the least squares residuals are not met, and thus the errors of the test hypothesis increase. The use of the reduced major axis regression is therefore more appropriate (Green, 2001).
3. **Analysis of Covariance (ANCOVA).** This is a statistical control technique which is used to isolate the effect of a variable. It has the advantage of integrating in a single procedure the regression analysis and the analysis of variance procedures. Some authors recommend the use of this method exclusively in order to eliminate the effect of the value of body mass (García-Berthou, 2001).

An example based on the Southern grey shrike shows the different conclusions reached using one method or another. The Southern grey shrike is a medium size bird (25 cm) whose sexes remain separate during the non-breeding period: males remain in the breeding territories and females occupy distant areas (Campos & Martín, 2010). Campos et al. (2008) analysed the seasonal variation in the relationship body size - body mass in agricultural areas of Northern Spain, separating males and females. For this, they used the residual index RI calculated on the body mass - tarsus length regression. Their conclusion was that during the non-breeding season, the RI value did not vary significantly between autumn (October and November) and winter (December to February), and neither did they vary significantly between sexes or within each sex.

In the present chapter, unpublished data to date on this variation in Southern grey shrike in the centre of Spain where the environmental conditions in the study area are similar to those

in Campos et al. (2008) are presented. The ANCOVA procedure was used to compare the relationship between body mass and body size for different season, sex, age (yearling or adult) and habitat (irrigation crops *vs* non-irrigated crops), each of them with two levels. The prototypic analytic model for these outcomes was a four-way ANCOVA using tarsus (indicator of body size) as covariate. Main effects and interactions that were not significant at $P > 0.05$ were removed so that the best model could be fitted to the data. Additionally two-way ANCOVA models were used to assess differences in mean between groups examined by the independent variables of season, habitat and sex, respectively.

The full and adjusted 4-way ANCOVA models for the body mass were significant (Table 2). Because sex was not significant either in the main effect or in the interaction effect, this variable was removed in the following analysis. In the adjusted model the two significant effects found were the main effect of season ($P < 0.001$) and the habitat x season x age interaction ($P = 0.026$).

When examined by habitat, the significant effect in this analysis was the main effect of season for both, non-irrigated crops and irrigated crops ($P = 0.029$ and $P < 0.001$, respectively, Table 3).

	Full factorial model			Adjusted model		
	df	F	P	df	F	P
Main effects						
Habitat	1	1.717	0.192	1	0.476	0.491
Season	1	11.054	0.001	1	15.886	<0.001
Age	1	2.064	0.152	--	--	--
Sex	1	0.068	0.795	1	0.067	0.795
Two-way interactions						
Habitat x Season	1	0.890	0.347	1	2.673	0.104
Habitat x Sex	1	0.370	0.544	--	--	--
Habitat x Age	1	0.606	0.437	1	1.275	0.261
Season x Sex	1	0.142	0.707	--	--	--
Season x Age	1	0.529	0.468	1	0.026	0.872
Sex x Age	1	0.489	0.485	--	--	--
Three-way interactions						
Habitat x Season x Sex	1	0.024	0.876	--	--	--
Habitat x Season x Age	1	4.431	0.037	1	5.004	0.026
Habitat x Sex x Age	1	0.239	0.626	--	--	--
Season x Sex x Age	1	0.504	0.478	--	--	--
Four-way interactions						
Habitat x Season x Sex x Age	1	0.820	0.366	--	--	--
Covariates						
Tarsus	1	22.820	<0.001	1	28.511	<0.001
Overall model	17	6589.764	<0.001	9	11776.667	<0.001

Table 2. Full and adjusted ANCOVA models taking into account body mass, body size, habitat, season, age and sex. Statistically significant interactions are in bold. In the adjusted model, no significant main and interaction effects ($P > 0.05$) were removed if they were not included in higher order interactions. df: degree of freedom. P: probability.

	Non-irrigated crops			Irrigated crops		
	df	F	P	df	F	P
Main effects						
Season	1	4.903	0.029	1	22.989	<0.001
Age	1	0.170	0.681	1	1.282	0.260
Two-way interactions						
Season x Age	1	0.873	0.352	1	2.145	0.146
Covariates						
Tarsus	1	20.152	<0.001	1	21.835	<0.001
Overall model	5	10586.661	<0.001	5	12193.111	<0.001

Table 3. Variables that influence body mass in accordance to habitat. df: degree of freedom. P: probability.

Significant differences in body mass of the shrikes in autumn and in winter were recorded in both types of crops (Table 4 by rows).

	Autumn	Winter	Difference
Non-irrigated crops	64.65 ± 2.79 (11)	63.88 ± 2.51 (55)	0.77
Irrigated crops	65.35 ± 2.67 (32)	61.76 ± 3.23 (20)	3.59

Table 4. Mean value ± SD of body mass in adult Southern grey shrikes according to habitat (non-irrigated crops, irrigated crops) and season (autumn, winter). Sample size in brackets. The difference between mean values is adjusted.

Furthermore, non significant main effects or interactions were found for autumn, but habitat was significant ($P = 0.045$) for winter (Table 5).

	Autumn			Winter		
	df	F	P	df	F	P
Main effects						
Habitat	1	0.045	0.832	1	4.099	0.045
Age	1	0.494	0.484	1	0.564	0.454
Two-way interactions						
Habitat x Age	1	1.603	0.209	1	1.182	0.279
Covariates						
Tarsus	1	19.229	<0.001	1	23.105	<0.001
Overall model	5	10455.906	<0.001	5	12478.849	<0.001

Table 5. Variables that influence body mass according to season. df: degree of freedom. P: probability.

Indeed, the mean value of body mass of the shrikes was greater in the birds captured in non-irrigated crops than in birds captured in irrigated crops (Table 4, by rows).

Finally, when examined by age, the main effect of season was significant for both, yearling and adult ($P = 0.05$ and $P < 0.001$, respectively, Table 6) and the interaction habitat x season was significant for adult ($P = 0.029$). Body mass mean value (\pm SD) varied significantly

between autumn and winter both in young shrikes (64.1 ± 3.1 N = 48 *vs* 62.2 ± 3.5 , N = 57) and in adults (65.1 ± 2.6 , N = 43 *vs* 63.3 ± 2.8 , N = 75).

	Yearling			Adult		
	df	F	P	df	F	P
Main effects						
Habitat	1	1.608	0.208	1	0.883	0.349
Season	1	8.322	0.005	1	14.428	<0.001
Two-way interactions						
Habitat x Season	1	0.116	0.735	1	4.907	0.029
Covariates						
Tarsus	1	33.048	<0.001	1	10.384	0.002
Overall model	5	9506.195	<0.001	5	14127.559	<0.001

Table 6. Variables that influence body mass according to age. df: degree of freedom. P: probability.

The body mass - body size relationship has also been used to analyse other ecological issues in birds such as offspring quality. The measurements obtained on nestlings are a good example to analyse bilateral asymmetry and to verify which factors have an influence on body development of their bilateral traits. In this case, the issue to be resolved is which tarsus or wing must be related to body mass.

5. Further research

It can be inferred from the paragraphs detailed above that the following issues should be analysed in more detail in future research on:

- The way in which to increase the accuracy of measurements, unifying measurement criteria until their use becomes universal. This will enable the comparison of data obtained from different researchers and will facilitate reaching valid conclusions in studies based on animals from different geographic origin.
- Sex determination from biometric traits so that accuracy is close to 100%. That way bird sex may be determined through simple, quick and cheap methods. The importance of knowing the sex of a bird in a wide type of ecological studies has been shown above.
- Variation of biometric characteristics of birds according to their distribution area also requires further studies. Variables which allow to accurately determine, for example, where the birds captured in a study area come from are required. This aspect appears to be essential in order to analyse the behaviour of migratory species.
- Biometric traits - body size relationship until an almost perfect adjustment is obtained. New biometric characteristics which so far have been poorly explored and that would enable a more accurate statistical adjustment should be studied. An example of this has been the use of the third primary of the wing (see paragraph 3) instead of the total length (maximum chord).
- Similarly, the body size - body mass relationship should be further studied until the most suitable biometric characteristics are found in order to analyse them statistically. To that effect, it would be convenient to detail what type of mathematical analyses should be applied in each type of study, so that their use can be generalized and comparable results may be obtained in any part of the world.

6. Acknowledgements

Our thanks to Luis Corrales, Miguel Miranda and Elsa Santos for their help in the field work. Jesús López-Fidalgo made valuable comments on the manuscript. This work was partially founded by the Obra Social de Caja España and the Obra Social de Caja Ávila. Data of trapped birds were obtained under official permission of the Regional Government of Castilla-León.

7. References

- Albrecht, G.H.; Gelvin, B.R. & Hartman, S.E. (1993). Ratios as a size adjustment in morphometrics. *American Journal of Physical Anthropology*, 91, 441-468, ISSN 0002-9483
- Andersson, M. & Norberg, R.A. (1981). Evolution of reversed sexual size dimorphism and role partitioning among predatory birds, with a size scaling of flight performance. *Biological Journal of the Linnean Society* 15, 105-130, ISSN 0024-4066
- Ardia, D.R. (2006). Glycated hemoglobin and albumin reflect nestling growth and condition in American kestrels. *Comparative Biochemistry and Physiology Part A*, 143, 62-66, ISSN 1095-6433
- Arizaga, J.; Campos, F. & Alonso, D. (2006). Variations in wing morphology among subspecies might reflect different migration distances in Bluethroat. *Ornis Fennica*, 83, 162-169, ISSN 0030-5685
- Arizaga, J.; Aidalur, A., Herrero, A. & Galicia, D. (2008). Sex differentiation of Yellow-legged Gull (*Larus michahellis lusitanicus*): the use of biometrics, bill morphometrics and wing tip coloration. *Waterbirds*, 31, 211-219, ISSN 1524-4695
- Arizaga, J.; Hernández, M.A., Rivas, J. & Miranda, R. (2009). Biometrics of Iberian Dippers *Cinclus cinclus*: an exploration to understand causes explaining differences among populations. *Ardea*, 97, 23-30, ISSN 0373-2266
- Ashton, K.G. (2002). Patterns of within-species body size variation of birds: strong evidence for Bergmann's rule. *Global Ecology & Biogeography*, 11, 505-523, ISSN 1466-8238
- Atchley, W.R.; Gaskins, C.T. & Anderson, D. (1976). Statistical properties of ratios. I. Empirical results. *Systematic Zoology*, 25, 137-148, ISSN 0039-7989
- Baldwin, S.; Oberholser, H.C. & Worley, L.G. (1931). *Measurements of birds*. Scientific Publications Cleveland Museum Natural History, 2, 1-165, USA
- Baker, K. (1993). *Identification guide to European Non-Passerines*, British Trust for Ornithology, ISBN 0903793180, Thereford
- Bantock, T.M.; Prys-Jones, R.P. & Lee, P.L.M. (2008). New and improved molecular sexing methods for museum bird specimens. *Molecular Ecology Resources*, 8, 519-528, ISSN 1755-0998
- Barret, R.T.; Anker-Nilssen, T., Bakken, V., Strøm, H. Krasnov, Y. & Aarvak, T. (2008). Biometrics as a determinant of the origins of seabirds killed in oil spills and other incidents. *Bird Conservation International*, 18, 229-241, ISSN 0959-2709
- Bavoux, C.; Burneleau, G. & Bretagnolle, V. (2006). Gender determination in the western marsh harrier (*Circus aeruginosus*) using morphometrics and discriminant analysis. *Journal of Raptor Research*, 40, 57-64, ISSN 0892-1016
- Belda, E.J.; Kvist, L., Monrós, J.S., Ponnikas, S. & Torralvo, C. (2009). Use of molecular techniques and discriminant analyses to determine with biometric measurements two subspecies of reed bunting *Emberiza schoeniclus*. *Ardeola*, 56, 85-94, ISSN 0570-7358

- Berkunsky, I.; Mahler, B. & Reboreda, J.C. (2009). Sexual dimorphism and determination of sex by morphometrics in Blue-fronted Amazons (*Amazona aestiva*). *Emu*, 109, 192-197, ISSN 0158-4197
- Berthold, P. & Fridrich, W. (1979). Die Federlänge ein neues nützliches Flü gelmass. *Vogelwarte*, 30 11-21, ISSN 0049-6650
- Bibby, C.J. & Thomas, D.K. (1984). Sexual dimorphism in size, moult and movements of Cetti's Warbler *Cettia cetti*. *Bird Study*, 31, 28-34, ISSN 0006-3657
- Bjordal, H. (1983). Effects of deep freezing, freeze-drying and skinning on body dimensions of House Sparrows *Passer domesticus*. *Fauna Norvegica Ser. C. Cinclus*, 6, 105-108, ISSN 0332-7701
- Bowlin, M.S. (2007). Sex, wingtip shape, and wing-loading predict arrival date at a stopover site in the swainson's thrush (*Catharus ustulatus*). *Auk*, 124, 1388-1396, ISSN 0004-8038
- Brady, R.S.; Paruk, J.D. & Kern, A.J. (2009). Sexing adult Northern Shrikes using DNA, morphometrics, and plumage. *Journal of Field Ornithology*, 80, 198-205, ISSN 0273-8570
- Brown, M.E. (1996). Assessing body condition in birds. In: *Current Ornithology, volume 13*, Nolan, V. & Ketterson, E.D. (Eds.), 67-135, Plenum Press, ISBN 9780306454738, New York
- Busse, P. (1983). Biometrical standards in the Operation Baltic. *Ring*, 116, 125-138, ISSN 0035-5429
- Campos, F.; Gutiérrez-Corcheró, F., López-Fidalgo, J. & Hernández, M. (2005a). A new criteria to sex European Dippers *Cinclus cinclus* in the Iberian Peninsula. *Revista Catalana d'Ornitología*, 21, 43-46, ISSN 1697-4697
- Campos, F.; Hernández, M., Arizaga, J., Miranda, R. & Amezcua, A. (2005b). Sex differentiation of Corn Buntings *Miliaria calandra* wintering in northern Spain. *Ring & Migration*, 22, 159-162, ISSN 0307-8698
- Campos, F.; Gutiérrez-Corcheró, F., Hernández, M.A., Rivas, J.M. & López-Fidalgo, J. (2005c). Biometric differences among Dipper *Cinclus cinclus* populations of Spain. *Acta Ornithologica*, 40, 87-93, ISSN 0001-6454
- Campos, F.; Gutiérrez-Corcheró, F., Hernández, M.A. & López-Fidalgo, J. (2008). Seasonal variation in the body size - body mass relationship in the Southern Grey Shrike *Lanius meridionalis*. *Acta Ornithologica*, 43, 139-143, ISSN 0001-6454
- Campos, F. & Martín, R. (2010). Spatial and temporal distribution of Southern Grey Shrikes *Lanius meridionalis* in agricultural areas. *Bird Study*, 57, 84-88, ISSN 0006-3657
- Catry, P.; Phillips, R.A. & Croxall, J.P. (2005). Sexual segregation in birds: patterns, processes and implications for conservation. In: *Sexual segregation in Vertebrates: Ecology of the two sexes*, K.E. Ruckstuhl & P. Neuhaus (Eds.), pp 351-378, Cambridge University Press, ISBN 0521835224, Cambridge
- Chandler, C.R. & Mulvihill, R.S. (1988). The use of wing shape indices: an evaluation. *Ornis Scandinavica*, 19, 212-216, ISSN 0908-8857
- Childress, B.; Harper, D., Hughes, B. & Ferris, C. (2005). Sex determination in the Lesser Flamingo (*Phoenicopterus minor*) using morphological measurements. *Ostrich*, 76, 148-153, ISSN 0030-6525
- Collar, N.J. (2005). Family Turdidae (Thrushes). In: *Handbook of the Birds of the World. Volume 10. Cuckoo-shrikes to Thrushes*, J. del Hoyo, A. Elliot & D. Christie (Eds.), pp 748-749, Lynx, ISBN 8487334725, Barcelona
- Cramp, S. & Simmons, K.E.L. (1977). *The Birds of the Western Palearctic, Volume 1*, Oxford University Press, ISBN 9780198573586, Oxford

- Cresswell, W. (1998). Diurnal and seasonal mass variation in blackbirds *Turdus merula*: consequences for mass-dependent predation risk. *Journal of Animal Ecology*, 67, 78-90, ISSN 0021-8790
- Cuervo, J.J.; Moller, A.P., De Lope, F. (2007). Haematocrit is weakly related to condition in nestling barn swallows *Hirundo rustica*. *Ibis*, 149, 128-134, ISSN 0019-1019
- Daniel, C.; Millar, C.D., Ismar, S.M.H., Stephenson, B.M. & Hauber, M.E. (2007). Evaluating molecular and behavioural sexing methods for the Australasian gannet (*Morus serrator*). *Australian Journal of Zoology*, 55, 377-382, ISSN 0004-959X
- Dmitrenok, M.; Puglisi, L., Demongin, L., Gilbert, G., Polak, M. & Bretagnolle, V. (2007). Geographical variation, sex and age in Great Bittern *Botaurus stellaris* using coloration and morphometrics. *Ibis*, 149, 37-44, ISSN 0019-1019
- Donald, P.F. (2007). Adult sex ratios in wild bird populations. *Ibis*, 149, 671-692, ISSN 0019-1019
- Ellrich, H.; Salewki V., Fiedler W. (2010). Morphological sexing of passerines: not valid over larger geographical scales. *Journal of Ornithology*, 151, 449-458, ISSN 0021-8375
- Esteban, L.; Campos, F. & Ariño, A.H. (2000). Biometrics amongst Dippers *Cinclus cinclus* in the north of Spain. *Ringling & Migration*, 20, 9-14, ISSN 0307-8698
- Eybert, M.C. ; Geslin, T., Questiau, S. & Beaufils, M. (1999). La Baie du Mont Saint-Michel: nouveau site de reproduction pour deux morphotypes de gorgebleu à miroir blanc (*Luscinia soecica namnetum* et *Luscinia soecica cyanecula*). *Alauda*, 67, 81-88, ISSN 0002-4619
- Fiedler, W. (2005). Ecomorphology of the external flight apparatus of blackcaps (*Sylvia atricapilla*) with different migration behaviour. *Annals of New York Academy of Sciences*, 1046, 253-263, ISSN 0077-8923
- Fletcher, K.L. & Hamer, K.C. (2003). Sexing terns using biometrics: the advantage of within-pair comparisons. *Bird Study*, 50, 78-83, ISSN 0006-3657
- García-Berthou, E. (2001). On the misuse of residuals in ecology: testing regression residuals vs. the analysis of covariance. *Journal of Animal Ecology*, 70, 708-711, ISSN 0021-8790
- Gosler, A.G.; Greenwood, J.J.D., Baker, J.K. & King, J.R. (1995). A comparison of wing length and primary length as size measures for small passerines. *Ringling & Migration*, 16, 65-78, ISSN 0307-8698
- Gosler, A.G.; Greenwood, J.J.D., Baker, J.K. & Davidson, N.C. (1998). The field determination of body size and condition in passerines: a report to the British Ringling Committee. *Bird Study* 45, 92-103, ISSN 0006-3657
- Green, A.J. (2001). Mass/length residual: measures of body condition or generators of spurious results? *Ecology*, 85, 1473-1483, ISSN 0012-9658
- Griffiths, R.; Double, M.C., Orr, K. & Dawson, R.J.G. (1998). A DNA test to sex most birds. *Molecular Ecology*, 7, 1071-1075, ISSN 0962-1083
- Guillemain, M. ; Mondain-Monval, J.Y., Johnson, A.R. & Simon, G. (2005). Long-term climatic trend and body size variation in teal *Anas crecca*. *Wildlife Biology*, 11, 81-88, ISSN 0909-6396
- Gunnarsson, T.G.; Gill, J.A., Goodacre, S.L., Gelinaud, G., Atkinson, P.W., Hewitt, G.M., Potts, P.M. & Sutherland, W.J. (2006). Sexing of Black-tailed godwits *Limosa limosa islandica*: a comparison of behavioural, molecular, biometric and field-based techniques. *Bird Study*, 53, 193-198, ISSN 0006-3657
- Gutiérrez-Corchero, F.; Campos, F., Hernández, M.A. & Amezcua, A. (2007a). Biometrics of the Southern Grey Shrike (*Lanius meridionalis*) in relation to age and sex. *Ringling & Migration*, 23, 141-146, ISSN 0307-8698

- Gutiérrez-Corchero, F.; Campos, F. & Hernández, M.A. (2007b). Sexual dimorphism in an insular Southern Grey Shrike subspecies *Lanius meridionalis koenigi*. *Ardeola*, 53, 327-330, ISSN 0570-7358
- Hansen, B.D.; Minton, C.D.T., Jessop, R. & Collins, P. (2009). Biometrics, sexing criteria, age-structure and moult of Sooty Oystercatchers in south-eastern and north-western Australia. *Emu*, 109, 25-33, ISSN 0158-4197
- Harper, D.G.C. (1994). Some comments on the repeatability of measurements. *Ringing & Migration*, 15, 727-738, ISSN 0307-8698
- Helander, B.; Hailer, F. & Vila, C. (2007). Morphological and genetic sex identification of white-tailed eagle *Haliaeetus albicilla* nestlings. *Journal of Ornithology*, 148, 435-442, ISSN 0021-8375
- Helm, B. & Albrecht, H. (2000). Human handedness causes directional asymmetry in avian wing length measurements. *Animal Behaviour*, 60, 899-902, ISSN 0003-3472
- Hermosell, I.G.; Balbontin, J., Marzal, A., Reviriego, M., & de Lope, F. (2007). Sex determination in barn swallows *Hirundo rustica* by means of discriminant analysis in two European populations. *Ardeola*, 54, 93-100, ISSN 0570-7358
- Hernández, M.A.; Campos, F. & Padilla, D.P. (2010). Tandem repeats in the mtDNA Control Region of the endemic Southern Grey Shrike from the Canary Islands. *Ardeola*, 57, 437-441, ISSN 0570-7358
- Herring, G.; Gawlik, D.E. & Beerens, J.M. (2008). Sex determination for the Great Egret and White Ibis. *Waterbirds*, 31, 298-303, ISSN 1524-4695
- Hobson, K.A. (2005). Stable isotopes and the determination of avian migratory connectivity and seasonal interaction. *Auk*, 122, 1037-1048, ISSN 0004-8038
- Hurley, V.G.; Hogan, F., White, J.G. & Cooke, R. (2007). A morphological model for sexing nestling peregrine falcons (*Falco peregrinus macropus*) verified through genetic analysis. *Wildlife Research*, 34, 54-58, ISSN 1035-3712
- Jacob, E.M.; Marshall, S.D. & Uetz, G.W. (1996). Estimating fitness: a comparison of body condition indices. *Oikos*, 77, 61-67, ISSN 0030-1299
- Jakober, H. & Stauber, W. (2000). Are Red-backed Shrikes becoming smaller? *Journal of Ornithology*, 141, 408-417, ISSN 0021-8375
- Jenni, L. & Winkler, R. (1989). The feather-length of small passerines: a measurement for wing-length in live birds and Museum skins. *Bird Study*, 36, 1-15, ISSN 0006-3657
- Johnsen, A.; Andersson, S., García-Fernández, J., Kempnaers, B., Pavel, V., Questiau, S., Raess, S., Rindal, E. & Lifjeld, J.T. (2006). Molecular and phenotypic divergence in the bluethroat (*Luscinia svecica*) subspecies complex. *Molecular Ecology*, 15, 4033-4047, ISSN 0962-1083
- Ketterson, E.D. & Nolan, V. (1983). The evolution of differential bird migration. In: *Current Ornithology, volume 1*, Johnston, R.F. (ed.), 357-402, Plenum Press, ISSN 0742-390X, New York
- Klassert, T.E.; Hernández, M.A., Campos, F., Infante, O., Almeida, T., Martell, N., Pestano, N. & Hernández, M. (2007). Mitochondrial DNA points to *Lanius meridionalis* as a polyphyletic species. *Molecular Phylogenetics and Evolution*, 47, 1227-1231, ISSN 1055-7903
- Kloskowski, J.; Grela, P., Krogulec, J., Gaska, M. & Tchorzewski, M. (2006). Sexing Red-necked grebes *Podiceps grisegena* by molecular techniques and morphology. *Acta Ornithologica*, 41, 176-180, ISSN 0001-6454
- Knox, A. (1980). Post-mortem changes in wing lengths and wing formulae. *Ringing & Migration*, 3, 29-31, ISSN 0307-8698

- Kuczynski, L.; Tryjanowski, P., Antczak, M., Skoracki, M. & Hromada, M. (2003). Repeatability of measurements and shrinkage after skinning: the case of the Great Grey Shrike *Lanius excubitor*. *Bonner Zoologische Beiträge*, 51, 127-130, ISSN 0006-7172
- Lefranc, N. & Worfolk, T. (1997). *Shrikes. A Guide to the Shrikes of the World*. Yale University Press, ISBN 0300073364, New Haven
- Lessells, C.M. & Boag, P.T. (1987). Unrepeatable repeatabilities: a common mistake. *Auk*, 104, 116-121, ISSN 0004-8038
- Lima, S.L. (1986). Predation risk and unpredictable feeding conditions: determinants of body mass in birds. *Ecology* 67, 377-385, ISSN 0012-9658
- Liordos, V. & Goutner, V. (2008). Sex determination of Great Cormorants (*Phalacrocorax carbo sinensis*) using morphometric measurements. *Waterbirds*, 31, 203-210, ISSN 1524-4695
- Lockwood, R.; Swaddle, J. P. & Rayner, M.V. (1998). Avian wingtip shape reconsidered: wingtip shape indices and morphological adaptations to migration. *Journal of Avian Biology*, 29, 273-292, ISSN 0908-8857
- MacLeod, R.; Barnett, P., Clark, J.A. & Cresswell, W. (2005). Body mass change strategies in blackbirds *Turdus merula*: the starvation-predation risk trade-off. *Journal of Animal Ecology*, 74, 292-302, ISSN 0021-8790
- MacLeod, R.; Clark, J. & Cresswell, W. (2008). The starvation-predation risk trade-off, body mass and population status in the Common Starling *Sturnus vulgaris*. *Ibis*, 150 (Suppl. 1), 199-208, ISSN 0019-1019
- Mayr, B. (1939). The sex ratio in wild birds. *American Naturalist*, 73, 156-179, ISSN 0003-0147
- McGregor, P.K. & Peake, T.M. (1998). The role of individual identification in conservation biology. In: *Behavioural Ecology and Conservation Biology*, T. Caro (Ed.), pp 31-55, Oxford University Press, ISBN 0195104900, New York & Oxford
- McNamara, J.M. & Houston, A.I. (1990). The value of fat reserves and the trade-off between starvation and predation. *Acta Biotheoretica*, 38, 37-61, ISSN 0001-5342
- Marchetti, K.; Price, T. & Richman, A. (1995). Correlates of wing morphology with foraging behaviour and migration distance in the genus *Phylloscopus*. *Journal of Avian Biology*, 26, 177-181, ISSN 0908-8857
- Meiri, S., & Dayan, T. (2003). On the validity of Bergmann's rule. *Journal of Biogeography*, 30, 331-351, ISSN 0305-0270
- Merilä, J. (1996). Genetic variation in offspring condition: an experiment. *Functional Ecology*, 10, 465-474, ISSN 0269-8463
- Mönkkönen, M. (1995). Do migrant birds have more pointed wings - A comparative study. *Evolutionary Ecology*, 9, 520-528, ISSN 0269-7653
- Moreno, J. (1989). Body-mass variation in breeding Northern Wheatears: a field experiment with supplementary food. *Condor*, 91, 178-186, ISSN 0010-5422
- Moreno-Rueda, G. & Rivas, J.J. (2007). Recent changes in allometric relationships among morphological traits in the dipper (*Cinclus cinclus*). *Journal of Ornithology*, 148, 489-494, ISSN 0021-8375
- Morgan, J.H. (2004). Remarks on the taking and recording of biometric measurements in birds ringing. *Ring*, 26, 71-78, ISSN 0035-5429
- Newton, I. (1979). *Population ecology of raptors*. T & AD Poyser, ISBN 9780856610233, Calton
- Newton, I. (1986). *The Sparrowhawk*. T & A.D. Poyser, ISBN 0856610410, Calton
- Ottwall, R. & Gunnarsson, G. (2007). Morphological and molecular sex identification of Redshanks *Tringa totanus*. *Bird Study*, 54, 127-129, ISSN 0006-3657
- Palmer, A.R. & Strobeck, C. (1986). Fluctuating asymmetry: measurement, analysis, patterns. *Annual Review of Ecology and Systematics*, 17, 391-421, ISSN 0066-4162

- Pérez-Tris, J. & Tellería, J.L. (2001). Age-related variation in wing shape of migratory and sedentary Blackcap *Sylvia atricapilla*. *Journal of Avian Biology*, 32, 207-213, ISSN 0908-8857
- Pitzer, S.; Hull, J., Ernest, H.B. & Hull, A.C. (2008). Sex determination of three raptor species using morphology and molecular techniques. *Journal of Field Ornithology*, 79, 71-79, ISSN 0273-8570
- Rands, S.A. & Cuthill, I.C. (2001). Separating the effects of predation risk and interrupted foraging upon mass changes in the Blue Tit *Parus cearuleus*. *Proceedings of the Royal Society of London, Series B*, 268, 1783-1790, ISSN 0370-1662
- Rands, S.A.; Houston, A.I. & Cuthill, I.C. (2006). Measurement of mass change in breeding birds: a bibliography and discussion of measurement techniques. *Ringing & Migration*, 23, 1-5, ISSN 0307-8698
- Redfern, C.P.; Slough, A.E.J., Dean, B., Brice, J.L. & Jones, P.H. (2000). Fat and body condition in migrating Redwings *Turdus iliacus*. *Journal of Avian Biology*, 31, 197-205, ISSN 0908-8857
- Reid, K.; Liddle, G.M., Prince, P.A. & Croxall, J.P. (1999). Measurement of chick provisioning in Antarctic Prions *Pachyptila desolata* using an automated weighing system. *Journal of Avian Biology*, 30, 127-134, ISSN 0908-8857
- Rodríguez, M.A.; Olalla-Tárraga, M.A. & Hawkins, B.A. (2008). Bergmann's rule and the geography of mammal body size in the Western Hemisphere. *Global Ecology & Biogeography*, 17, 274-283, ISSN 1466-8238
- Searcy, W.A.; Peters, S. & Nowicki, S. (2004). Effects of early nutrition on growth rate and adult size in song sparrows *Melospiza melodia*. *Journal of Avian Biology*, 35, 269-279, ISSN 0908-8857
- Senar, J.C.; Lleonart, J. & Metcalfe, N.B. (1994). Wing-shape variation between resident and transient wintering Siskins *Carduelis spinus*. *Journal of Avian Biology*, 25, 50-54, ISSN 0908-8857
- Shealer, D.A. & Cleary, C.M. (2007). Sex determination of adult black terns by DNA and morphometrics: test of sample size, temporal stability and geographical specificity in the classification accuracy of discriminant function models. *Waterbirds*, 30, 180-188, ISSN 1524-4695
- Sokal, R.R. & Rohlf, F.J. (1981). *Biometry*. W.H. Freeman & Co., ISBN 8472141667, San Francisco
- Summers, R.W.; Underhill, L.G., Waltner, M. & Swann, R.L. (2010). Differences in biometrics and moult of non-breeding Red Knots *Calidris canutus* in southern Africa and Scotland reflect contrasting climatic conditions. *Ibis*, 152, 127-135, ISSN 0019-1019
- Svagej, W.S. & Quintana, F. (2007). Sexual size dimorphism and sex determination by morphometrics measurements in breeding imperial shags (*Phalacrocorax atriceps*). *Waterbirds*, 30, 97-102, ISSN 1524-4695
- Svensson, L. (1992). *Identification Guide to European Passerines*, ISBN 9163011182, Stockholm
- Szép, T.; Barta, Z., Tóth, Z. & Sívári, Z. (1995). Use of an electronic balance with Bank Swallow nests: a new field technique. *Journal of Field Ornithology*, 66, 1-11, ISSN 0273-8570
- Szulc, B. (1964). A tentative assessment of certain morphological measurement used in ornithology. *Ekologia Polska, Seria B*, 10, 19-25, ISSN 0013-2969
- Van de Pol, M.; Oosterbeek, K., Rutten, A.L., Ens, B.J., Tinbergen, J.M. & Verhulst, S. (2009). Biometric sex discrimination is unreliable when sexual dimorphism varies within

- and between years: an example in Eurasian Oystercatchers *Haematopus ostralegus*. *Ibis*, 151, 171-180, ISSN 0019-1019
- Vanhooydonck, B.; Herrel, A., Gabela, A. & Podos, J. (2009). Wing shape variation in the medium ground finch (*Geospiza fortis*): an ecomorphological approach. *Biological Journal of the Linnean Society*, 98, 129-138, ISSN 0024-4066
- Vepsäläinen, K. (1968). Wing length of Lapwing (*Vanellus vanellus*) before and after skinning with remarks on measuring methods. *Ornis Fennica*, 45, 124-126, ISSN 0030-5685
- Vogeli, M.; Serrano, D., Tella, J.L., Méndez, M. & Godoy, J.A. (2007). Sex determination of Dupont's Lark *Chrisophilus duponti* using molecular sexing and discriminant functions. *Ardeola*, 54, 69-79, ISSN 0570-7358
- Wang, L.C.; Chen, C.T., Lee, H.Y., Li, S.H., Lir J.T., Chin, S.C., Pu, C.E., Wang, C.H. (2006). Cut feather containing rachis as a sampling way for avian sexing. *Zoo Biology*, 25, 279-283, ISSN 1098-2361
- Wiklund, C.G. (1990). Offspring protection by Merlin *Falco columbarius* females; the importance of brood size and expected offspring survival for defense of young. *Behavioral Ecology and Sociobiology*, 26, 217-223, ISSN 0340 5443
- Wiklund, C.G. (1996). Body length and wing length provide univariate estimates of overall body size in the merlin. *Condor*, 98, 581-588, ISSN 0010-5422
- Winker, K. (1993). Specimen shrinkage in Tennessee Warblers and "Traill's" Flycatchers. *Journal of Field Ornithology*, 64, 331-336, ISSN 0273-8570
- Winker, K. (1996). Specimen shrinkage versus evolution: I'iwi morphology. *Conservation Biology*, 10, 657-658, ISSN 0888-8892
- Witter, M.S. & Cuthill, I.C. (1993). The ecological cost of avian fat storage. *Philosophical Transactions of the Royal Society London B*, 340, 93-125, ISSN 0962-8436
- Witter, M.S.; Cuthill, I.C. & Bonser, R.H.C. (1994). Experimental investigations of mass-dependent predation risk in the European Starling, *Sturnus vulgaris*. *Animal Behaviour*, 48, 201-222, ISSN 0003-3472
- Wyllie, I. & Newton, I. (1994). Latitudinal variation in the body-size of Sparrowhawks *Accipiter nisus* within Britain. *Ibis*, 136, 434-440, ISSN 0019-1019
- Xirouchakis, S.M. & Poulakakisi, N. (2008). Biometrics, sexual dimorphism and gender determination of Griffon Vultures *Gyps fulvus* from Crete. *Ardea*, 96, 91-98, ISSN 0373-2266
- Yom-Tov, Y. (1993). Does the rock hyrax, *Procapra capensis*, conform with Bergmann's rule? *Zoological Journal of Linnean Society*, 108, 171-177, ISSN 0024-4082
- Yom-Tov, Y. & Yom-Tov, J. (2005). Global warming, Bergmann's rule and body size in the masked shrew *Sorex cinereus* Kerr in Alaska. *Journal of Animal Ecology*, 74, 803-808, ISSN 0021-8790
- Yom-Tov, Y.; Yom-Tov, S., Wright, J., Thorne, C.J.R. & Du Feu, R. (2006). Recent changes in body weight and wing length among some British passerine birds. *Oikos*, 112, 91-101, ISSN 0030-1299

Toward An Efficient Fingerprint Classification

Ali Ismail Awad¹ and Kensuke Baba²

¹*Graduate School of Information Science and Electrical Engineering,
Kyushu University,*

²*Library, Kyushu University,
Japan*

1. Introduction

Biometrics technology is keep growing substantially in the last decades with great advances in biometric applications. An accurate personal authentication or identification has become a critical step in a wide range of applications such as national ID, electronic commerce, and automated and remote banking. The recent developments in the biometrics area have led to smaller, faster, and cheaper systems such as mobile device systems. As a kind of human biometrics for personal identification, fingerprint is the dominant trait due to its simplicity to be captured, processed, and extracted without violating user privacy.

In a wide range of applications of fingerprint recognition, including civilian and forensics implementations, a large amount of fingerprints are collected and stored everyday for different purposes. In Automatic Fingerprint Identification System (AFIS) with a large database, the input image is matched with all fields inside the database to identify the most potential identity. Although satisfactory performances have been reported for fingerprint authentication (1:1 matching), both time efficiency and matching accuracy deteriorate seriously by simple extension of a 1:1 authentication procedure to a 1:N identification system (Manhua, 2010). The system response time is the key issue of any AFIS, and it is often improved by controlling the accuracy of the identification to satisfy the system requirement. In addition to developing new technologies, it is necessary to make clear the trade-off between the response time and the accuracy in fingerprint identification systems. Moreover, from the versatility and developing cost points of view, the trade-off should be realized in terms of system design, implementation, and usability.

Fingerprint classification is one of the standard approaches to speed up the matching process between the input sample and the collected database (K. Jain et al., 2007). Fingerprint classification is considered as indispensable step toward reducing the search time through large fingerprint databases. It refers to the problem of assigning fingerprint to one of several pre-specified classes, and it presents an interesting problem in pattern recognition, especially in the real and time sensitive applications that require small response time. Fingerprint classification process works on narrowing down the search domain into smaller database subsets, and hence speeds up the total response time of any AFIS. Even for

fingerprint recognition, a large number of classification methods have been proposed (summarized in Section 2).

This chapter proposes a novel method for fingerprint classification using simple and established image processing techniques. The processing time of the proposed method is dramatically decreased with a small effect on the resulted classification accuracy. The processing time and the accuracy of the proposed classification method have been evaluated by intensive experiments over different standard fingerprint databases. The time-accuracy optimization is not trivial task for every biometrics based practical systems from theoretical to practical implementations of the classification algorithm. For example, selecting extremely complex features for performing classification might increase the processing time in a pattern matching, and hence, reducing the overall system performance. The total accuracy of any identification system depends on the distribution of the features in addition to the classification accuracy.

In the rest of this chapter, first we shade light on the existing classification methods. In common, fingerprint classification algorithms extract features from the interleaved ridge and valley flows on fingerprints. In terms of the previous features, fingerprints are classified by Sir Henry (Maltoni et al., 2009) into the common five classes, Arch, Tented Arch, Left Loop, Right Loop, and Whorl. One of the standard approaches for fingerprint classification is to use the information extracted by frequency domain analysis of input images. Some standard calculations on frequency domain are well studied, hence we can benefit from the refined algorithms and Application Specific Integrated Circuit (ASIC) for implementation. Our algorithm works different from any other approach in the literature by dividing a fingerprint image into four sub-images, and then applies the standard frequency-based algorithm to each sub images to extract distinguished feature based on ridge (*periodicity and directionality*) inside each sub image. Then, the classification process uses those extracted features to exclusively classify it into four classes (Tented Arch is regarded as Arch). We have implemented the algorithm, evaluated its processing time and classification accuracy on two standard databases.

The contribution of this chapter falls under the possibility to maximize time-accuracy trade-off by implementing simple techniques to build an effective fingerprint classification. The novelty of the classification method falls under the extraction of distinguished patterns from frequency domain representation of the fingerprint. Due to its simplicity, it is expected that the method may be combined with other advanced technologies such as machine learning (Yao et al., 2003) to improve both its robustness and efficiency.

2. Review of fingerprint classification

Fingerprint classification is still a hot research topic in the area of biometric authentication. Generally, the advantage of classification is that it provides an indexing mechanism and facilities the matching process over the large databases. Without a robust classification algorithm, identification performs exhaustive matching processes to an input with all of the available elements in the database, which is computationally demanding. Fingerprint classification is usually based on global features such as global ridge structure and core or delta singular points. The core point is defined as the topmost point of the innermost curving ridge, where the delta point is defined as the centre of triangular regions where three different direction flows meet (Espinosa-Dur, 2001).

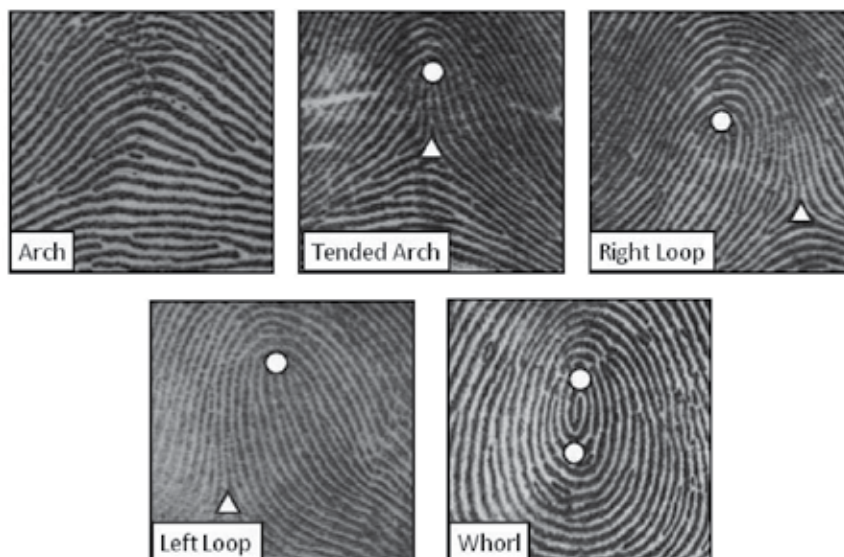


Fig. 1. Common five classes of fingerprints with singular points (*Circle-Core, Triangle-Delta*)

Fingerprint classification methods can be grouped into two main categories: continuous classification and exclusive classification (Maltoni et al., 2009). Figure 1 shows examples of exclusive fingerprint classes with related singular core and delta points (Amin & Neil, 2004).

2.1 Continuous fingerprint classification

In general, continuous classification overcomes some defects of exclusive classification by representing each fingerprint by a vector which summarizes its main features, instead of assigning them into a single class. (Lumini et al., 1997) proposed a continuous classification scheme which characterizes each fingerprint with a numerical vector. Apparently, continuous classification does not allow some tasks to be executed such as fingerprint labelling according to a given classification scheme. The continuous classification approach is more preferable than the classical exclusive approach if we want to classify fingerprints only for improving the fingerprint retrieval efficiency.

2.2 Exclusive fingerprint classification

Exclusive fingerprint classification groups fingerprint images into some predefined classes according to their global features. Most of fingerprint identification systems use that exclusive fingerprint classification approach (Cappelli et al., 1999) to improve the total response time. Global patterns of ridges and furrows in the central region of the fingerprint form special configuration, see Figure 1, which have a certain amount of intraclass variability. These variations are sufficiently small which allows a systematic classification of fingerprint (Wang et al., 2006). Galton (K. Jain et al., 2007) has made the first scientific studies on fingerprint classification area. He exclusively divided fingerprint into three major classes: Loop, Arch, and Whorl. Galton's algorithm is then refined by increasing the number of classes into eight classes: Plain Arch, Tended Arch, Right Loop, Left Loop, Plain Whorl, Central Pocket, Twin Loop, and Accidental Whorl.

Arch is a special type of fingerprint configuration, as less than 5% of all fingerprints is arches. Plain Arch is defined as a "type of fingerprint in which ridges enter one side and

flow out of the other with the rise of wave in the center". In Tented Arch, most of the ridges enter one side and flow out of the other with rise wave in the center and the rest of the ridges form a definite angle (Maltoni et al., 2009). Arch and Tented Arch classes are grouped into one class due to the small intra-class variations. Loop class is defined as a "type of fingerprints in which one or more of the ridges enter on fingerprint side, recurve, and touch or pass an imaginary line drawn from the delta to the core, and terminate or tend to terminate on or toward the same side from which such ridge or ridges entered" (Maltoni et al., 2009). A Whorl is "that type of fingerprint in which at least two deltas are present with a recurve in front of each". However, these preceding definitions are very general, but they catch the essence of the category. The performance of the exclusive classification strongly depends on the number of classes and the distribution of fingerprints. Unfortunately, in exclusive system the number of classes is small and fingerprints are not uniformly distributed. Also there are many ambiguous fingerprints whose exclusive classes that can not reliably be stated even by human experts. Exclusive classification allows the efficiency of the 10-print based identification to be improved, since the knowledge of the classes of the ten fingerprints can be used as a code for limiting the number of minutiae comparisons.

2.2.1 Graph based classifications

Graph based method, represented in Figure 2, is an example of spatial domain based classifiers. The basic idea of graph based classification scheme is partitioning the directional fingerprint image into homogenous regions, and these regions and the relations among them contain information useful for classification. The approach in (Maltoni & Maio, 1996) is divided into four main steps: computation of the directional image, segmentation of the directional image, construction of the relational graph, and the graph matching process. The relational graph is built by creating a node for each region and an arc for each pair of adjacent regions. Produced graph structure summarizes the topological features of the fingerprint by appropriately labeling the nodes and arcs of the graph. Although graph based approaches have interesting properties such as robustness to image rotation, displacement, and its ability to handle partial fingerprints, it is not easy to accurately partition the orientation image into homogeneous regions, especially in a poor quality fingerprint images. Producing good directional fingerprint image also needs preprocessing, binarization, and thinning which are time exhaustive operations that may impose impact on the overall system performance.

2.2.2 Dynamic mask approach

(Cappelli et al., 1999) have extended the graph based method, explained in the preceded paragraph, using dynamic mask approach that controls the freedom of fingerprint image segmentation process. A set of dynamic masks, directly derived from the most dominant fingerprint classes, are used to guide the image partitioning process. For every input fingerprint image, an application cost function is calculated for each dynamic mask. Intuitively, the application cost function measures how well mask fits with the input fingerprint image. A dynamic mask is built for only five fingerprint classes: Arch, Left Loop, Right Loop, Tented Arch, and Whorl. The smaller cost function value is the closer to the true fingerprint class.

There are many fingerprint classifications described in the literature (Maltoni et al., 2009). They can be grouped based on the used features and the type of the proposed classifiers.

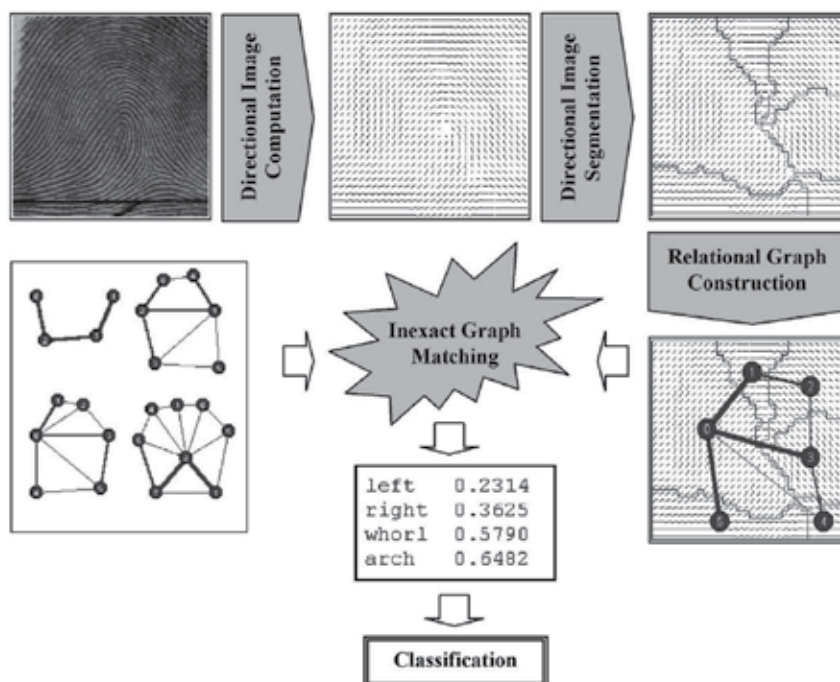


Fig. 2. Flowchart of graph based fingerprint classification technique, (Maltoni & Maio, 1996)

The most important types of classification techniques include Neural Network classifiers as in (Senior, 2001; Wang et al., 2006), the statistical based approach can be found in (Cappelli et al., 2002; K. Jain & Minut, 2002; Yao et al., 2003), and the rule-based classification approaches (K. Jain et al., 1999) that may use the numbers and relations of the singular points as a base for fingerprint classification process.

3. An efficient fingerprint classification

The proposed novel classification method is presented in this section. There are some classification methods exist which apply the idea of Fast Fourier Transform (FFT) to extract features from fingerprint images such as (Green & Fitz, 1996), (Sarbadhikari et al., 1998), and (Park & Park, 2005). These methods used the frequency representation of the full fingerprint image in the classification process. However, these methods come with a new idea, but they failed to achieve good results because the classes overlapping. The proposed method is novel and overcomes the classes overlapping problem, it also facilitates the texture property of fingerprint image by building four different patterns for each class using image division process. The main idea behind our method is that fingerprint images are divided into four sub-images, and then a standard FFT is applied to each sub-image to extract the class discriminant features. The prototype of the proposed algorithm can found in (Awad et al., 2008).

3.1 Outline

In our method, we consider classification of fingerprint images with four classes, Arch, Left Loop, Right Loop, and Whorl. The novel method consists of the following stages; Figure 3 introduces the algorithm flowchart that describes the following steps:

1. Calculation of standard classes patterns (four selected classes from a given database),
2. Acquisition of the input fingerprint image,
3. Division of the input image into four sub-images,
4. Transformation of the sub-images into frequency domain,
5. Patterns extraction for the input image,
6. Matching of the calculated pattern with the standard patterns calculated in step (1),
7. Decision making for the four classes.

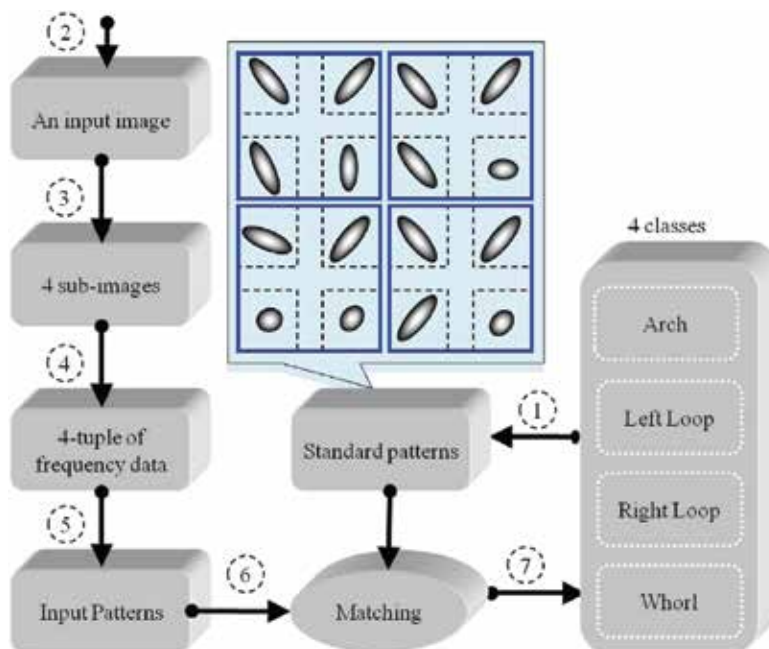


Fig. 3. Block diagram of patterns based fingerprint classification algorithm

The classification algorithm supports input fingerprint image in different formats, and the images size can be up to (512×512) pixels. Since the algorithm is an exclusive classifier the input image will be matched only with the standard classes to detect the correct class. The proposed algorithm can easily accept shifted, rotated, and even the poor quality images.

3.2 Division of Fingerprint image

At step (3), the input image is divided into four sub-images (a sub-image is sometimes called a "block" in the rest of this chapter) based on (x, y) lengths. Figure 4 shows an example of the division process. Fingerprint partitioning provides the ability to process fingerprint image as four different blocks with its own ridge frequency and direction. The number of blocks (four) has been selected due to processing time and computational complexity considerations. Four blocks selection compromising the trade-off between processing time and accepted algorithm's performance. Although the accuracy and the processing time of a classification method depend on the patterns (features) and the procedure of matching in general, roughly speaking, it is expected that the accuracy is better, but the processing time is get worse when the number of the sub images is being increased.

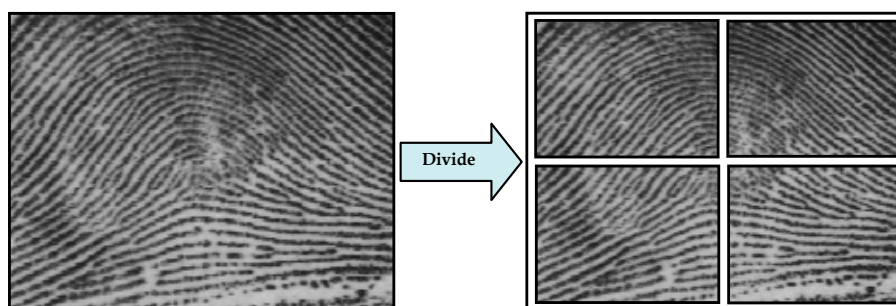


Fig. 4. A divided fingerprint image into four blocks (Input image was Arch)

3.3 Transformation into frequency domain

The simplest method to transform fingerprint images from spatial domain to frequency domain is 2D-FFT (Gonzalez et al., 2009). The FFT-based approach for estimating the frequency and direction of an image is an established method (Sherlock et al., 1994; Sarbadhikari et al., 1998; Park & Park, 2005; Gonzalez et al., 2009). In general, fingerprints have a definite periodicity of ridges or valleys, therefore the periodicity and directionality of ridges obtained by FFT could be a quantifier of the fingerprint texture in different directions. For the various fingerprint classes, FFT components are likely to be different. Moreover, since these frequency features are global in nature, they are likely to be less sensitive to shift, rotation, and noise. In our method, a 2D-FFT is applied individually to each sub-image. Since the ridge's direction and frequency of the fingerprint image are not constant in overall image, they will be different from one block to another. The key issue of the proposed method is to use these distinguished outputs to generate patterns for matching with the standard classes. We found the combinations of the frequency patterns of four blocks which realize a classification into the four common fingerprint classes. Figure 5 shows the FFT representation of all sub-images of a fingerprint in the Arch class. The frequency pattern in each block is clearly observed as a different pattern from the others in the senses of the size and the direction. These patterns will be extracted from FFT images in the next step.

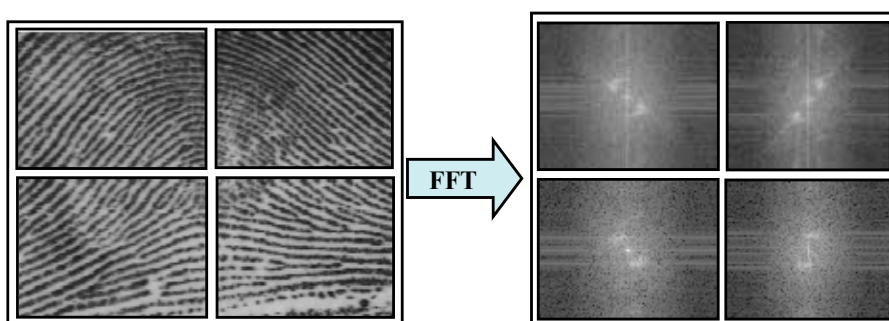


Fig. 5. Frequency domain representation for each sub image using 2D-FFT

3.4 Extraction of frequency patterns

Patterns extraction is the most important stage in our proposed classification method. The pattern of each class is constructed from the FFT outputs of four sub-images; therefore, the pattern of a single image is a 4-tuple of patterns. First, standard patterns of the four standard classes are extracted once and stored in a system buffer. The calculation of the standard patterns is based on the direction and shape of the FFT output. By considering the combination of 4 patterns, the proposed method achieves an accurate classification results. In the matching stage the system compares the 4-tuple of patterns of an input image with the 4-tuples of the standard classes. Figure 6 shows the frequency representation of the four fingerprint classes.

In this chapter, we considered simply the image of the FFT output as a frequency pattern. However, there is scope for further study about the representation of the pattern. We describe an idea of the representation in the rest of this subsection. In the pattern extraction, we considered that the output of FFT can be affected by three parameters: (i) ridge direction, (ii) ridges frequency or pitch, and (iii) the brightness variation in the block. The direction of output frequency is perpendicular to the total ridges direction in the block, while the ridge frequency appears in the frequency representation as a white spots on the line, the distance between these spots are inversely proportional to the ridges frequency. The pattern extraction process may consist of the following steps:

- Numbering each block,
- Computing the frequency orientation, and
- Deriving the output shape of FFT using simple morphological operations.

Figure 7 shows an example of the expected patterns corresponds to the FFT output in Figure 6.

3.5 Patterns matching

As we mentioned in the previous subsection, each element of the 4-tuple for a pattern is an image of the FFT output. Pattern padding process guarantees that the image is (300×300) pixels. We implemented the pattern matching of blocks by two methods, the absolute image difference and the 2D image correlation. To confirm that the two methods should be able to recognize each class, we operated prior experiments for the both methods.

3.5.1 Difference-based matching

The output of the matching process is held as a matrix with the same dimensions of the block used in matching process, that is, the matrix has the (300×300) elements. Figure 8 and Figure 9 both show a part of the results of the comparison based on the absolute image difference. Figure 8 is for the comparison of a Whorl pattern with the standard patterns of the four classes, where Figure 9 is of a random image.

We selected only the maximum values inside the matrix to show the results in appreciate format. Full patterns matching produces result that makes the decision maker able to classify different input images into its appreciating classes. In the graphs, the horizontal axis shows the columns of the output matrix, where we selected only the maximum values inside the matrix to show the results in appreciate format, therefore the length is come to 150. The vertical axis shows the summation of the elements of each column for the four blocks, that is, the total of the (300×4) elements. By the result, we can see that the difference-based matching is applicable for the pattern marching.

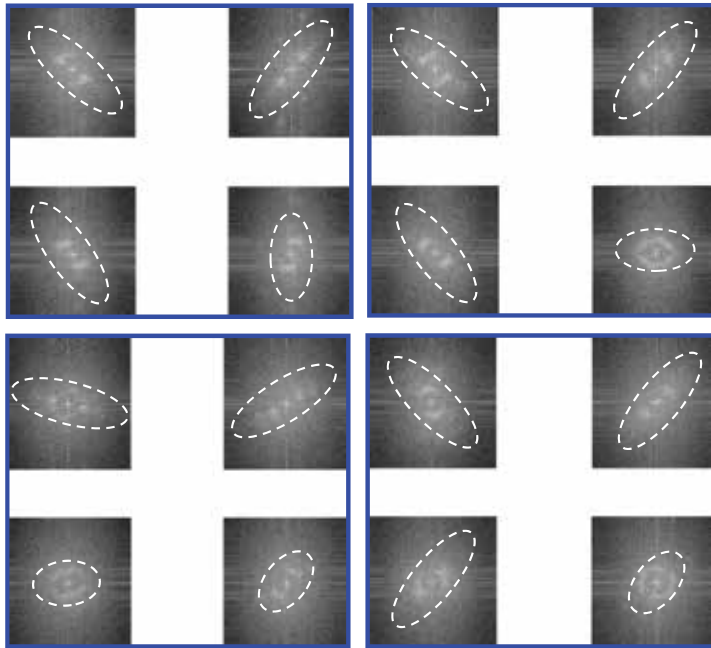


Fig. 6. Frequency transformation for each class (up left (Arch), up right (Left Loop), down left (Right Loop), down right (Whorl))

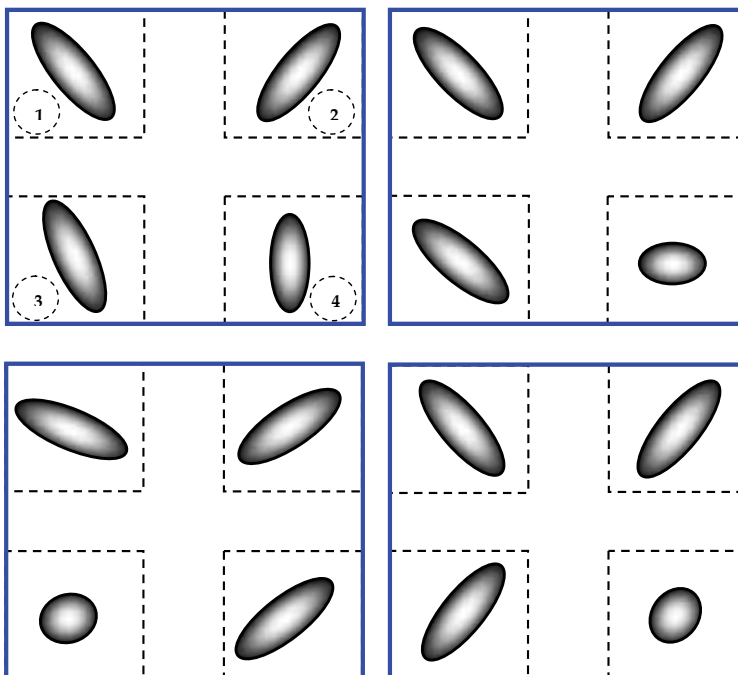


Fig. 7. The patterns extracted from sub images in Figure 6. Numbers inside the dashed circles are representing the block order

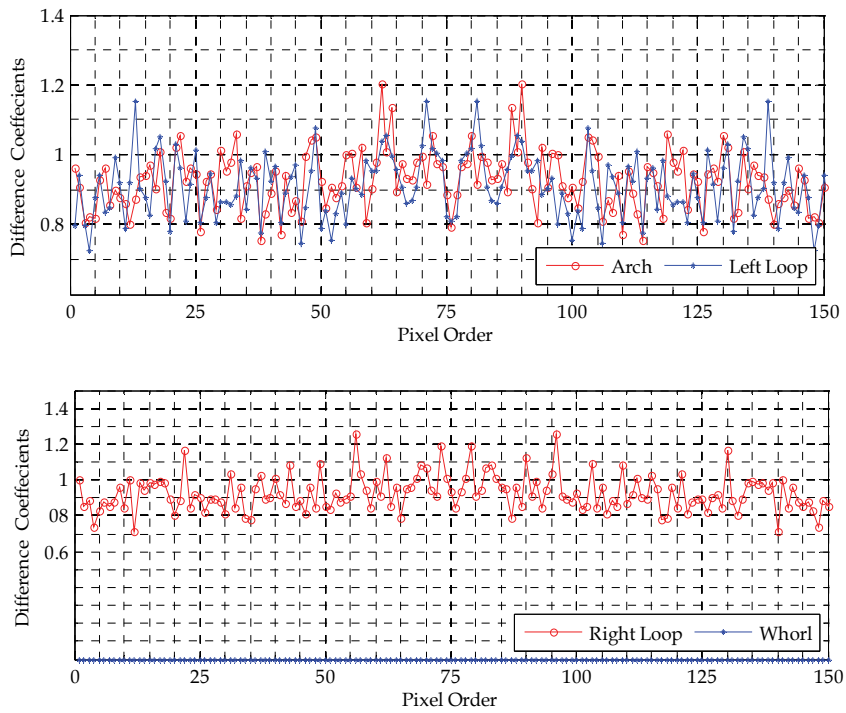


Fig. 8. Difference between the standard patterns and a Whorl patterns

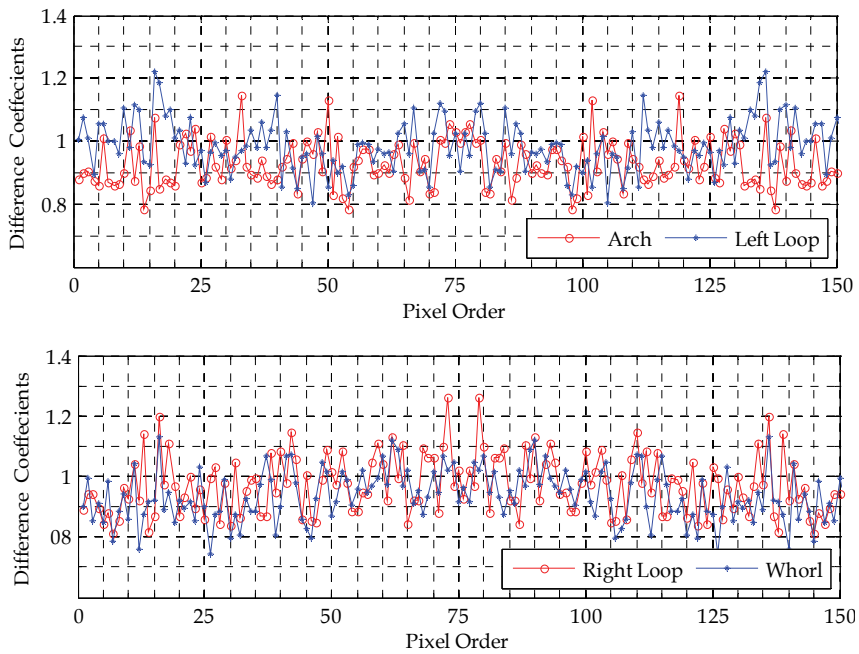


Fig. 9. Difference between the standard patterns and a random input pattern (Arch)

3.5.2 Correlation-based matching

Image correlation is much easier especially in frequency domain. The key issue of our proposed algorithm is the response time. We conducted intensive experimental work on performing pattern matching in frequency domain to make the processing time as short as possible. Comparing to image difference method, image correlation give a shorter response time with high matching accuracy. Also, the output data of the correlation process is little and it could be plotted or represented easily. Figure 10 and Figure 10 both are the result of the comparisons with the standard pattern with a Whorl pattern and a random pattern, respectively. In the graphs, "Block Number" is the number for the order of the four sub-images, (1), (2), (3), and (4) for up left, up right, down left, and down right, respectively. By the result, we can see that the correlation-based matching is also applicable for the patterns matching.

3.6 Decision making

The decision maker is responsible for selecting the final class from the information provided by the pattern matching stage. As we mentioned before, the matching results are stored in a matrix, and we use only the maximum matching values to preserve memory and plot the results in acceptable format. In the experiments in the following section, we considered the simple total of the elements of the matrix. However, there is scope for further study also about strategy of the decision making.

3.7 Singular points detection

In (Section 3.2), fingerprint images are divided simply based on the length. However, we are considering that the accuracy of classification should be improved by an ingenious scheme of the division. Figure 1 shows the common classes of fingerprint with core and delta points. The most popular approach for detecting fingerprint singularities is the method based on the Poincaré index. We describe the basic idea of the method in the rest of this subsection.

Since the Poincaré index is working on the direction changes, then the first step before calculating the Poincaré index is to extract the directional (orientation) image corresponding to the input fingerprint, Figure 12 shows the orientation filed for both core and delta point. To increase the accuracy of the orientation image, some enhancement techniques such as (Awad et al., 2007a) and (Awad et al., 2007b) can be implemented prior to the directional field estimation. We assume that $\theta(i, j)$ is pixel orientation of any directional image element pixel (i, j) , where $0 \leq \theta(i, j) \leq \pi$. Let (i_k, j_k) for $0 \leq k \leq N - 1$ is the element selected for calculating the Poincaré index of a point (i, j) . Then, the Poincaré index is defined as:

$$Poincaré(i, j) = \frac{1}{2\pi} \sum_{k=0}^{N-1} \Delta(k), \quad (1)$$

where

$$\Delta(k) = \begin{cases} \delta(k) & \text{if } |\delta(k)| < \pi / 2 \\ \pi + \delta(k) & \text{if } \delta(k) \leq -\pi / 2 \\ \pi - \delta(k) & \text{otherwise} \end{cases} \quad (2)$$

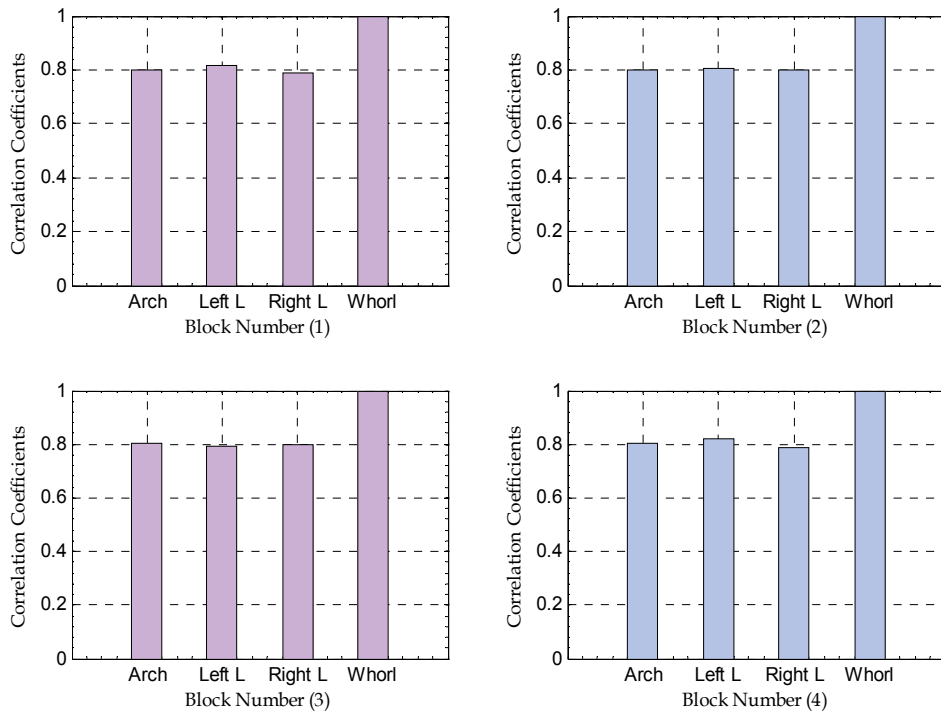


Fig. 10. Correlation results between input image patterns (Whorl) and other four classes

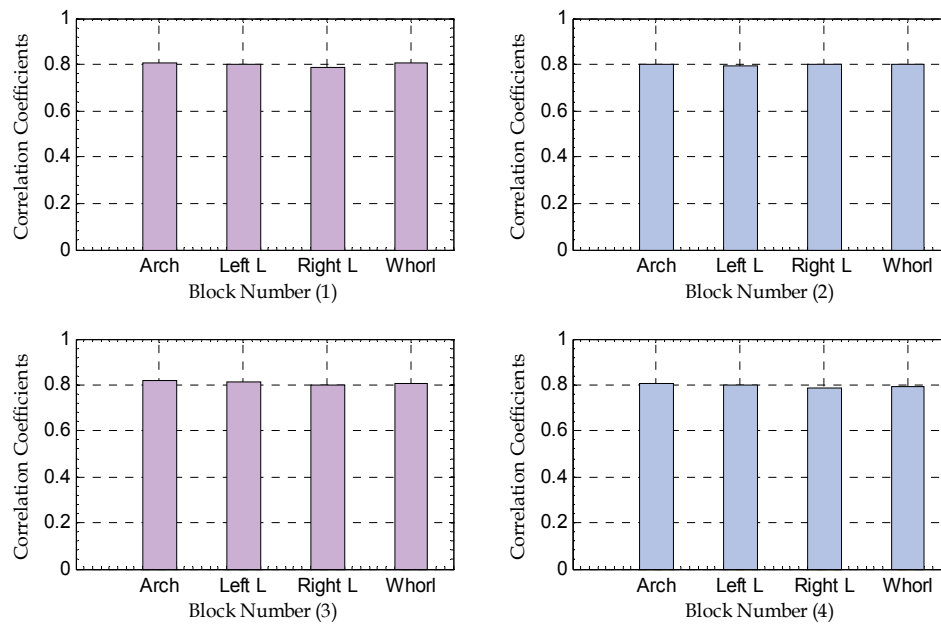


Fig. 11. Correlation results between input image patterns (Arch) and other four classes

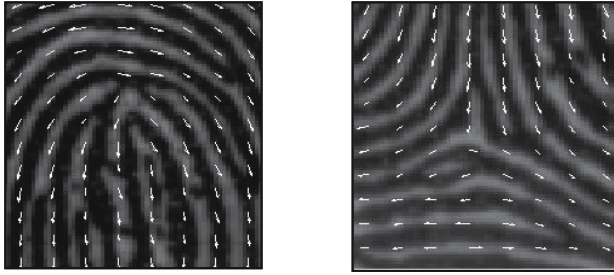


Fig. 12. Estimated orientation field for core (left) and delta (right)

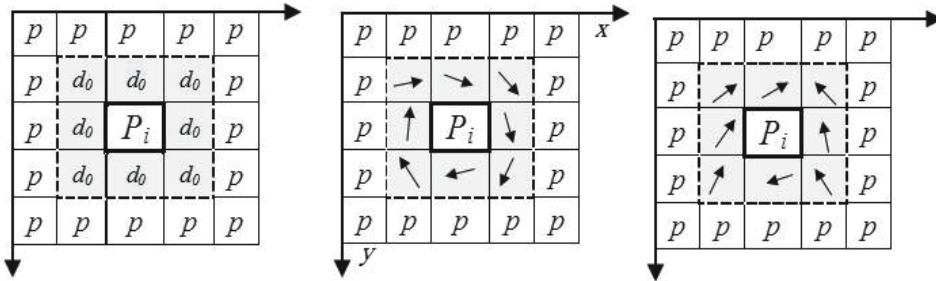


Fig. 13. Poincaré index representation: general Poincaré index calculation (left), Poincaré index = 1.0 (middle), Poincaré index = -0.5 (right)

and

$$\delta(k) = \theta(x_{(k+1) \bmod(N)}, y_{(k+1) \bmod(N)}) - \theta(x_k, y_k) \quad (3)$$

Then, the Poincaré index may have four kinds of values: 0 which means no singular point available in the area, $1/2$ for a core point, $-1/2$ for a delta point, and 1 which means that the selected area may have two singular points. Figure 13 shows the differences between the most dominant values of Poincaré index.

4. Experimental results and evaluations

The proposed efficient fingerprint classification algorithm has been intensively evaluated through different conducted experiments. The overall consumed processing time has been optimized to enhance the overall algorithm performance. We have implemented two matching algorithms; pattern matching by image difference and by image correlation also. Optimization process tried to reduce the algorithm's response time to its minimum value.

4.1 Data sets

In general, standard databases are used for implementation and evaluation of fingerprint recognition or classification methods. We used NIST-4 for evaluating the accuracy of the classification by the proposed method. Actually, in a lot of related work the evaluation is operated on the whole or a part of NIST-4. As for the processing time, in addition to NIST-4, four subsets of Fingerprint Verification Competition 2004 (FVC2004) (Maltoni et al., 2009) were used (Figure 14 shows samples of fingerprint images in different subsets of FVC2004).

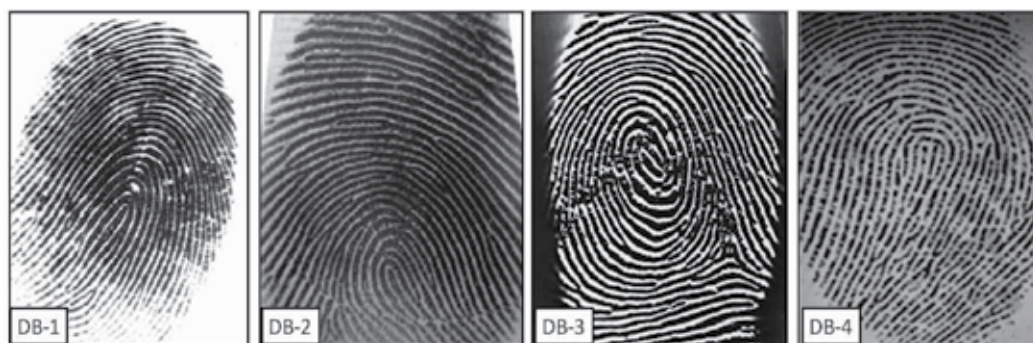


Fig. 14. Sample fingerprint images taken form FVC2004 available databases.

The selection criterion of the databases was interested to choose variety of fingerprints collected by different methods including optical, thermal sweeping sensors, and synthetic fingerprints. FVC2004 includes four sub databases: two categories generated by optical sensors, "V300" by CrossMatch and "U.are.U 4000" by Digital Persona respectively, the third one generated by thermal sensor "FingerChip FCD4B14CB" by Atmel, and the fourth is a synthetic by SFinGe proposed by (Cappelli, 2009).

4.2 Accuracy

The idea of confusion matrix is a common way to measure the performance of fingerprint classification algorithms. In a confusion matrix, a row and a column correspond to each actual class and each predicted class, respectively. Therefore, the diagonal elements are corresponding to the fingerprints that have been correctly classified. Table 1 is the confusion matrix resulted from applying the proposed method on NIST-4 database.

True Classes	Assigned Classes			
	A	R	L	W
Arch	912	37	36	5
R Loop	7	672	3	23
L Loop	10	6	780	33
Whorl	1	30	35	731

Table 1. The confusion matrix of implementing proposed method on NIST-4 database

Table 2 is the same result expressed in terms of the ratio, where "Fail Reject" and "Fail Accept" correspond to the ratios of the samples classified correctly in each row and column, respectively (therefore, both of them have the same value in "Total"). The result 6.9% of the error rate for the proposed method should be compared with the result 6% in (Park & Park, 2005) which uses FFT and NIST-4. The error rate is slightly worse compared to the existing method, however the calculation time is extremely small compared to the same existing method.

True Classes	Rates		
	Correct	Fail Reject	Fail Accept
Arch	92.1	7.9	1.9
R Loop	94.1	5.9	8.7
L Loop	95.3	4.7	9.8
Whorl	91.7	8.3	7.7
Total	93.1	6.9	-

Table 2. Classification results with False Acceptance and False Rejection rates of the proposed algorithm (%)

4.3 Processing time

Response time is the key issue in all fingerprint classification methods. We evaluated the processing time of the proposed method with respect to each step. The experiments were operated with Intel® Pentium 4 Core 2 Due™ processor (T9300, 2.5 GHz), 3 GB RAM, and Matlab® R2009b version. Table 3 represents the results of the processing time of the proposed method for one input fingerprint image. Each value is the average of the results for the fingerprint images in the databases. In the process of pattern matching, we evaluated the time of the difference-based method for the comparison with the correlation-based method. Note that the processes of “Division” and “FFT” are common in both methods. By the results, the total processing time of the proposed method is generally short in the sense of an application as an identification system. The correlation-based method is improving the computing time for the pattern matching process from the difference-based method.

	Division	FFT	Pattern Matching	Total
Correlation-based matching	0.0200	0.0229	0.0134	0.0552
Difference-based matching	0.0200	0.0229	0.0620	0.1056

Table 3. The processing time of the proposed method for each step (seconds)

4.4 Considerations

As we mentioned in (Subsection 3.6), we should consider a better scheme for the pattern matching and decision making. Does every block has the same weight during the matching process? Our experimental results proved that the answer is “No”, and hence each block pattern may have different weight during the matching step. Furthermore, we go to optimize the matching processing by selecting number of blocks to be matched in each class. In other words, instead of matching the four blocks of each class, we match only the most dominant blocks in each class. The weight for each block and order are different for each fingerprint class, therefore the problem was how to select the most optimum blocks for each class. The problem has been solved empirically by testing each block individually and tried to assign it a weight related to the correlation matching results of the input image with its corresponding standard one. The test has been extended to check two by two blocks, and three by three blocks as well.

Matching two by two blocks, blocks number (1) and (2) of the input image with the corresponding blocks in the standard classes as an example, is considered as the summation of matching two individual blocks. It helped us to find the block that produces maximum correlation in the patterns matching process. As a case study, we input a Left Loop fingerprint, the classification algorithm detected that the input image is a Left loop correctly; however less block difference or high correlation score was our goal. Through the optimization process, we are trying to achieve as much smaller difference between blocks as possible, and hence we were seeking to get a less block difference between the input image patterns and the Left Loop standard class patterns. The first graph in Figure 15 represents the matching results come from using two blocks matching in the matching process and discarding the others. From that figure we see that the blocks (2) and (4) produce a minimum block difference that considered as a high matching score, where blocks (1) and (3) produce a low matching score. Therefore, we conclude that blocks (2) and (4) are the first optimum choice as a base blocks for matching the input image with the Left Loop class.

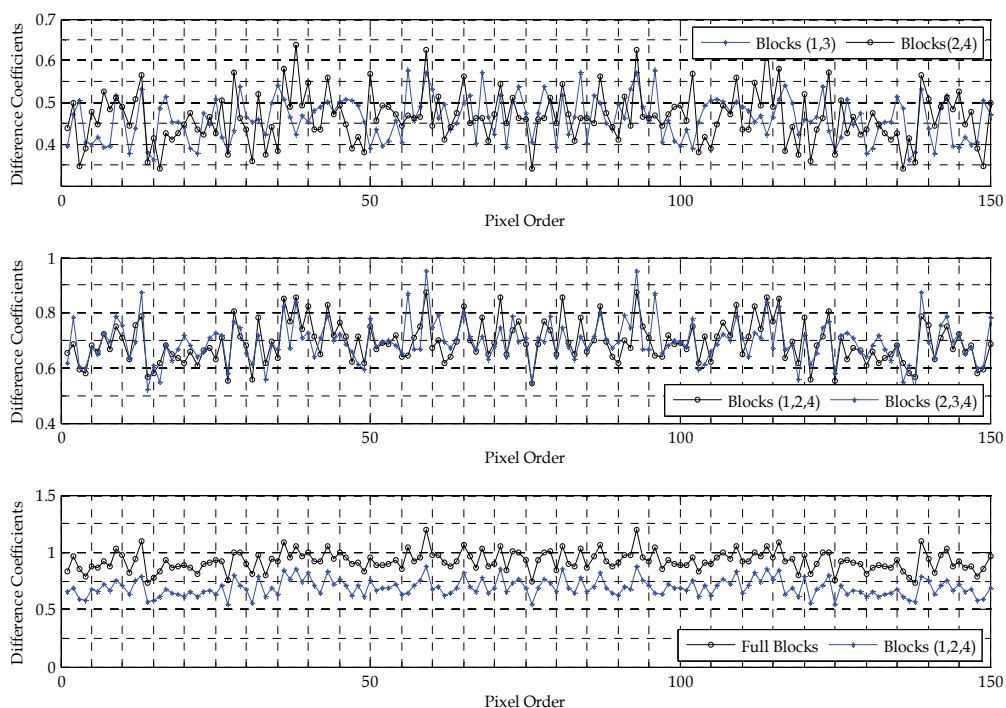


Fig. 15. Patterns matching optimization by two by two blocks (Upper), three by three blocks (Middle), and the amount of gained matching optimization (Bottom)

Due to the overlapping between classes, we found that two blocks are not sufficient for classification decision. Therefore, they are only used as a base blocks for determination of the maximum and minimum matching scores. To prevent overlapping between different fingerprint classes, we used three blocks from total patterns in matching process. The second graph in Figure 15 shows that the matching score produced from using blocks (1), (2), and (4) is much more than one produced by using blocks (2), (3), and (4). This figure leads us to conclude that blocks number (1), (2), and (4) are the optimum choice when

matching any input fingerprint image with the standard Left Loop class, and the third block may be discarded. The third graph in Figure 15 shows the amount of optimizations achieved by selecting three particular blocks in the matching process comparing to use the full patterns (four blocks). As we mentioned before, this experiments are operated as a case study and the order of blocks is different for each class. Therefore, we are going further to extend in the future the optimization process to include the remaining three classes to complete the general weight for our pattern matching.

5. Conclusions and future work

In this chapter, we proposed an efficient method which classifies fingerprints into the standard four classes. The basic idea of the proposed algorithm is dividing a fingerprint image into four sub-images before applying the standard FFT-based method of fingerprint classification. We have evaluated the proposed method in terms of the accuracy and the processing time by experiments with the standard databases NIST-4 and FVC2004. As the result, the proposed method achieves an extremely speed-up with a small classification losses compared to the simple techniques we have used. One of our future works is an improvement of the classification accuracy. For the improvement, we are considering that these is scope for further study about the division of images, the calculation of patterns, and the decision making. We mentioned about some ideas for the topics in Section 3.7, Section 3.4, and Section 4.4 respectively.

6. Acknowledgement

This work was supported by the Grant-in-Aid for Young Scientists (B) No. 22700149 of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) from 2010 to 2012.

7. References

- Amin, A. & Neil, Y. (2004). Fingerprint Classification: a Review. *Springer Verlag, Pattern Analysis & Applications*, Vol.7, No.1, (April, 2004), 77-93.
- Awad, A. I.; Mustafaa, M. & Moness, M. (2008). A New Fingerprint Classification Approach Based on Fast Fourier Transformer. *Proceedings of the 5th International Conference on Informatics and Systems*, pp. 78-83, March 2008, Faculty of Computers & Information-Cairo University, Cairo, Egypt.
- Awad, A. I.; Mustafaa, M. & Moness, M. (2007a). Fingerprint Image Enhancement based on Frequency Domain Representation. *Journal of Al Azhar University Engineering Sector (JAUES)* , Vol.2, No.12, (April, 2007a), 483-495.
- Awad, A. I.; Mustafaa, M. & Moness, M. (2007b). Fingerprint Image Enhancement based on Wavelet Filters. *Proceedings of Al Azhar Engineering Ninth International Conference*, pp. 100-113, April 2007b, Al Azhar, Cairo, Egypt.
- Cappelli, R. (2009). SFinGe. in *Stan Li (Editor), Encyclopedia of Biometrics*. Springer.
- Cappelli, R.; Lumini, A.; Maio, D. & Maltoni, D. (1999). Fingerprint Classification by Directional Image Partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , Vol.21, No.5, (May, 1999), 402-412.

- Cappelli, R.; Maio, D. & Maltoni, D. (2002). A Multi-Classifer Approach to Fingerprint Classification. *Springer Verlag London, Pattern Analysis & Applications*, Vol.5, No.2, (May, 2002), 136-144.
- Espinosa-Dur, V. (2001). Minutiae Detection Algorithm for Fingerprint Recognition. *Proceedings of IEEE 35th International Carnahan Conference on Security Technology*, pp. 264-266, ISBN 0-7803-6636-0, October 2001, IEEE, London, UK.
- Gonzalez, R. C.; Woods, R. E. & Eddins, S. L. (2009). *Digital Image Processing Using Matlab*, Second Edition, Gatesmark Publishing, ISBN 978-0982085400.
- Green, R. J. & Fitz, A. P. (1996). Fingerprint Classification using a Hexagonal Fast Fourier Transform. *Pattern Recognition*, Vol.29, No.10, (October, 1996), 1587-1597.
- K. Jain, A.; Flynn, P. & A. Ross, A. (2007). *Handbook of Biometrics*, Springer, ISBN 978-0387710402.
- K. Jain, A. & Minut, S. (2002). Hierarchical Kernel Fitting for Fingerprint Classification and Alignment. *Proceedings of IEEE 16th International Conference on Pattern Recognition (ICPR'02)*, p. 20469, ISBN 0-7695-1695-X, August 2002, IEEE, Quebec City, QC, Canada.
- K. Jain, A.; Prabhakar, S. & Hong, L. (1999). A Multichannel Approach to Fingerprint Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21, No.4, (April, 1999), 348-359.
- Lumini, A.; Maio, D. & Maltoni, D. (1997). Continuous versus Exclusive Classification for Fingerprint Retrieval. *Pattern Recognition Letters*, Vol.18, No.10, (October, 1997), 1027-1034.
- Maltoni, D. & Maio, D. (1996). A Structural Approach to Fingerprint Classification. *Proceedings of 13th international Conference on Pattern Recognition (ICPR)*, ISBN 0-8186-7282-X, August 1996, IEEE Computer Society, Vienna, Austria.
- Maltoni, D.; Maio, D.; K. Jain, A. & Prabhakar, S. (2009). *Handbook of Fingerprint Recognition*, Second Edition, Springer Verlag, ISBN 978-1848822535.
- Manhua, L. (2010). Fingerprint classification based on Adaboost learning from singularity features. *Pattern Recognition*, Vol.43, No.3, (March, 2010), 1062-1070.
- Park, C. H. & Park, H. (2005). Fingerprint classification using fast Fourier transform and nonlinear discriminant analysis. *Pattern Recognition*, Vol.38, No.4, (April, 2005), 495 - 503.
- Sarbadhikari, S.; Basak, J.; Pal, S. & Kundu, M. (1998). Noisy Fingerprints Classification with Directional Based Features Using MLP. *Springer-Verlag London Limited, Neural Computing & Applications*, Vol.7, (1998), 180-191.
- Senior, A. (2001). A Combination Fingerprint Classifier. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.10, (October, 2001), 1165 - 1174.
- Sherlock, B. G.; Monro, D. M. & Millard, K. (1994). Fingerprint Enhancement by Directional Fourier Filtering. *IEE Proceedings of Vision, Image, and Signal Processing*, Vol.141, No.2, (May, 1994), 87-94.
- Wang, W.; Li, J. & Chen, W. (2006). Fingerprint Classification Using Improved Directional Field and Fuzzy Wavelet Neural Networks. *Proceedings of the IEEE Sixth World Congress on Intelligent control and Automation*, pp. 9961-9964, January 2006, IEEE, Dalian, China.
- Yao, Y.; Marcialis, G. L.; Pontil, M.; Frasconi, P. & Roli, F. (2003). Combining Flat and Structured Representations for Fingerprint Classification with Recursive Neural Networks and Support Vector Machines. *Pattern Recognition*, Vol.36, 2003, 397- 406.

Dental Biometrics for Human Identification

Aparecido Nilceu Marana¹, Elizabeth B. Barboza², João Paulo Papa³,
Michael Hofer⁴ and Denise Tostes Oliveira⁵

^{1,2,3}*Dept. of Computing, School of Sciences, São Paulo State University (UNESP), Bauru*

⁴*Institute of Medical Engineering, Graz University of Technology*

⁵*Bauru School of Dentistry, University of São Paulo (USP), Bauru*

^{1,2,3,5}*Brazil*

⁴*Austria*

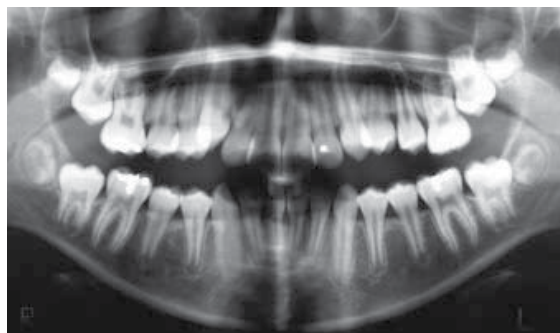
1. Introduction

In the history of civilization, the human identification based on dental information was first reported in the Roman Empire, when Nero's mother, Agripina, ordered the killing of Loilla Paulina, who was later identified by her dental caries and bad dental occlusion (Couto, 2009). The first treatise on human identification using dental records was conducted, in 1897, by Dr. Oscar Amoedo Valdés (1863-1945), a Cuban doctor, president of the French Dental Society and professor of the Paris Dental School, who applied a dental-based identification technique in order to reveal the identity of victims of a disaster which occurred in Paris (Amoedo, 1897).

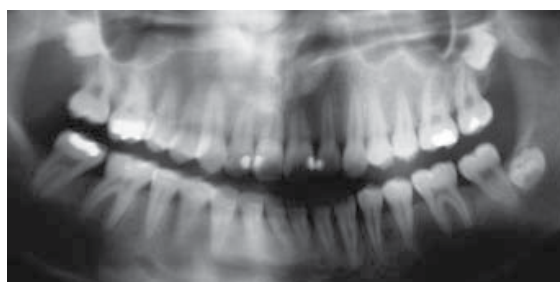
Since Dr. Amoedo's work, the Forensic Dentistry has attracted much attention, and the importance of using dental records for human identification is nowadays accepted worldwide (Chen & Jain, 2005). During the last decade, dental records have been extensively used in order to identify the victims of massive disasters, such as the 9/11 terrorist attack in New York (O'Shaughnessy, 2002) and the tsunami in Asia (Thepgumpanat, 2005).

In Forensic Dentistry, the human experts perform manual comparisons of ante-mortem (AM) and post-mortem (PM) dental records, looking for similarities (Jain & Chen, 2004). During this manual approach, the main characteristics used to compare dental records are: the presence or absence of a specific tooth, the morphology and dental restoration of the teeth, periodontal tissue characteristics, pathologies and other anatomical features. Figure 1 shows panoramic radiographs of two distinct individuals. One can easily observe several differences between these radiographs. In contrast to other popular biometric characteristics, dental features do change over time, causing great difficulties during the identification task. The teeth can change in appearance as a result of dental restorations, or can be missing altogether due to an accident which occurred after the AM records were taken. For this reason, although accepted in courts of law, dental based identification is considered less reliable than other biometric methods (Jain et al., 2003). Figure 2 shows panoramic radiographs of the same individual taken at two separate occasions during a period of three years. The dental patterns are almost the same, but changes due to dental restorations can be observed.

Despite known drawbacks of these identification methods, dental information may be the only available mean for identification in many disaster scenarios and mass accidents like fire and plane crashes. The other popular methods of identification are impossible since physical traits like faces and fingerprints are, in general, completely destroyed in such events



(a)



(b)

Fig. 1. Panoramic radiographs of two distinct adult individuals. It is possible to observe several details in teeth patterns that can be used to distinguish the individuals by their dental radiographs.

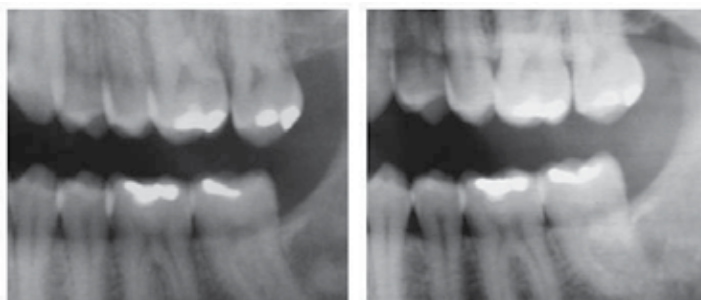


Fig. 2. Panoramic radiographs of the same person. (a) Acquired in the year 2000; (b) Acquired in the year 2003.

(Tang et al., 2009). The teeth and their dental restorations are very resistant to modest force effects and high temperatures. The teeth need a high temperature to be annealed and the components of the dental restorations likewise have a very high melting point. Savio et al. (2006) carried out an in-vitro experiment where, after exposing teeth to fire, periapical radiographs of all the teeth were taken. They reported a number of significant radiographic characteristics of the teeth were conserved after the exposure: the composite fillings were in place maintaining the shape up to 600°C (1112°F), the amalgam fillings were in

place maintaining the shape up to 1000°C (1832°F) and the endodontic treatments were recognizable up to 1100°C (2012°F).

Even when the person has never undergone dental treatment, the Forensic Dentistry can provide additional important information for the identification of the victim, such as species, racial group, gender, age, height, possible profession, and other private information which can facilitate police investigations.

When adult dentition is complete, dental-based identification process can provide high recognition rates, since no two individuals share the same teeth structure and characteristics. However, despite its accuracy, the traditional manual dental records comparison method demands too much time, and is not applicable in large scale identification, like mass disasters. Therefore, the development of techniques and systems that facilitate human identification through automated teeth recognition has become a necessity.

>From the Pattern Recognition and Computer Vision point of view, the problem of person identification based on dental records can be defined as an image matching and retrieval problem. That is, given an input dental image (usually a PM radiograph), the system searches the database in order to find the best matching AM radiograph (Jain et al., 2003).

The goal of this chapter is to introduce the problem of human identification based on dental biometrics, to summarize several techniques proposed in the specialized literature for automated dental recognition, to describe in detail an original method for dental recognition based on a new biometric descriptor, called dental code, and to propose a new method for dental recognition using the Image-Foresting Transform (Falcao et al., 2004) and the Shape Context (Belongie et al., 2000).

2. Automated Dental Identification Systems

Recent mass disasters, like the 9/11 terrorist attack and the Asian Tsunami, have highlighted the significance of automated dental identification systems (ADIS). In both these disasters, many victims were identified only by parts of their jaw bones. In the Asian tsunami, for instance, about 75% of the victims were identified using dental records, compared to just 0,5% victims which were identified using DNA (NewScientists, 2005). However, since the method of manual identification was used, it took several months to identify only a small part of the victim groups (only 20% of the 9/11 attack victims were identified in the first 12 months, and only 1,15% of the Asian tsunami victims were identified in the first 9 months). Therefore, we can conclude that manual dental identification is a very efficient post-mortem identification tool, but is also a time consuming process.

Besides being a humanitarian issue, a fast and precise post-mortem human identification is also crucial in solving problems related to heritage, proprietorship, insurance policies, pension charging, etc. Thus, the development of Automated Dental Identification Systems (ADIS) is a necessity. Automating dental identification methods will enhance the process of human identification in catastrophic events where the use of biometric identifiers such as face and fingerprints may not be possible (Abaza et al., 2009).

According to Fahmy et al. (2004), a typical architecture of an ADIS is composed of three main components: dental record preprocessing, search and retrieval, and image comparison. Figure 3 illustrates main phases of a person identification system based on dental records. In the first phase, the query radiograph is preprocessed in order to enhance its contrast, remove its noises, and select the areas of interest. The segmentation of the teeth and the normalization of the image regarding discrepancies in scale, rotation and illumination are also carried out at this stage. Next, a template (or model) image is retrieved from the database and is registered with the query image, for matching. In the following, decision making phase, the features

extracted from the teeth on both images are compared by using a proper distance function. The system output is, in general, a score proportional to the probability of both radiograph images being of the same individual.

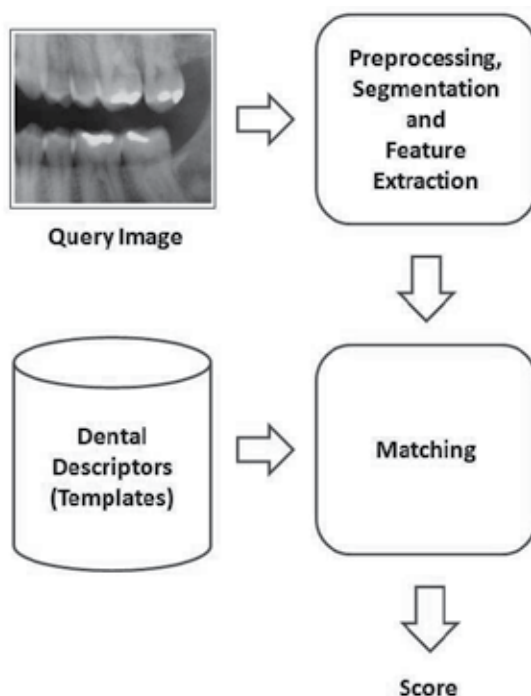


Fig. 3. Typical ADIS components.

Chen & Jain (2005) proposed an ADIS that has two main stages: feature extraction and matching. The feature extraction stage uses anisotropic diffusion to enhance the images and a Mixture of Gaussians model to segment the dental restorations. The matching stage has three sequential steps: tooth-level matching, computation of image distances, and subject identification. In the tooth-level matching step, tooth contours are matched using a shape registration method, and the dental restorations are matched on overlapping areas. The distance between the tooth contours and the distance between the dental restorations are then combined using posterior probabilities. In the second step, the tooth correspondences between the given query (post-mortem) radiograph and the database (ante-mortem) radiograph are established. A distance based on the corresponding teeth is then used to measure the similarity between the two radiographs. Finally, all the distances between the given post-mortem radiographs and the ante-mortem radiographs that provide candidate identities are combined to establish the identity of the subject associated with the post-mortem radiographs.

Zhou & Abdel-Mottaleb (2005) presented a system to assist human identification using dental radiographs. The goal of their system is to archive ante-mortem (AM) dental images and enable content-based retrieval of AM images that have similar teeth shapes to a given post-mortem (PM) dental image. During archiving, the system classifies the dental images to bitewing, periapical, and panoramic views. It then segments the teeth and the bones in the bitewing images, separates each tooth into the crown and the root, and stores the contours

of the teeth in the database. During retrieval, the proposed system retrieves from the AM database the images with the teeth most similar to the PM image based on Hausdorff distance measure between the teeth contours.

Abaza et al. (2009) considered archiving and retrieving dental records from large databases to be a challenging task which has not received adequate attention in the literature. Therefore, they propose an efficient method for retrieving dental records from a database in order to assist the forensic expert in identifying deceased individuals in a rapid manner. The proposed method is an appearance-based technique that consolidates the evidence presented by individual teeth in a dental record, that is, it moves from tooth-to-tooth in order to render a record-to-record matching score. The proposed method is shown to reduce the searching time of record-to-record matching by a factor of a hundred.

Hofer & Marana (2007) proposed a method for human identification based on dental restorations observed in panoramic radiographs. The proposed method has three main processing steps: automatic segmentation of dental restorations, creation of a dental code, and matching. In the segmentation step, seed points of the dental restorations are detected by thresholding. The final segmentation is obtained with a snake (active contour) algorithm. The dental code is defined from the location and size of the dental restorations, and the distance between them. The matching stage is performed with the generalized edit distance (Levenshtein distance) (Navarro, 2001). Details of this method are presented next.

3. Proposed method for human identification based on dental restorations information

The new method proposed by Hofer & Marana (2007) for human identification based on dental restorations consists of three main steps: preprocessing of the dental radiographs and segmentation of the dental restorations (DRs); creation of a dental code (DC) out of the information of the detected DRs including the size, the location and the distance between them; and matching of a query DC with a template DC taken from the database.

3.1 Preprocessing and Segmentation

The dental radiograph image (RGB image) is converted into a gray-scale image and a median filtering is performed to reduce noise. Because of different contrast conditions in the dental radiographs, the image is subdivided into 2 regions of interest (ROIs), as illustrated in Figure 4.

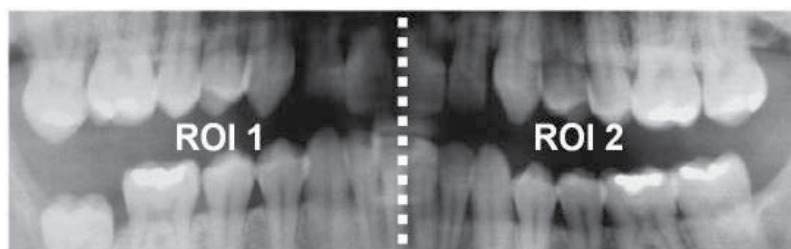


Fig. 4. ROIs in a dental radiograph, left part (ROI 1), right part (ROI 2).

The algorithm determines a gray value threshold in the left and in the right ROIs of the dental radiograph. Typically, the DRs feature the highest intensities in the image and appear as a distinct, relatively small but pronounced mode in the upper range of the gray-scale histogram. After smoothing the histogram with a moving average filter, the threshold is set

to the gray-value at the location of the left valley at the rightmost mode, which indicates the DRs, as shown in Figure 5.

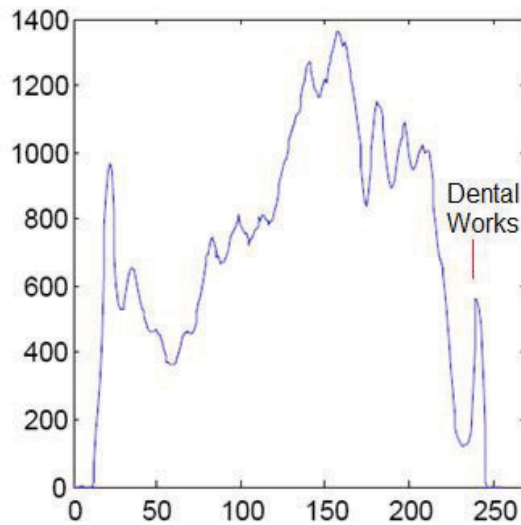


Fig. 5. Smoothed histogram of the ROI 2 in Figure 4 (threshold value is 232).

The threshold value is used to binarize the gray-level image. The results of the conversion are used to initialize the contour for the segmentation method. Each region represents a possible DR.

A snake (active contour) algorithm, as defined by Xu & Prince (1998), is used to perform the final segmentation of the DRs. Snakes can be used to segment objects with fuzzy border contours where traditional edge-detection (Ziou & Tabbone, 1998) will fail. Snakes are curves that can move under the influence of internal forces (elasticity and bending forces) coming from within the curve itself, and external forces (potential forces) computed from the image data. The internal and external forces are defined so that the final snake will conform to an object boundary (Xu & Prince, 1998). The external force field is computed from the gradient image. A snake needs to be initialized with an initial curve (e.g. circle) and is an iterative procedure which stops after a defined number of iterations. The better the initialization curve, the better the performance of the algorithm and the final segmentation results.

Each DR is segmented with a separate snake. To improve the segmentation and to speed up the algorithm, the initial curves for all DRs are computed from the binary mask. The borders of the detected regions are used as initial curves. The evaluation of the snake is shown in Figure 6.

Finally, a binary mask of the image including all detected DRs is created. This mask is called a dental restorations mask (DRM).

3.2 Creation of the Dental Code

Based on the DRM, a dental code (DC) is created. The DC incorporates information about the location, the size of the DRs and the distance between two DRs next to each other.

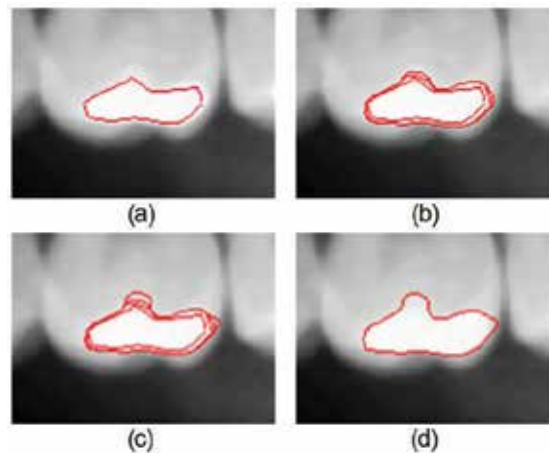


Fig. 6. Segmentation process. (a) Initial curve; (b) Curve transformation after 2 iterations; (c) After 5 iterations; (d) Final segmentation result after 30 iterations.

3.2.1 Locations of the dental restorations

An algorithm was implemented to sort all DRs into the DRM moving from left to right, based on the leftmost pixel of each individual DR, as illustrated in Figure 7.



Fig. 7. Dental restorations mask with sorted dental restorations from left to right.

An important factor for the creation of the DC is whether the tooth on which the DR is located belongs to the upper or to the lower jaw. Therefore, a border between the upper and the lower row of teeth has to be defined. A stripe in the intensity image is cut with the width of the current region. Next, the intensity sum of all horizontal rows in the stripe is calculated. The highest intensity represents the area of the DR, as illustrated in Figure 8.

The algorithm detects the first valley on the left and on the right side of the highest intensity point. The valley with the lower intensity represents the border between the upper and the lower row of teeth. If the position of this valley is above the DR in the image, the DR belongs to the lower row of teeth. If the position is below the DR, the DR belongs to the upper row of teeth. The locations of the dental restorations are represented in the DC with the letter "L" or "U" (Letter "L": DR on the lower row of teeth, Letter "U": DR on the upper row of teeth).

3.2.2 Size of the dental restorations

The proposed method uses registered dental radiographs images which are all resized to be the same size. Consequently, the amount of pixels in a dental radiograph is always the same, which means that the size of a DR (amount of pixels) is a percentage of the total amount of pixels in the dental radiograph.

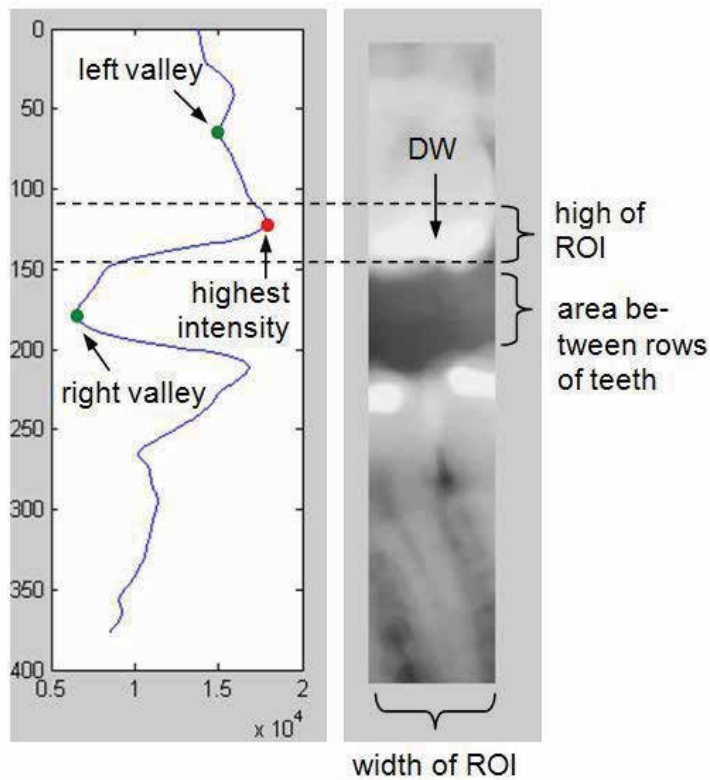


Fig. 8. Cut stripe and sum of intensities; right valley represents lower intensity which indicated that the DR belongs to the upper row of teeth, (dental code = "U").

3.2.3 Distance between two dental restorations

To make the matching algorithm more sensitive, the distance (amount of pixels) between the center of the mass point of a DR and its left neighbor is also included in the DC, as illustrated in Figure 9. The distance of the leftmost DR is set to zero. The value for the distance is given in percentage of the total width of the dental radiograph.



Fig. 9. Distances between dental restorations.

The DC is created out of this information as follows: 1) the position of the DR ("L" or "U"); 2) the size of the DR; 3) the distance between the current DR and the previous DR (from left to right). An example of a finalized DC for a panoramic dental radiograph is shown in Figure 10.

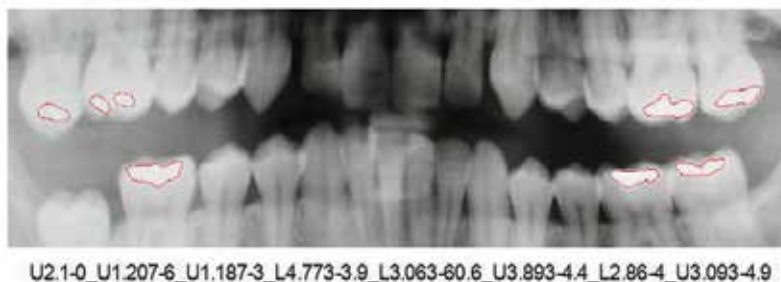


Fig. 10. Dental code (DC) obtained from a panoramic dental radiograph.

3.3 Matching

After the DC is created, it can be compared to other DCs in a database. These can be different codes of the same person, due to possible changes in dental restorations, as illustrated in Figure 2, or codes of different subjects. Matching between radiographs of the same subject is called "genuine matching" and matching between radiographs belonging to two different subjects is called "impostor matching". An algorithm was implemented based on the generalized edit distance (Levenshtein distance). The edit distance between two strings is given by the minimum number of operations needed to transform one string into the other, where an operation is an insertion, a deletion, or a substitution (Navarro, 2001).

In the edit distance, every edit operation is associated with certain costs. For instance, to transform the string *hitten* into *sitting*, it is necessary to substitute 'h' for 's', 'e' for 'i', and finally, to insert 'g' at the end. The overall Edit distance cost in this example is 3, using the same cost for each operation. It is also possible to change the cost of different operations (insertion, deletion, or substitution).

The edit distance has received a lot of attention because its generalized version is powerful enough for a wide range of applications (Navarro, 2001).

Because of the structure of the DC, it was necessary to adjust the edit distance costs. Not only the letters "U" and "L" have to be compared, but also the size of the DRs and the distance between two DRs. The costs of the insertion, deletion, and substitution were changed to improve the results of the matching algorithm. In the case of insertion and deletion, the matching cost is 60 (see Table 1). In the case of substitution the cost for comparing two DRs is given by the sum of two costs: the cost of comparing the size, and the cost of comparing the distance between two DRs. If the compared sizes differ more the 100%, the cost is set to 25. Otherwise, the cost for comparing the size is set according to the percentage difference of the two compared DRs (see Table 1). If the compared distances differ more than 15%, the cost is set to 25. Otherwise, the cost for comparing the distances is set according to the percentage difference of compared DRs (see Table 1).

3.4 Assessment of the proposed method

In order to assess the proposed dental biometric method, some experiments were carried out on a database including 68 panoramic dental radiographs: a pair of panoramic radiographs for each of 22 adult subjects (44 panoramic radiographs) plus a single panoramic radiograph

Operation	Cost
Insertion	60
Deletion	60
Substitution	0, difference between 0 ... 10%
Comparing	1, difference between 10 ... 20%
Size	⋮
	10, difference between 90 ... 100%
	25, difference > 100%
Substitution	0, difference between 0 ... 1%
Comparing	1, difference between 1 ... 2%
Distances	⋮
	15, difference between 14 ... 15%
	25, difference > 15%

Table 1. Edit distance costs for insertion, deletion and substitution.

for each of other 24 subjects. For the purpose of the study, the older radiographs of the 22 subjects with two panoramic radiographs were considered AM (ante-mortem) while their newer radiographs were considered PM (post-mortem). The radiographs of the 24 subjects with only one panoramic radiograph were considered AM.

In the experiments, the images were manually registered to obtain comparable conditions. In cases of over-segmentation or under-segmentation of the DRs, the segmentation results were manually corrected. Likewise, if a DR could not be detected by thresholding, a ROI was manually selected in the radiograph to perform local thresholding. Segmentation results and their corresponding DCs for two panoramic radiographs are shown in Figure 11.

In order to test the matching performance of the method, an algorithm was implemented to compare panoramic radiographs of the genuine class (two radiographs of the same person) and panoramic radiographs of the impostor class (two radiographs of different persons). Figure 12(a) shows a receiver operating characteristic (ROC) curve, which plots the false acceptance rate (FAR) versus the false rejection rate (FRR) for different threshold values. The ROC curve shows that the proposed method obtained 11% of equal error rate (EER) for the used database, which is a good result. The equal error rate (EER) is the value where the FAR is equal to the FRR. The lower the EER, the better the performance of the biometric system. Figure 12(b) shows the accuracy curve obtained when the 22 PM radiographs were matched to the 46 AM radiographs from the database. Using the top-1 retrieval, the accuracy was 19/22 (= 86%). Using top-8 retrievals, the retrieving accuracy was 95%. The accuracy reached 100% when the top-11 retrievals were used.

4. Dental recognition based on image-foresting transform and shape context

The assessment of the proposed method for human identification based on dental restorations information (Section 3) brought us to a conclusion that the segmentation step is crucial. Based on the good results obtained by Falguera et al. (2008), who used other biometrics characteristics from radiographs, such as the frontal sinus, we opted to use the Differential Image-Foresting Transform (DIFT), proposed by Falcao & Bergo (2004), for dental segmentation. The main idea is to design an interactive ADIS, where the user has to indicate only a few points (seeds) inside and outside of the dental restorations.



Fig. 11. Segmentation results of two dental radiographs including their dental codes.

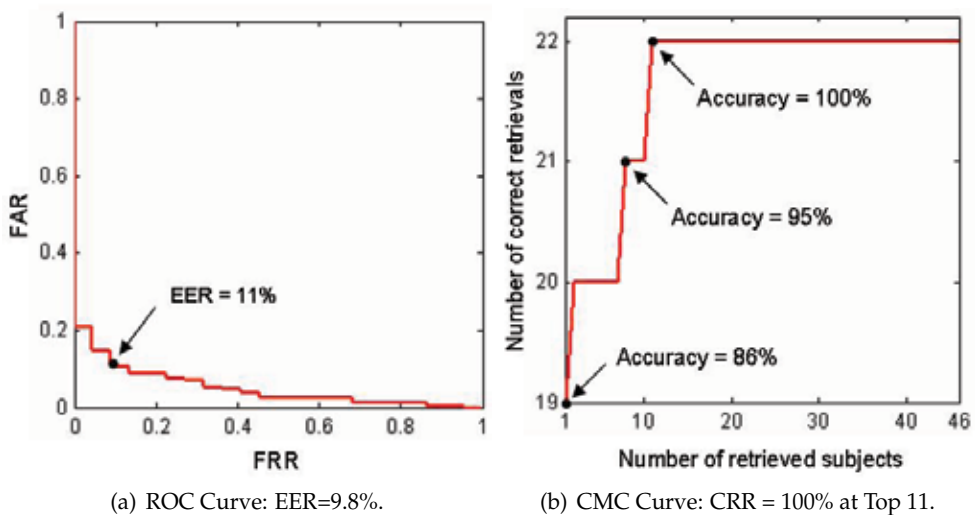


Fig. 12. Results obtained with the proposed method based on dental restorations recognition.

Figure 13 presents the dental restorations segmentation obtained from a panoramic radiograph using the DIFT-based segmentation algorithm. One can observe that the algorithm obtained very good dental restorations segmentations with minimal user intervention using only a few external (white spots) and internal seeds (yellow spots).

Also, in order to improve the recognition rates, we propose the fusion of the dental restorations based method with a more precise method based on teeth shapes segmented

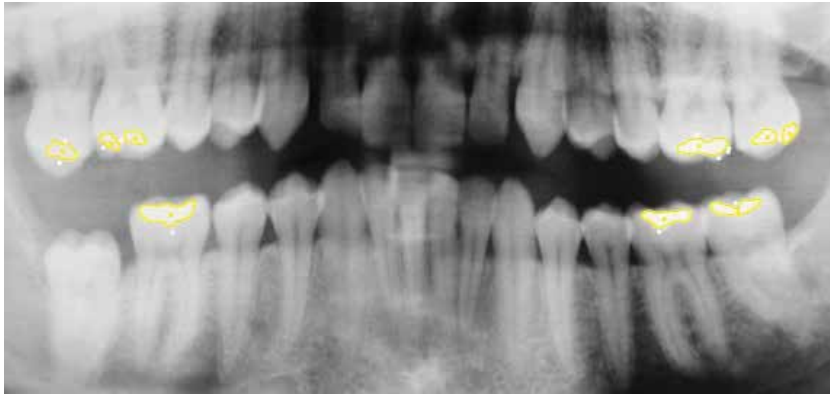


Fig. 13. Automated dental restorations segmentation using the DIFT-based segmentation algorithm. White spots: external seeds. Yellow spots: internal seeds.

in the panoramic radiograph. In this multibiometrics method, we compute the query and template radiographs matching score as a weighted sum of dental restorations matching score and the teeth shapes matching score. The teeth shapes matching score is computed by using the shape context method, proposed by Belongie et al. (2000).

Figure 14 shows the interface of our interactive ADIS, during the matching stage, in which the shape of a molar tooth from the query radiograph is being compared with the shape of the same molar tooth, but from a template radiograph.

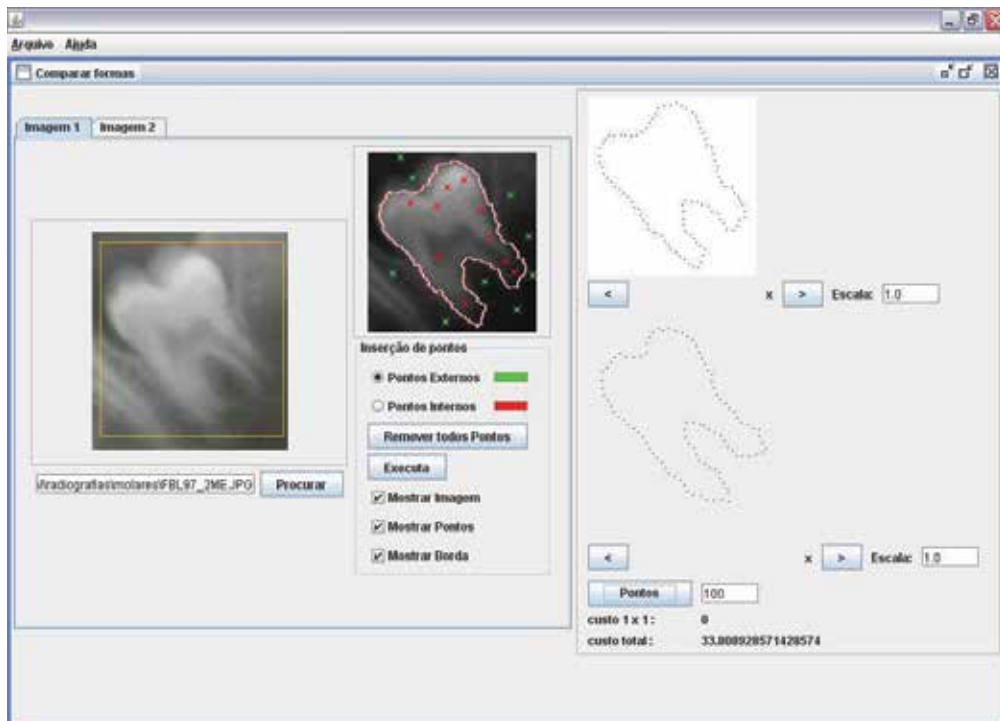


Fig. 14. Proposed ADIS (Automated Dental Identification System) for dental recognition.

In the following sections, the DIFT and the Shape Context methods for segmentation and shape comparison are summarized respectively.

4.1 Differential image-foresting transform segmentation method

The image segmentation algorithm based on Differential Image Foresting Transform (DIFT) was proposed by Falcao & Bergo (2004). This algorithm reduces the image segmentation problem to the calculation of minimum cost path forests in the graph derived from the image. The search for minimum cost paths is restricted to paths that have origin in a set of initial pixels called seeds.

The function that determines the cost of the path between two nodes of the graph is the difference between the gray levels of the pixels that are represented by graph nodes. This way, pixels with similar gray levels will have smaller path costs and will tend to be connected to each other, originating plateaus in the image.

Therefore, using this algorithm, the human intervention during the image segmentation is reduced to the choice of a few seeds (pixels) inside and outside of the object to be segmented. Then, the seeds will compete for the graph nodes in each interaction, creating, in the end, internal and external regions related to the object of interest, in which the terminal pixels delimitate the object (Miranda, 2006).

Figure 15(a) presents a graph of a gray level two-dimensional image, in a 4-neighborhood system. The numbers correspond to the intensities $I(p)$ of the pixels and the biggest circles represent two seeds: one inside and the other outside the object which is to be segmented (in this example, a rectangle in the center of the image) (Miranda, 2006).

Figure 15(b) shows the minimum cost path forest, obtained using the cost function $C(p, q) = |I(q) - I(p)|$, where $I(p)$ is the intensity of a pixel and $I(q)$ is the intensity of the preceding pixel. The shown numbers are the values of the calculated costs. The segmentation of the rectangle at the center of the image is obtained by the tree rooted in the internal seed. The arrows points to the predecessor of each node in the optimum path (Miranda, 2006).

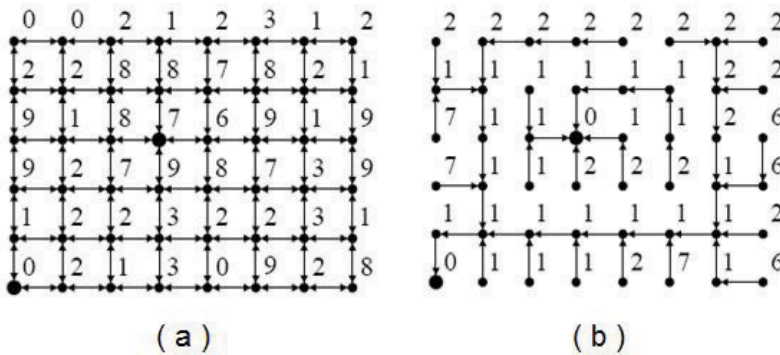


Fig. 15. (a) Graph of a gray level 2D image in a 4-neighborhood. (b) Minimum cost path forest obtained for the graph under (a) (Miranda, 2006).

In case of a non-satisfactory segmentation, the user can add more inside or outside seeds and re-run the algorithm. Instead of calculating a new minimum cost path forest every time the seed set is changed, the DIFT-based segmentation algorithm only updates the segmentation result in a differential way, within a time frame proportional to the number of nodes inside the changed regions of the forest, reducing the execution period up to ten times compared to the original IFT-based segmentation algorithm, proposed in (Miranda, 2006).

4.2 Shape context description method

Shape Context is a method for shape description proposed by Belongie et al. (2000). This description method calculates the polar-logarithmic distribution of shape border points in relation to a certain point p , which belongs to the shape border. This provides a global and discriminative characterization of the shape. Thus, the corresponding border points of two similar shapes will have similar shape contexts. The similarity measurement between two objects is computed as being the minimum cost of alignment among the border points of their shapes. The more different the shapes of the objects, the greater their alignment cost.

Figure 16(a) presents the polar-logarithmic histogram representation used to calculate the shape context for each point of the shape border in respect to the other points. In this example, the histogram is composed by 60 bins (12 angles and 5 distances). Figure 16(b) shows the polar-logarithmic histogram obtained for the border point of the shape selected in the Figure 16(a).

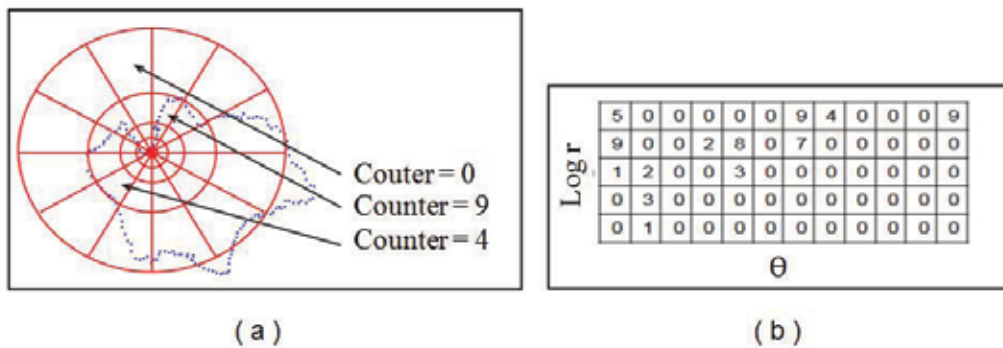


Fig. 16. (a) Illustration of the log-polar histogram computation for one shape; (b) Shape context histogram computed for the shape point shown in (a).

Figure 17 illustrates a comparison between histograms of three distinct points of two shapes. One can note that the two points located at similar regions (top-left) of the border present similar histograms. On the other hand, the third point, located at a different location (bottom-right) presents a very different histogram. To make the visualization easier, the values of the histograms were mapped into gray levels: the darker the position at the histogram, the greater the occurrence of border points at that angle and distance.

The similarity between two shapes is computed as the minimum cost of alignment obtained for them, using the shape context descriptors for every pixel belonging to each shape. The smaller the cost, the larger the similarity score between the shapes. Given the pixels p_i and p_j , belonging to each border shape that are being compared, once the shape contexts of each pixel are distributions represented by histograms, it is possible to utilize the statistic test χ^2 , given by equation 1, to calculate the matching cost C_{ij} between p_i and p_j , where $h_i(k)$ and $h_j(k)$ denote the normalized histograms of p_i and p_j , respectively, and K represents the number of histogram bins.

$$C_{ij} = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (1)$$

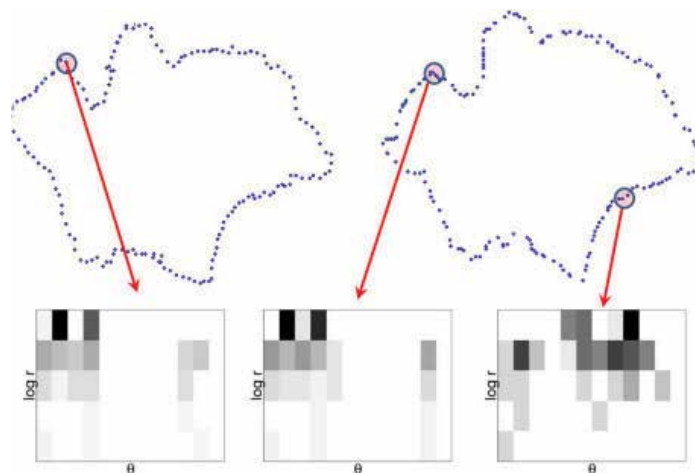


Fig. 17. Comparison of histograms of three distinct points on two shapes obtained from the same source.

5. Conclusion

In Forensic Dentistry, human experts generally perform manual comparisons of ante-mortem and post-mortem dental records, looking for differences among teeth patterns. Despite its accuracy, this manual comparison method demands great amount of time, which is not admissible in large scale identification. Therefore, the development of techniques and systems that facilitate human identification through automated teeth recognition is required.

In the last years, some ADIS (Automated Dental Identification Systems) have been proposed in the literature. In our research in this area, we have proposed an ADIS based on dental restorations recognition and teeth shapes analysis. The results obtained with the dental restorations features can be considered good, when the high variability of this biometric characteristic is observed. To cope with this high variability, we proposed a very promising distance function based on the generalized edit distance.

In order to improve the recognition rates, we also proposed the fusion of a teeth restorations based method and a teeth shapes based method, in which the teeth shapes are extracted from panoramic radiographs images using a Differential Image-Foresting Transform (DIFT) segmentation algorithm and the shape analysis is carried out using the Shape Context method. The DIFT segmentation algorithm and the Shape Context method were employed by the authors for another kind of biometrics trait, the frontal sinus, providing very good results. The same good performance level is expected to be reached in the application of these methods on panoramic dental radiographs. Results obtained from the fusion with both methods will be published soon.

6. References

- Abaza, A., Ross, A. & Ammar, H. (2009). Retrieving dental radiographs for postmortem identification, *Proc. of IEEE International Conference on Image Processing (ICIP)*.
- Amoedo, O. (1897). The role of the dentists in the identification of the victims of the catastrophe of the 'bazar de la charité', paris, 4th of may, 1897, *The Dental Cosmos* 39: 905–12.

- Belongie, S., Malik, J. & Puzicha, J. (2000). Shape context: A new descriptor for shape matching and object recognition, *Neural Information Processing Systems Conference (NIPS)*, pp. 831–837.
- Chen, H. & Jain, A. K. (2005). Dental biometrics: Alignment and matching of dental radiographs, *IEEE Trans. Pattern Anal. Mach. Intell.* 27(8): 1319–1326.
- Couto, M. I. S. B. G. (2009). *Importância dos Registos Dentários em Situações de Grandes Catástrofes*, Master Dissertation in Forensic Medicine, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Portugal.
- Fahmy, G., Nassar, D., Haj-Said, E., Chen, H., Nomir, O., Zhou, J., Howell, R., Ammar, H. H., Abdel-Mottaleb, M. & Jain, A. K. (2004). Towards an automated dental identification system (adis), in D. Zhang & A. K. Jain (eds), *Biometric Authentication*, Vol. 3072 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 1–41.
- Falcao, A., Stolfi, J. & de Alencar Lotufo, R. (2004). The image foresting transform: theory, algorithms, and applications, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26(1): 19 – 29.
- Falcao, A. X. & Bergo, F. P. G. (2004). Interactive volume segmentation with differential image foresting transforms, *IEEE Trans. on Medical Imaging* 23(9): 1100–1108.
- Falguera, J. R., Falguera, F. P. S. & Marana, A. N. (2008). Frontal sinus recognition for human identification, *SPIE*, p. 69440S.
- Hofer, M. & Marana, A. N. (2007). Dental biometrics: Human identification based on dental work information, *Proc. XX Brazilian Symposium on Computer Graphics and Image Processing, 2007 (SIBGRAPI 2007)*., pp. 281 –286.
- Jain, A. K. & Chen, H. S. (2004). Tooth contour extraction for matching dental radiographs, *Proc. 17th ICPR, vol. III*, Cambridge, UK, pp. 522–525.
- Jain, A. K., Chen, H. S. & Minut, S. (2003). Dental biometrics: Human identification using dental radiographs, *Proc. 4th Int. Conf. on Audio and Video-Based Biometric Person Authentication (AVBPA)*, Guildford, UK, pp. 429–437.
- Miranda, P. A. V. (2006). *Segmentação de Imagens pela Transformada Imagem-Floresta*, Master Dissertation, Institute of Computing, UNICAMP, Campinas, Brazil.
- Navarro, G. (2001). A guided tour to approximate string matching, *ACM Comput. Surveys* 33(1): 31–88.
- NewScientists (2005). Dental records beat dna in tsunami ids.
URL: [NewScientists.com News Service](http://www.alertnet.org)
- O’Shaughnessy, P. (2002). More than half of victims idd, *New York Daily News* .
- Savio, C., Merlati, G., Danesino, P., Fassina, G. & Menghini, P. (2006). Radiographic evaluation of teeth subjected to high temperatures: Experimental study to aid identification processes, *Forensic Science International* 158(2): 108–116.
- Tang, J. P., Hu, D. Y., Jiang, F. H. & Yu, X. (2009). Assessing forensic applications of the frontal sinus in a chinese han population, *Forensic Science International* 183(1-3): 104.e1 – 104.e3.
- Thepgumpanat, P. (2005). Thai tsunami forensic centre produces first ids, *Reuters* .
URL: <http://www.alertnet.org/>
- Xu, C. & Prince, J. L. (1998). Snakes, shapes, and gradient vector flow, *IEEE Transactions on Image Processing* 7(3): 359–369.
- Zhou, J. & Abdel-Mottaleb, M. (2005). A content-based system for human identification based on bitewing dental x-ray images, *Pattern Recognition* 38(11): 2132 – 2142.
- Ziou, D. & Tabbone, S. (1998). Edge detection techniques - an overview, *International Journal of Pattern Recognition and Image Analysis* 8: 537–559.

Facial Expression Recognition

Bogdan J. Matuszewski, Wei Quan and Lik-Kwan Shark
ADSIP Research Centre, University of Central Lancashire
UK

1. Introduction

Facial expressions are visible signs of a person's affective state, cognitive activity and personality. Humans can perform expression recognition with a remarkable robustness without conscious effort even under a variety of adverse conditions such as partially occluded faces, different appearances and poor illumination. Over the last two decades, the advances in imaging technology and ever increasing computing power have opened up a possibility of automatic facial expression recognition and this has led to significant research efforts from the computer vision and pattern recognition communities. One reason for this growing interest is due to a wide spectrum of possible applications in diverse areas, such as more engaging human-computer interaction (HCI) systems, video conferencing, augmented reality. Additionally from the biometric perspective, automatic recognition of facial expressions has been investigated in the context of monitoring patients in the intensive care and neonatal units for signs of pain and anxiety, behavioural research, identifying level of concentration, and improving face recognition.

Automatic facial expression recognition is a difficult task due to its inherent subjective nature, which is additionally hampered by usual difficulties encountered in pattern recognition and computer vision research. The vast majority of the current state-of-the-art facial expression recognition systems are based on 2-D facial images or videos, which offer good performance only for the data captured under controlled conditions. As a result, there is currently a shift towards the use of 3-D facial data to yield better recognition performance. However, it requires more expensive data acquisition systems and sophisticated processing algorithms. The aim of this chapter is to provide an overview of the existing methodologies and recent advances in the facial expression recognition, as well as present a systematic description of the authors' work on the use of 3-D facial data for automatic recognition of facial expressions, starting from data acquisition and database creation to data processing algorithms and performance evaluation.

1.1 Facial expression

Facial expressions are generated ... skin texture" (Pantic & Rothkrantz, 2000)" should be replaced by "Expressions shown on the face are produced by a combination of contraction activities made by facial muscles, with most noticeable temporal deformation around nose, lips, eyelids, and eyebrows as well as facial skin texture patterns (Pantic & Rothkrantz, 2000). Typical facial expressions last for a few seconds, normally between 250 milliseconds and five seconds (Fasel & Luetten, 2003). According to psychologists Ekman and Friesen

(Ekman & Friesen, 1971), there are six universal facial expressions, namely: anger, disgust, fear, happiness, sadness, and surprise, as shown in Figure 1. Among these universal expressions, some, such as happiness, can be accurately identified even if they are expressed by members of different ethnic groups. Others are more difficult to recognise even when expressed by the same person.

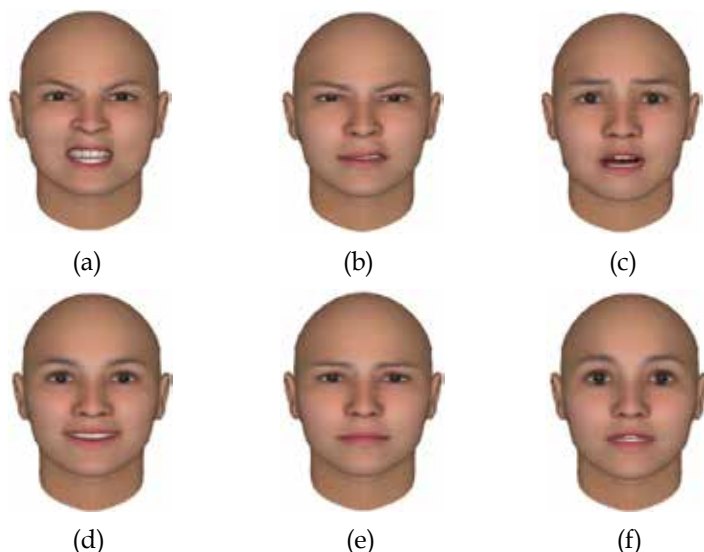


Fig. 1. Examples of six universal computer simulated expressions: (a) anger, (b) disgust, (c) fear, (d) happiness, (e) sadness, and (f) surprise (FaceGen, 2003).

In computer vision and pattern recognition, facial expression recognition is often confused with human emotion recognition. While facial expression recognition uses purely visual information to group facial articulation into abstract classes, emotion recognition is based on many other physiological signals, such as voice, pose, gesture and gaze (Fasel & Luetten, 2003). It is noteworthy to mention that emotions (or in general person's mental state) are not the only cause of facial expressions. Facial expressions could also be manifestations of physiological activities and aid verbal and non-verbal communication. These, for example, may include physical symptoms of pain and tiredness, or listener responses during verbal communication. Therefore emotion recognition requires not only interpretation of facial expression but also understanding of the full contextual information.

1.2 Applications

Facial expression analysis can be traced back to the nineteenth century with Darwin's theory on the similarity of facial expressions across different cultures and between different species (Darwin, 1872). From then on, most of research was conducted by psychologists until 1978 with Suwa et al. (Suwa et al., 1978) presenting a preliminary investigation on automatic facial expression analysis from an image sequence. Currently automatic facial expression analysis and recognition have become an active research area associated with a wide range of applications, such as human-machine interaction, video conferencing, virtual reality, and biometrics.

Automatic recognition of facial expressions can act as a component of human-machine interface or perceptual interface (van Dam, 2000; Pentland, 2000). This kind of interface would be able to support the automated provision of services that require a good appreciation of the emotion from the user (Bartlett et al., 2003). For example, it could provide the user's intentions and feelings to a machine or robot to enable it to respond more appropriately during the service (Essa & Pentland, 1997). In the application of robot-assisted learning whereby a robot is used to teach the user by explaining the content of the lesson and question the user afterwards, understanding the human emotion will enable the robot to progress from one lesson to the next when the user is ready (Wimmer et al., 2008).

Video conferencing, tele-presence and tele-teaching require transmission of a large amount of data, and data compression is often needed in order to reduce the storage and bandwidth requirements. Facial expression analysis offers an approach to achieve data compression for video conferencing (Pearson, 1995; Aizawa & Huang, 1995). Using the video frames recorded by the cameras in front of the people attending the video conference, facial expressions and motions of each person can be estimated at the transmitting side as a set of parameters that describe the current appearance of each person, thereby reducing the amount of data to be transmitted. At the other side of the video conference, the set of parameters received is used to render the facial model to present the approximate appearance of each person (Eisert & Girod, 1998). The technique of automatic facial expression analysis used in video conferencing can also be applied to virtual reality to provide realistic synthesis of facial articulation (Morishima & Harashima, 1993).

In terms of biometrics, automatic expression recognition has been investigated in the context of monitoring patients in the intensive care and neonatal units (Brahnam et al., 2006) for signs of pain and anxiety, behavioural research on children's ability to learn emotions by interacting with adults in different social contexts (Pollak & Sinha, 2002), identifying level of concentration (Vural et al., 2007), for example detecting driver tiredness, and finally in aiding face recognition.

1.3 Challenges

Despite of the significant progress made in both research and application development, automatic facial expression recognition remains a particularly challenging problem (Wang & Yin, 2007; Bartlett et al., 2003). Broadly speaking, there are two major obstacles. One is related to a robust capture of facial expressions, and the other is associated with machine learning.

With facial data being the information source for the facial expression recognition task, the processing complexity and expression recognition performance are strongly dependent on the data capture technique used. Although simple and low cost, current expression recognition systems based on 2-D imaging are only able to achieve good recognition performance in constrained environments due to difficulties in handling large variations in illumination and view angle (Quan, 2009a). These difficulties arise from the fact that the human face is a three-dimensional surface rather than a two-dimensional pattern, resulting in its 2-D projection being sensitive to changes in lighting and head pose (Pantic & Rothkrantz, 2000). This has led to the increased use of 3-D facial data capture systems, since it is largely immune to changes in pose and illumination (Yin et al., 2006; Quan, 2009a). However, it needs to be stressed that the 3-D face acquisition does not solve all the problems as for example it does not help to alleviate issues associated with occlusions, where typical examples of facial occlusions include subjects wearing glasses, or having long hair.

The machine learning challenges are related to facial feature extraction and classification to achieve a high performance of the facial expression recognition. The extracted features should represent different types of facial expressions in a way which is not significantly affected by age, gender, or ethnic origin of the subject. The classification method must be capable to define appropriate rules in order to derive a specific type of facial expression from the facial features provided, even when the output from the preceding processing stages, such as facial data acquisition and facial feature extraction, is noisy or incomplete (Wimmer et al., 2008).

2. Facial expression recognition systems

In general, a typical facial expression recognition system, whether automatic or semi-automatic, consists of three main processing stages: acquisition, representation and classification. The general framework of a facial expression recognition system is illustrated in Figure 2 with each processing stage discussed in the following subsections. Facial data acquisition addresses how to detect, locate and acquire facial data in complex 2-D or 3-D scenes. Facial expression representation is concerned with the extraction of representative facial features for different types of expressions to give the required accuracy and robustness. Facial expression classification is tasked with finding a suitable classification algorithm to categorise facial expressions in terms of the facial features provided by the facial expression representation stage.

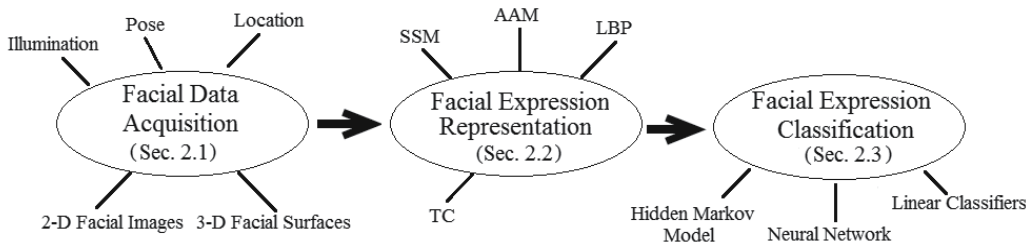


Fig. 2. A general framework of facial expression recognition system

2.1 Facial data acquisition

Facial data can be acquired in 2-D or 3-D, both in a static or dynamic mode. A static 2-D facial image represents a scene at a fixed moment in time and as such does not contain information about temporal evolution of the scene; a dynamic 2-D facial data set is a time ordered collection of images and therefore can provide information about expression evolution which can aid recognition. For faces appeared in complex scenes with cluttered backgrounds, face detection algorithms are required to locate the facial area in each image (Heisele et al., 2001; Jones & Viola, 2003), since most of the face expression analysis methods need the exact position of the face in order to extract facial features of interest (Lanitis et al., 1997; Hong et al., 1998; Steffens et al., 1998). Furthermore, the captured facial images usually need to be filtered so as to reduce the lighting variation as well as the specular reflection on eyes, teeth and skin (Black et al., 1998). Often for the subsequent feature extraction some assumptions need to be made about face appearance, these may include size, face orientation and intensity variations mentioned above. If the actual detected face does not

fulfil these restrictions, e.g. it is too small, it needs to be re-rendered. Such process is called face normalisation and is often applied after the face detection step and before the feature extraction stage (Flanelli et al., 2010). Interestingly some authors proposed to use 3-D cues for facial image normalisation even though the actual recognition process is based on 2-D information (Niese et al., 2007).

A static 3-D facial data set is usually represented as a set of 3-D points or surface patches. Such data is normally captured by 3-D imaging systems (often called 3-D scanners). They scan a real-world object generating a geometric point cloud corresponding to samples taken from the observed 3-D surface. Apart from surface geometry, such 3-D scanners can often provide information about the corresponding 3-D point appearance e.g. colour. In general, there are two major types of 3-D scanners, contact and non-contact (Curless, 2000). Contact 3-D scanners, used mostly in manufacturing, are seldom used for facial expression analysis. Despite the fact that they provide accurate surface measurements, they require prohibitively long acquisition time. Non-contact scanners are much more suitable for 3-D facial data acquisition. They can be further divided into two broad categories of active and passive scanners.

Active scanners measure object surface by emitting a light pattern and detecting its reflection (Zhang & Huang, 2006). Active scanners commonly found in applications of facial data analysis use structured or random patterns which are projected on the face surface. The surface reconstruction is based on detected geometrical distortions of the pattern. Such scanners often include one or more cameras and a projector.

Passive 3-D scanners use the principles of multi-view geometry (Bernardini & Rushmeier, 2002) utilising information from multiple cameras placed around the face surface without emitting any kind of radiation. The most popular type of passive scanners uses two cameras to obtain two different views of the face surface. By analysing the position differences between the corresponding surface points seen by each camera, the distance of face surface point can be determined through triangulation. A passive scanner from the Dimensional Imaging which has been used by the authors for the research on facial expression analysis is shown in Figure 3. This 3-D scanner uses six digital cameras with three cameras on each side to capture six views of the face surface. Four of these images are used to reconstruct 3D facial geometry and the other two images provide the textural information for accurate 3D face rendering.

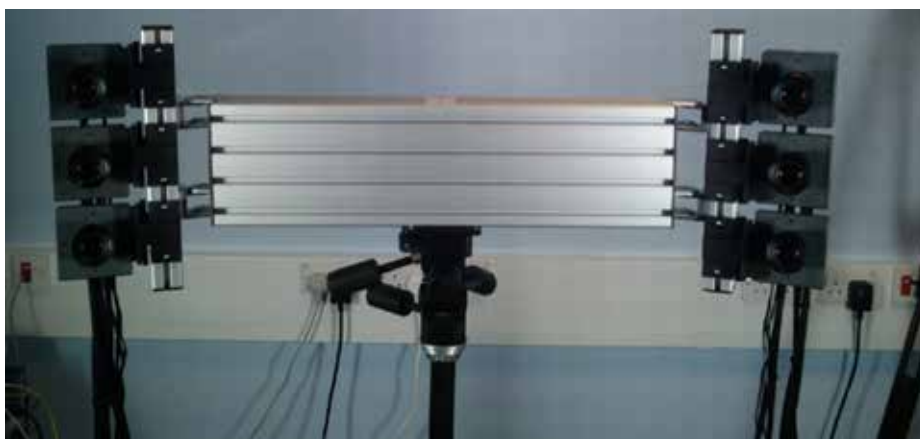


Fig. 3. A passive 3-D scanner from Dimensional Imaging.

Similar to the dynamic 2-D data, a dynamic 3-D facial data set is a time ordered collection of 3-D surfaces. This, among others, enables temporal tracking of facial points' motion in the 3-D space. The scanner shown in Figure 3 is capable of capturing dynamic data up to 60 frames per second (fps). Each second of the facial data with a resolution of 20,000 vertices captured at 60 fps requires around 10 Gigabytes of storage. This example illustrates one of the main disadvantages of 3-D dynamic scanners, namely: required availability of significant storage and computational resources.

2.2 Facial expression representations

Facial expression representation is essentially a feature extraction process, which converts the original facial data from a low-level 2-D pixel or 3-D vertex based representation, into a higher-level representation of the face in terms of its landmarks, spatial configuration, shape, appearance and/or motion. The extracted features usually reduce the dimensionality of the original input facial data (Park & Park, 2004) (a noticeable example to the contrary would be Haar or Gabor features calculated as an input for the AdaBoost training algorithm, where dimensionality of the feature vector could be higher than the dimensionality of the original data). Presented in the following are a number of popular facial expression representations.

A landmark based representation uses facial characteristic points, which are located around specific facial areas, such as edges of eyes, nose, eyebrows and mouth, since these areas show significant changes during facial articulation. Kobayashi and Hara (Kobayashi & Hara, 1997) proposed a geometric face model based on 30 facial characteristic points for the frontal face view as shown in Figure 4(a). Subsequently, the point-based model was extended to include 10 extra facial characteristic points on the side view of the face (Pantic & Rothkrantz, 2000) as shown in Figure 4(b). These points on the side view are selected from the peaks and valleys of the profile contours.

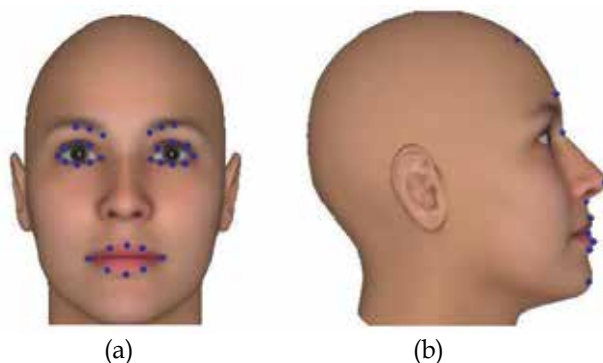


Fig. 4. Point-based model: (a) 30 facial points selected from the frontal view, (b) 10 facial points selected from the side view.

The localised geometric model could be classified as a representation based on spatial configuration derived from facial images (Saxena et al., 2004). The method utilises a facial feature extraction method, which is based on classical edge detectors combined with colour analysis in the HSV (hue, saturation, value) colour space to extract the contours of local facial features, such as eyebrows, lips and nose. As the colour of the pixels representing lips, eyes and eyebrows differ significantly from those representing skin, the contours of these

features can be easily extracted from the hue colour component. After facial feature extraction, a feature vector built from feature measurements, such as the brows distance, mouth height, mouth width etc., is created.

Another representation based on spatial configuration is topographic context (TC) that has been used as a descriptor for facial expressions in 2-D images (Wang & Yin, 2007). This representation treats an intensity image as a 3-D terrain surface with the height of the terrain at pixel (x,y) represented by its image grey scale intensity $I(x,y)$. Such image interpretation enables topographic analysis of the associated surface to be carried out leading to a topographic label, calculated based on a local surface shape, being assign to each pixel location. Resulting TC feature is an image of such labels assigned to all facial pixels of the original image. Topographic labels include: peak, ridge, saddle, hill, flat, ravine and pit. In total, there are 12 types of topographic labels (Trier et al., 1997). In addition, hill-labelled pixels can be divided into concave hill, convex hill, saddle hill (that can be further classified as a concave saddle hill or a convex saddle hill) and slope hill; and saddle-labelled pixels can be divided into ridge saddle or ravine saddle. For a facial image, the TC-based expression representation requires only six topographic labels shown in Figure 5 (Yin et al., 2004).

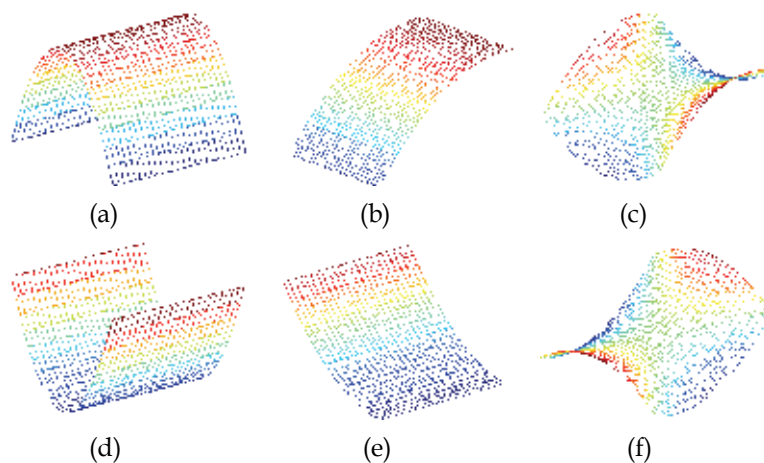


Fig. 5. A subset of the topographic labels: (a) ridge, (b) convex hill, (c) convex saddle hill, (d) ravine, (e) concave hill, and (f) concave saddle hill.

Similar to TC, the local binary patterns (LBP) have been also used to represent facial expressions in 2-D images (Liao et al., 2006). LBP as an operator was first proposed by Ojala et al. (Ojala et al., 2002) for texture description. An example of LBP calculation is illustrated in Figure 6. For a given pixel (shown in red in Figure 6), its value is subtracted from all the neighbouring pixels and the sign of the results is binary coded. After the clockwise grouping of the binary bits, starting from the top left pixel, the arranged binary string is converted to a decimal number as the final LBP result for that pixel. The LBP operator is an intensity invariant texture measure, which captures the directions of the intensity changes in an image. Using the LBP operator, each pixel of a facial image can be encoded by a LBP value which preserves the intensity difference with respect to its local neighbours. This encoded image can be used for the classification of facial expressions. Figure 7 shows a facial image and its corresponding LBP encoded image. The original LBP algorithm has also been

modified to encode the information of depth differences and applied to 3-D face representations (Huang et al., 2006).

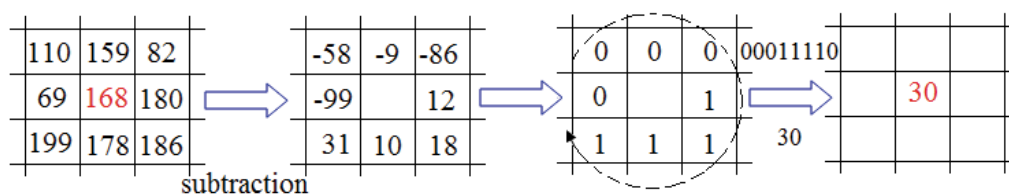


Fig. 6. An example of LBP calculation.

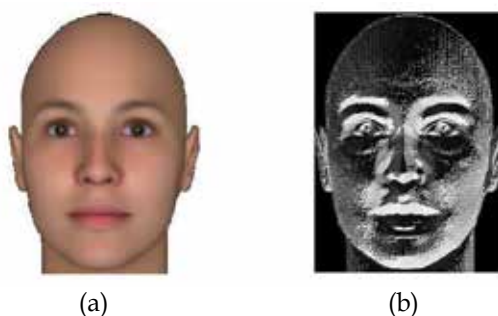


Fig. 7. A facial image and its LBP encoded image: (a) original, and (b) encoded.

Among various shape based representations, the facial expression representation method based on the B-spline surface was proposed by Hoch (Hoch et al., 1994) for 3-D facial data. The B-spline surface is a parametric model which can efficiently represent 3-D faces using a small number of control points. Using B-splines in combination with the facial action coding system (FACS) (Ekman & Friesen, 1978), different facial expressions can be automatically created on the faces by moving the specific control points which correspond to the different action units of the FACS. The FACS is a human observer based system with the ability to detect subtle changes in facial features and it forms a widely accepted theoretical foundation for the control of facial animation (Hager et al., 2002). There are 46 action units defined by the FACS with each action unit representing a specific basic action performable on the human face.

Another often used shape based representation is the statistical shape model (SSM) (Cootes et al., 1995). The SSM can be built based on facial points, landmarks or other parameters (e.g. B-spline coefficients) describing face shape. Given a set of training faces with known correspondences between them, the SSM finds a mean representation of these faces, e.g. mean configuration of the facial landmarks, and their main modes of variation (so called eigen-faces). With such SSM each new face can be approximately represented as a linear superposition of the mean face and the weighted eigen-faces. The weights, calculated through projection of the observed face on the subspace spanned by the eigen-faces, can be subsequently used as an approximate representation of the face.

Using the SSM constructed, a 3-D face in the training data set is represented by a shape space vector (SSV) of much smaller dimensionality than the dimensionality of the original data space vector (Blake & Isard, 1998). To produce the SSV for an unseen new face, a modified iterative closest point (ICP) method is often incorporated with the least-squares

projection to register the SSM to the new face (Cootes & Taylor, 1996). For 3-D facial expression representation, it has been proved experimentally by the authors (Quan et al., 2007a; Quan et al., 2007b) that the SSV is able to encode efficiently different facial expressions.

A combination of shape and appearance based representations yields the active appearance model (AAM), which could be classified as another statistical model and has been used for facial expression representation (Hong et al., 2006; Edwards et al., 1998). It models the shape as well as grey levels textures and it is mainly used for 2-D facial images in facial expression representation applications. The model is built using a set of training facial images with corresponding landmarks selected manually and localised around the prominent facial features. In order to build the AAM, the selected landmarks representing the shape of each training face are aligned into a common coordinate system with each training facial image normalised to the same size. Subsequently, principal component analysis (PCA) is applied to the aligned landmarks and the normalised facial images to yield a vector of parameters which represents both the shape and grey level variations of the training facial images (Cootes et al., 1998).

The 3-D morphable model is an extension of the AAM for 3-D data. Instead of using the manually selected landmarks, the 3-D morphable model uses all the data points of 3-D facial scans to represent the geometrical information (Banz & Vetter, 2003). Optic flow is often used for establishing the dense correspondences of points between each training face and a reference face in texture coordinate space. This model has been used for representing the facial expressions which are embedded in the 3-D faces (Ramanathan et al., 2006).

Finally, a motion based representation is the local parameterised model, which is calculated using the optical flow of image sequences (Black & Yacoob, 1997). Two parametric flow models are available for image motion, namely: affine and affine-plus-curvature. They not only model non-rigid facial motion, but also provide a description about the motion of facial features. Given the region of a face and locations of facial features of interest, such as eyes, eye brows and mouth, the motion of the face region between two frames in the image sequence is first estimated using the affine model. Subsequently the estimated motion is used to register the two images via warping, and the relative motions of the facial features are computed using the affine-plus-curvature model. Using the estimated motions, the region of the face and the locations of the facial features are predicted for the next frame, and the process is repeated for the whole image sequence. The parameters of estimated motions provide a description of the underlying facial motions and can be used to classify facial expressions. The region of the face and the locations of the facial features can be manually selected in the initial processing stage and automatically tracked thereafter.

2.3 Facial expression classification

In the context of facial expression analysis, classification is a process of assigning observed data to one of predefined facial expression categories. The specific design of this process is dependent on the type of the observation (e.g. static or dynamic), adopted data representation (type of the feature vector used to represent the data) and last but not the least the classification algorithm itself. As there is a great variety of classification algorithms reported in literature, this section will focus only on those, which are the most often used, or which have been recently proposed in the context of facial expression recognition. From that perspective, some of the most frequently used classification methods include nearest

neighbour classifiers, Fisher's linear discriminant (also known as linear discriminant analysis), support vector machines, artificial neural networks, AdaBoost, random forests, and hidden Markov models.

Nearest neighbour classifier (NNC) is one of the simplest classification methods, which classifies objects based on closest training examples in the feature space. It can achieve consistently high performance without a prior assumption about the distribution from which the training data is drawn. Although there is no explicit training step in the algorithm, the classifier requires access to all training examples and the classification is computationally expensive when compared to other classification methods. The NNC assigns a class based on the smallest distances between the test data and the data in the training database, calculated in the feature space. A number of different distance measures have been used, including Euclidean and weighted Euclidean (Md. Sohail and Bhattacharya, 2007), or more recently geodesic distance for features defined on a manifold (Yousefi et al., 2010).

Linear discriminant analysis (LDA) finds linear decision boundaries in the underlying feature space that best discriminate among classes, i.e., maximise the between-class scatter while minimise the within-class scatter (Fisher, 1936). A quadratic discriminant classifier (Bishop, 2006) uses quadratic decision boundary and can be seen, in the context of Bayesian formulation with normal conditional distributions, as a generalisation of a linear classifier in case when class conditional distributions have different covariance matrices.

In recent years, one of the most widely used classification algorithms are support vector machines (SVM) which performs classification by constructing a set of hyperplanes that optimally separate the data into different categories (Huang et al., 2006). The selected hyperplanes maximise the margin between training samples from different classes. One of the most important advantages of the SVM classifiers is that they use sparse representation (only a small number of training examples need to be maintained for classification) and are inherently suitable for use with kernels enabling nonlinear decision boundary between classes.

Other popular methods are the artificial neural networks. The key element of these methods is the structure of the information processing system, which is composed of a large number of highly interconnected processing elements working together to solve specific problems (Padgett et al., 1996).

AdaBoost (Adaptive Boosting) is an example of so called boosting classifiers which combine a number of weak classifiers/learners to construct a strong classifier. Since its introduction (Freund and Schapire, 1997), AdaBoost is enjoying a growing popularity. A useful property of these algorithms is their ability to select an optimal set of features during training. As results AdaBoost is often used in combination with other classification techniques where the role of the AdaBoost algorithm is to select optimal features which are subsequently used for classification by another algorithm (e.g. SVM). In the context of facial expression recognition Littlewort (Littlewort et al., 2005) used the AdaBoost to select best Gabor features calculated for 2D video which have been subsequently used within SVM classifier. Similarly in (Ji and Idrissi, 2009) authors used a similar combination of AdaBoost for feature selection and SVM for classification with LBP calculated for 2D images. In (Whitehill et al., 2009) authors used a boosting algorithm (in that case GentleBoost) and the SVM classification algorithm with different features including Gabor filters, Haar features, edge orientation histograms, and LBP for detection of smile in 2D stills and videos. They demonstrated that when trained on real-life images it is possible to obtain human like smile recognition accuracy. Maalej (Maalej

et al., 2010) successfully demonstrated the use of Adaboost and SVM, utilising different kernels, with feature vector defined as geodesic distances between corresponding surface's patches selected in the input 3D static data.

More recently random forest (Breiman, 2001) classification techniques have gained significant popularity in the computer vision community. In (Flanelli et al., 2010) authors used random forest with trees constructed from a set of 3D patches randomly sampled from the normalised face in 2D video. The decision rule used in each tree node was based on the features calculated from a bank of log-Gabor filters and estimated optical flow vectors.

Hidden Markov Models (HMM) are able to capture dependence in a sequential data and therefore are often the method of choice for classification of spatio-temporal data. As such they have been also used for classification of facial expressions from 2D (Cohen et al., 2003) and 3D (Sun & Yin, 2008) video sequences.

3. Expression recognition from 3-D facial data

As discussed in the previous section, automatic recognition of facial expressions can be achieved through various approaches by using different facial expression representation methods and feature classification algorithms. In this section, more details are given for one of the techniques previously proposed by the authors. The method uses the shape space vector (SSV) as the feature for facial expression representation, which is derived from the statistical shape model (SSM) and has shown promising results in facial expression recognition (Quan et al., 2009b).

3.1 Statistical shape models

The basic principle of the SSM is to exploit a redundancy in structural regularity of the given shapes, thereby enabling a shape to be described with fewer parameters, i.e., reducing the dimensionality of the shapes presented in the original spatial domain. In order to achieve this dimensionality reduction, the principal component analysis (PCA) is usually used. Given a set of M 3-D facial surfaces, $\{\mathbf{Q}_i, i \in [1, M]\}$, where \mathbf{Q}_i denotes a column vector representing the set of N corresponding vertices in the i -th face with $\mathbf{Q}_i \in \mathbb{R}^{3N}$, the first step of the PCA is to construct the mean vector, denoted by $\bar{\mathbf{Q}}$, and calculated from all the available data.

$$\bar{\mathbf{Q}} = \frac{1}{M} \sum_{i=1}^M \mathbf{Q}_i \quad (1)$$

Let \mathbf{C} be defined as a $3N \times 3N$ covariance matrix calculated from the training data set.

$$\mathbf{C} = \frac{1}{M} \sum_{i=1}^M (\mathbf{Q}_i - \bar{\mathbf{Q}})(\mathbf{Q}_i - \bar{\mathbf{Q}})^T \quad (2)$$

By building matrix \mathbf{X} of "centred" shape vectors, with $(\mathbf{Q}_i - \bar{\mathbf{Q}})$ as the i -th column of matrix \mathbf{X} , covariance matrix \mathbf{C} can be calculated as

$$\mathbf{C} = \mathbf{X}\mathbf{X}^T \quad (3)$$

Since the number of faces, M , in the training data set is usually much smaller than the number of data points on a face, the eigenvectors \mathbf{u}'_i and eigenvalues λ'_i of $M \times M$ matrix $\mathbf{X}^T \mathbf{X}$ are calculated first, and the corresponding eigenvectors \mathbf{u}_i and eigenvalues λ_i of the $\mathbf{X}\mathbf{X}^T$ are subsequently calculated from $\mathbf{u}_i = \mathbf{X}\mathbf{u}'_i / \|\mathbf{X}\mathbf{u}'_i\|$ and $\lambda_i = \lambda'_i$ (Vrtovec et al., 2004).

By using these eigenvalues and eigenvectors, the data points on any 3-D face in the training data set can be approximately represented using a linear model of the form

$$\hat{\mathbf{Q}} = \mathbf{W}\mathbf{b} + \bar{\mathbf{Q}} \quad (4)$$

where $\mathbf{W} = [\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_K]$ is a $3N \times K$ so called Shape Matrix of K eigenvectors, or "modes of variation", associated with the largest eigenvalues; and $\mathbf{b} = [b_1, \dots, b_i, \dots, b_K]^T$ is the shape space vector (SSV), which controls the contribution of each eigenvector, \mathbf{u}_i , in the approximated surface $\hat{\mathbf{Q}}$ (Cootes et al., 1995).

Projection of a facial surface denoted by \mathbf{Q} on to the shape space using the shape matrix is then given by

$$\mathbf{b} = \mathbf{W}^T (\mathbf{Q} - \bar{\mathbf{Q}}) \quad (5)$$

Based on the SSM built using 450 randomly selected 3-D faces from the BU-3DFE database (see Section 4), Figure 8 shows variability of the first three elements of the SSV. The faces in each row have been obtained by adding to the mean face the corresponding eigenvector multiplied by a varying weight.

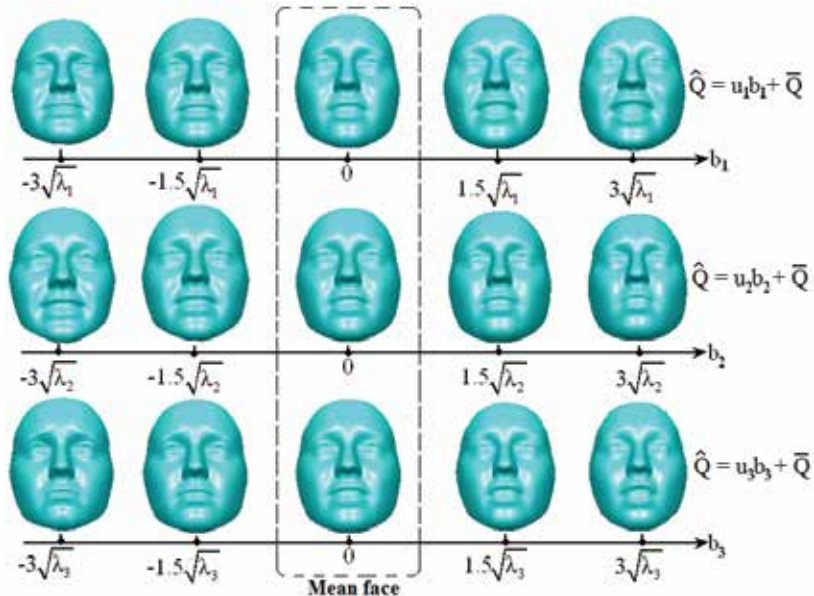


Fig. 8. Facial articulations controlled by the first three elements of the SSV.

In order to achieve the facial expression recognition using the extracted SSV, two processing stages are required, namely: model building and model fitting. The aim of the model building is to construct a 3-D SSM from a set of training data, whereas the model fitting is to estimate the SSV describing the new input 3-D faces so the SSM matches the new face.

3.2 Model building

A fundamental problem when building an SSM lies in correct determination of point correspondences between 3-D faces in the training data set. This is critical, because incorrect correspondences can either introduce too much variations or lead to invalid cases of the model (Cootes et al., 1995). The dense point correspondence estimation can be achieved in three steps: (i) facial landmark determination, (ii) thin-plate spline warping, and (iii) closest points matching. The first step is to identify the sparse facial landmarks on the training faces and a selected reference face; the second step is to warp the reference face to each of the training faces using the thin-plate spline transformation; and the last step is to use the warped reference face to estimate dense point correspondences for each of the training faces.

3.2.1 Facial landmark determination

This step is to identify the sparse facial landmarks which are around areas sensitive to facial expressions. These landmarks can be either extracted automatically or selected manually. Although automated identification of facial landmarks could be fast, without requiring user input, it is likely to fail at some point and has low accuracy due to data uncertainties. On the other hand, manual selection of facial landmarks can be used to handle data uncertainties, but it is tedious, time-consuming and user dependent.

A number of authors have proposed an automated identification method by using a template with annotated landmarks. Through registration of the template with a face, the landmarks pre-selected in the template can be propagated to the face (Frangi et al., 2002; Rueckert et al., 2003). Optical flow is another technique that can be used to identify corresponding landmarks between 3-D facial surfaces through their 2-D texture matching (Blanz & Vetter, 1999).

For the results presented later in this section the BU-3DFE database was used to construct the SSM. Each face in that database has 83 manually selected landmarks and these landmarks were used to establish a dense point correspondence. Figure 9 shows an example of faces in the BU-3DFE database with the selected facial landmarks.



Fig. 9. BU-3DFE faces with selected facial landmarks (Yin et al., 2006).

3.2.2 Thin-plate spline warping

Using the sparse facial landmarks, the non-rigid transformations between the reference and training faces are estimated so that the reference face can be warped or deformed to match the shape of each training face. The thin-plate spline transformation model is used to achieve the non-rigid transformation (Bookstein, 1989), since the thin-plate spline is an effective tool for modeling deformations by combining global and local transformations and it has been applied successfully in several computer vision applications (Johnson & Christensen, 2002). In this step, the coefficients of the thin-plate spline transformation model are first computed using the sparse corresponding facial landmarks in two facial surfaces, and the transformation with the computed coefficients are applied to all the points of one facial surface so as to deform it to match to the other facial surface.

3.2.3 Closest point matching

After thin-plate spline warping, the reference face is deformed to match each of the training faces. Since the pose and shape of the deformed reference face are very close to the associated training face, the correspondences of all points between them can be determined using the Euclidean distance metric (Besl & McKay, 1992).

Let $\mathcal{S} = \{\mathbf{p}_i, i \in [1, N]\}$ be the point set of one of the training faces. The point $\hat{\mathbf{p}}(\mathbf{q})$ on that training face estimated to be closest to any given point \mathbf{q} in the reference face is given by

$$\hat{\mathbf{p}}(\mathbf{q}) = \arg \min_{\mathbf{p} \in \mathcal{S}} d(\mathbf{p}, \mathbf{q}) \quad (6)$$

where $d(\mathbf{p}, \mathbf{q})$ is the Euclidean distance between points \mathbf{p} and \mathbf{q} . When the correspondences between points on the reference face and each of the training faces are found, the whole training data set is aligned and brought into correspondence. With the aligned training data set, the SSM can be built directly using PCA as discussed in Section 3.1. Figure 10 shows examples on how the reference face is deformed to match five, randomly selected, training faces. These deformed reference faces are subsequently used to select points on the training faces which are in correspondence, across all the faces in the training database. It is seen that the shapes of deformed reference faces are close enough to the training faces to allow for selection of dense correspondences.

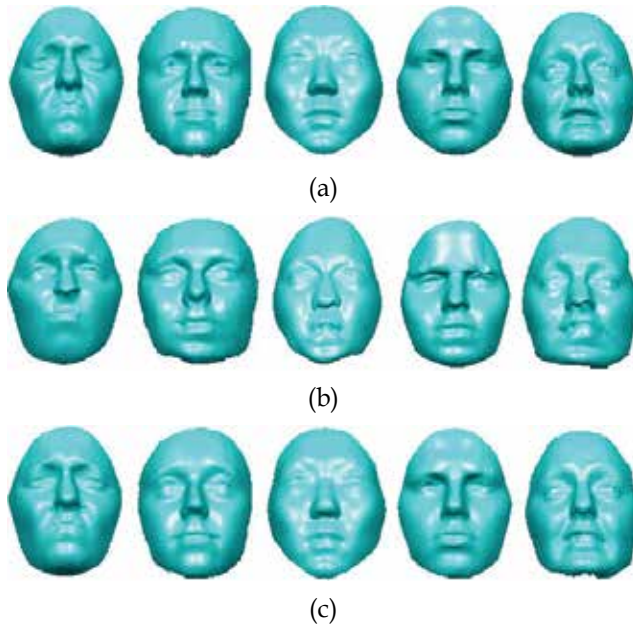


Fig. 10. Alignment of training images using thin-plate spline warping: (a) training faces, (b) reference faces after the thin-plate spline warping, and (c) training faces after re-selection of points which are in correspondence across all faces in the training database

3.3 Model fitting

The model fitting process is treated as 3-D surface registration, which consists of initial model fitting followed by refined model fitting. While the former provides a global

alignment, using similarity transformation, between the new input face and the SSM model constructed, the latter refines the alignment by iteratively deforming the SSM model to match the input face.

3.3.1 Initial model fitting

The initial model fitting stage is achieved by using the iterative closest point (ICP) method with similarity transformation (Besl & McKay, 1992), and is implemented in two steps.

The first step is to roughly align the mean face, of the built SSM, and the input face based on their estimated centroids. This alignment does not have to be very accurate. An example of such alignment is shown in Figure 11(b), where the blue surface indicates the mean face of the SSM and the yellow surface represents the input face.

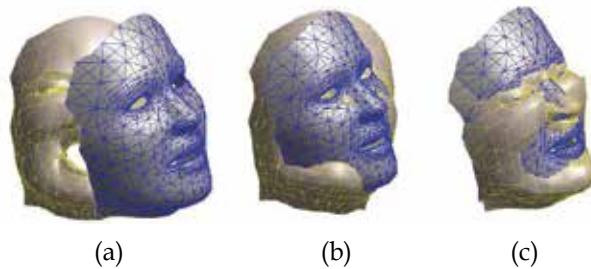


Fig. 11. An example of initial model fitting: (a) starting poses, (b) after rough alignment, and (c) after initial model fitting.

In the second step this rough alignment is iteratively refined by alternately estimating point correspondences and finding the best similarity transformation that minimises a cost function between the corresponding points. In the specific implementation of the algorithm described here the cost function is defined as

$$E = \frac{1}{N} \sum_{i=1}^N \|\mathbf{q}_i - (s\mathbf{R}\mathbf{p}_i + \mathbf{t})\|^2 \quad (7)$$

where N is the number of point correspondence pairs between the two surfaces, \mathbf{q}_i and \mathbf{p}_i are respectively corresponding points on the model and new data surfaces, \mathbf{R} is a 3×3 rotation matrix, \mathbf{t} is a 3×1 translation vector, and s is a scaling factor. For the given point correspondence the pose parameters minimising cost function E can be found in a closed form (see (Umeyama, 1991) for more details).

The iterations in the second step could be terminated either when: (i) the alignment error between the two surfaces is below a fixed threshold, (ii) the change of the alignment error between the two surfaces in two successive iterations is below a fixed threshold, or (iii) the maximum number of iterations is reached. An option, which is often used, is to combine the second and third conditions.

Furthermore, since the search of the dense point correspondences is one of the most time-consuming tasks, a multi-resolution correspondence search has been developed to reduce the computation time. It starts searching for the point correspondences with a coarse point density, and gradually increases the point density in a series of finer resolutions as the registration error is reduced. To make this process more robust, random sampling is used to

sub-sample the original surface data to obtain a sequence of lower resolution data sets, with each sub-sampling reducing the number of the surface points by a quarter. At the beginning of the correspondence search, the lowest resolution data set is used, which allows for estimation of the large coarse motion of the model to match the input face. As the point density increases with the subsequent use of the finer data sets, it gradually restricts the amount of motion allowed and produces finer and finer alignment. Hence, the use of the multi-resolution correspondence search not only saves computation time, but also improves robustness of the alignment by avoiding local minima. Figure 11(c) shows an example result produced by the initial model fitting stage.

3.3.2 Refined model fitting

After the initial model fitting stage, the SSM and the input face are globally aligned. In the next stage, the SSM is iteratively deformed to better match the input face. This is achieved in the refined model fitting stage that consists of two main intertwined iterative steps, namely: shape and correspondence updates.

Based on the correspondence established in the initial model fitting stage, the new face denoted here by vector $\mathbf{P}^{(k)}$, with k representing the iteration index, and built by concatenating all surface vertices $\mathbf{p}_i^{(k)}$ into a single vector, is projected onto the shape space to produce the first estimate of the SSV, $\mathbf{b}^{(k)}$ and the first shape update, $\hat{\mathbf{Q}}^{(k)}$, as described in equations (8-9).

$$\mathbf{b}^{(k)} = \mathbf{W}^T (\mathbf{P}^{(k)} - \bar{\mathbf{Q}}) \quad (8)$$

$$\hat{\mathbf{Q}}^{(k)} = \mathbf{W}\mathbf{b}^{(k)} + \bar{\mathbf{Q}} \quad (9)$$

In the subsequent correspondence update stage, as described in equations (10-12), the data vertices, $\mathbf{p}_i^{(k)}$, are matched against vertices $\mathbf{q}_i^{(k)}$ from the updated shape, $\hat{\mathbf{Q}}^{(k)}$. This includes similarity transformation of the data vertices to produce new vertices position $\tilde{\mathbf{p}}_i^{(k+1)}$, and updating of the correspondence, represented by re-indexing of the data vertices, $\mathbf{p}_{j(i)}^{(k+1)}$.

$$\{ \hat{\mathbf{R}}, \hat{\mathbf{T}}, \hat{s} \} = \arg \min_{\{ \mathbf{R}, \mathbf{T}, s \}} \left(\sum_i \left\| \hat{\mathbf{q}}_i^{(k)} - s\mathbf{R}\mathbf{p}_i^{(k)} - \mathbf{T} \right\|^2 \right) \quad (10)$$

$$\tilde{\mathbf{p}}_i^{(k+1)} = \hat{s}\hat{\mathbf{R}}\mathbf{p}_i^{(k)} + \hat{\mathbf{T}} \quad (11)$$

$$\tilde{\mathbf{p}}_i^{(k+1)} \rightarrow \mathbf{p}_{j(i)}^{(k+1)} \quad (12)$$

This process iterates till either the change of the alignment error between the updated shape and the transformed data vertices is below a fixed threshold, or the maximum number of iterations is reached. The final result of the described procedure is the shape state vector $\mathbf{b}^{(K)}$, where K denotes the last iteration, representing surface deformations between the model mean face and the input facial data.

From equation 8, it is seen that the size of the SSV is fixed which in turn determines the number of the eigenvectors from the shape matrix to be used. Although using a fixed size of the SSV can usually provide a reasonable final result of matching if a good initial alignment is achieved, it may mislead the minimisation of the cost function towards a local minimum in the refined model fitting stage. This can be explained by the fact that when a large size of the SSV is used, the SSM has a high degree of freedom enabling it to iteratively deform to a shape representing a local minimum, if it happens that the SSM was instantiated in the basing of this minimum. On the other hand, the SSM with a small size of the SSV has a low degree of freedom which constrains the deformation preventing it from converging to shapes associated with the local minima of the cost function, thereby limiting the ability of the SSM to accurately represent the data (Quan et al., 2010a). Figure 12 shows some examples of model matching failures caused by the size of the SSV being too large or too small.

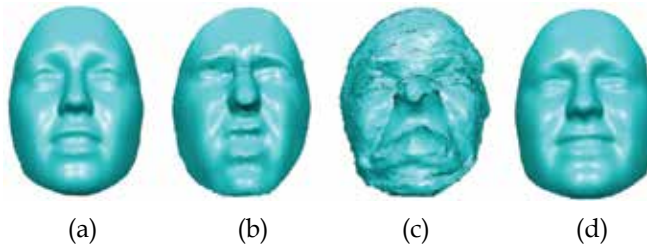


Fig. 12. Failed examples of model refinement by using a fixed size of SSV: (a) model, (b) input face, (c) using large size of SSV, and (d) using small size of SSV.

In order to solve the problem caused by the fixed size of SSV, the multi-level model deformation approach could be employed by using an adaptive size for the SSV. In the beginning of the refined model fitting stage, the SSV has a small size. Although it may result in a large registration error, it allows the algorithm to provide a rough approximation of the data. When the registration error is decreased, the size of SSV is gradually increased to provide more shape flexibility and allow the model to match the data. To implement this adaptive approach the fixed sized shape space matrix \mathbf{W} in equations (8-9) needs to be replaced with \mathbf{W}_k , where index k signifies that the size of the matrix (number of columns) is changing (increasing) with the iteration index. A reasonable setting for the adaptive SSV is to start with

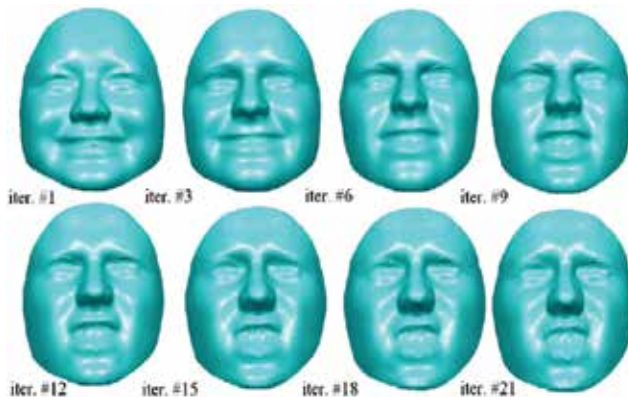


Fig. 13. Example of intermediate results obtained during iteration of refined model fitting with adaptive size of SSV.

an SSV size that enables the SSM to contain around 50% shape variance of the training data set (this corresponds to five eigenvectors in the case of 450 BU-3DFE training faces), to increase the SSV size by one in every two iterations during the refined model fitting stage, and to terminate the increase when the SSV size enables the SSM to contain over 95% shape variance. Some intermediate results taken from 21 iterations performed using the described procedure are shown in Figure 13. In that figure, the model and the input face used are the same as those shown in Figure 12. It can be seen that the multi-level model deformation approach not only provides a smooth transition between iterations during the model refinement stage but also enables the model to match the appropriate shape accordingly.

4. Facial expression databases

In order to evaluate and benchmark facial expression analysis algorithms, standardised data sets are needed to enable a meaningful comparison. Based on the type of facial data used by an algorithm, the facial expression databases can be categorised into 2-D image, 2-D video, 3-D static and 3-D dynamic. Since facial expressions have been studied for a long time using 2-D data, there is a large number of 2-D image and 2-D video databases available. Some of the most popular 2-D image databases include CMU-PIE database (Sim et al., 2002), Multi-PIE database (Gross et al., 2010), MMI database (Pantic et al., 2005), and JAFFE database (Lyons et al., 1999). The commonly used 2-D video databases are Cohn-Kanade AU-Coded database (Kanade et al., 2000), MPI database (Pilz et al., 2006), DaFEx database (Battocchi et al., 2005), and FG-NET database (Wallhoff, 2006). Due to the difficulties associated with both 2-D image and 2-D video based facial expression analysis in terms of handling large pose variation and subtle facial articulation, there is recently a shift towards the 3-D based facial expression analysis, however this is currently supported by a rather limited number of 3-D facial expression databases. These databases include BU-3DFE (Yin et al., 2006), and ZJU-3DFED (Wang et al., 2006b). With the advances in 3-D imaging systems and computing technology, 3-D dynamic facial expression databases are beginning to emerge as an extension of the 3-D static databases. Currently the only available databases with dynamic 3-D facial expressions are ADSIP database (Frowd et al., 2009), and BU-3DFE database (Yin et al., 2008).

4.1 2-D image facial expression databases

CMU-PIE initial database was created at the Carnegie Mellon University in 2000. The database contains 41,368 images of 68 people, and the facial images taken from each person cover 4 different expressions as well as 13 different poses and 43 different illumination conditions (Sim et al., 2002). Due to the shortcomings of the initial version of CMU-PIE database, such as a limited number of subjects and facial expressions captured, the Multi-PIE database has been developed recently as an expansion of the CMU-PIE database (Gross et al., 2010). The Multi-PIE includes more than 750,000 images from 337 subjects, which were captured under 15 view points and 19 illumination conditions. MMI database includes hundreds of facial images and video recordings acquired from subjects of different age, gender and ethnic origin. This database is continuously updated with acted and spontaneous facial behaviour (Pantic et al., 2005), and scored according to the facial action coding system (FACS) (Ekman & Friesen, 1978). JAFFE database contains 213 images of 6 universal facial expressions plus the neutral expression (Lyons et al., 1999). This database was created with a help of 10 Japanese female models. Examples of the six universal expressions from that database are shown in Figure 14.

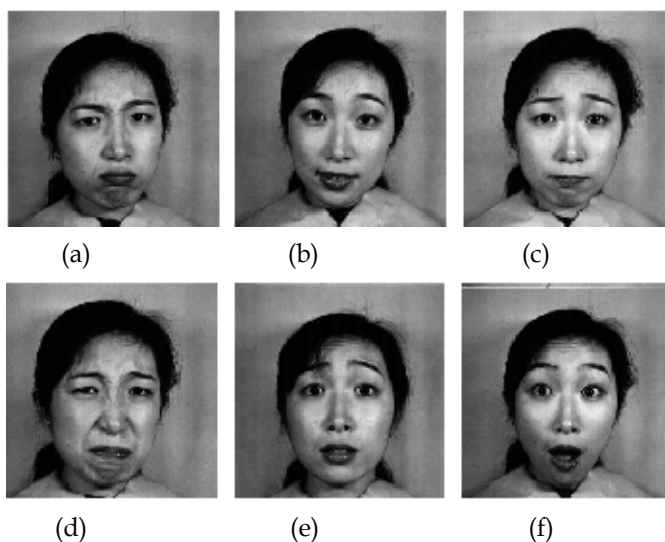


Fig. 14. Examples of the six universal facial expressions from the JAFFE database: (a) anger, (b) happiness, (c) sadness, (d) disgust, (e) fear, and (f) surprise.

4.2 2-D video facial expression databases

Cohn-Kanade AU Coded database is one of the most comprehensive 2-D video databases. Its initial release consisted of 486 video sequences from 97 subjects who were university students (Kanade et al., 2000). They ranged in age from 18 to 30 years old, with a variety of ethnic origins, including African-American, Asian and Hispanic. The peak expression for each sequence is coded according to the FACS (Ekman & Friesen, 1978). The second version of the database is an expansion of its initial release, which includes both posed and spontaneous facial expressions, with increased number of video sequences and subjects (Lucey et al., 2010). The third version of the database is planned to be published in 2011, and will contain, apart from the frontal view, additional synchronised recordings taken at 30 degrees angle. MPI database was developed at the Max Planck Institute for Biological Cybernetics (Pilz et al., 2006). The database contains video sequences of four different expressions: anger, disgust, surprise and gratefulness. Each expression was record from five different views simultaneously as shown in Figure 15. DaFEx database includes 1,008 short videos containing 6 universal facial expressions and the neutral expression from 8

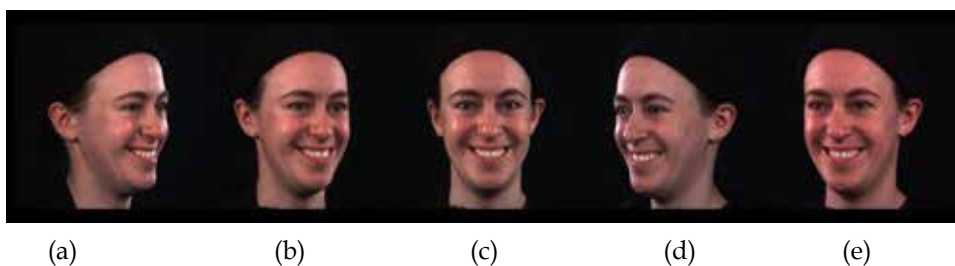


Fig. 15. Examples of the MPI database showing five views: (a) left 45 degree, (b) left 22 degree, (c) front, (d) right 22 degree, (e) right 45 degree.

professional actors (Battocchi et al., 2005). Each universal facial expression was performed at three intensity levels. FG-NET database has 399 video sequences which were gathered from 18 individuals (Wallhoff, 2006). The emphasis in that database is put on recording a spontaneous behaviour of the subjects with emotions induced by showing participants suitable selected video clips. Similar to the DaFEx database, it covers 6 universal facial expressions plus the neutral expression.

4.3 3-D static facial expression databases

BU-3DFE database was developed at the Binghamton University for the purpose of 3-D facial expression analysis (Yin et al., 2006). The database contains 100 subjects, with ages ranging from 18 to 70 years old, with a variety of ethnic origins including White, Black, East-Asian, Middle-East Asian, Indian and Hispanic. Each subject performed seven expressions, which include neutral and six universal facial expressions at four intensity levels. With 25 3-D facial scans containing different expressions for each subject, there is a total of 2,500 facial scans in the database. Each 3-D facial scan in the BU-3DFE database contains 13,000 to 21,000 polygons with 8,711 to 9,325 vertices. Figure 16 shows some examples from the BU-3DFE database.



Fig. 16. Examples from the BU-3DFE database (Yin et al., 2006).

ZJU-3DFED database is a static 3-D facial expression database, which was developed in the Zhe Jiang University (Wang et al., 2006b). Compared to other 3-D facial expression databases, the size of ZJU-3DFED is relatively small. It contains 360 facial models from 40 subjects. For each subject, there are 9 scans with four different kinds of expressions.

4.4 3-D dynamic facial expression databases

BU-4DFE database (Yin et al., 2008) is a 3-D dynamic facial expression database and an extension of the BU-3DFE database to enable the analysis of the facial articulation using dynamic 3-D data. The 3D facial expressions are captured at 25 frames per second (fps), and the database includes 606 3D facial expression sequences captured from 101 subjects. For each subject, there are six sequences corresponding to six universal facial expressions (anger, disgust, happiness, fear, sadness, and surprise). A few 3-D temporal samples from one of the BU-4DFE sequences are shown in Figure 17.

ADSIP database is a 3-D dynamic facial expression database created at the University of Central Lancashire (Frowd et al., 2009). The first release of the database (ADSIPmark1) was completed in 2008 with help from 10 graduates from the School of Performing Arts. The use of actors, and trainee actors enables capture of fairly representative and accurate facial expressions (Nusseck et al., 2008). Each subject performed seven expressions: anger, disgust, happiness, fear, sadness, surprise and pain, at three intensity levels (mild, normal and extreme). Therefore, there is a total of 210 3D facial sequences in that database. Each sequence was captured at 24 fps and lasts for around three seconds. Additionally each 3D

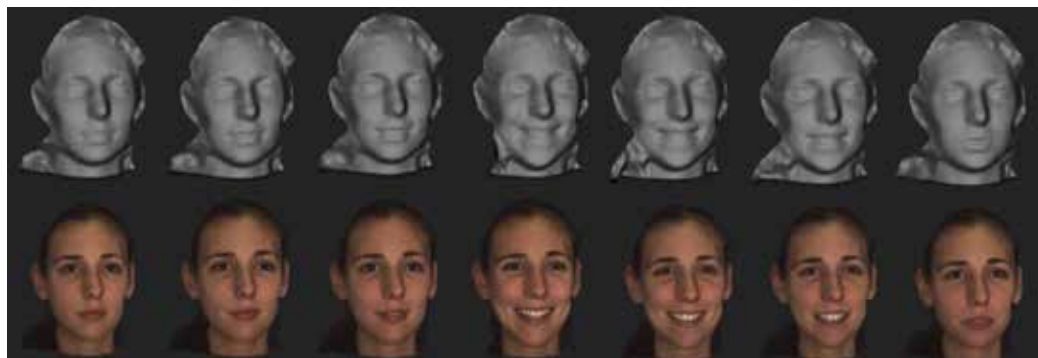


Fig. 17. Temporal 3-D samples from one of the sequences in the BU-4DFE database (Yin et al., 2008).

sequence is accompanied by a standard video recording captured in parallel with the 3D sequence. This database is unique in the sense that it has been independently validated as all the recordings in the database have been assessed by 10 independent observers. These observers assigned a score against each type of the expression for all the recordings. Each score represented how confident observers were about each sequence depicting each type of the expression. Results of this validation are summarised in the next section. Figure 18

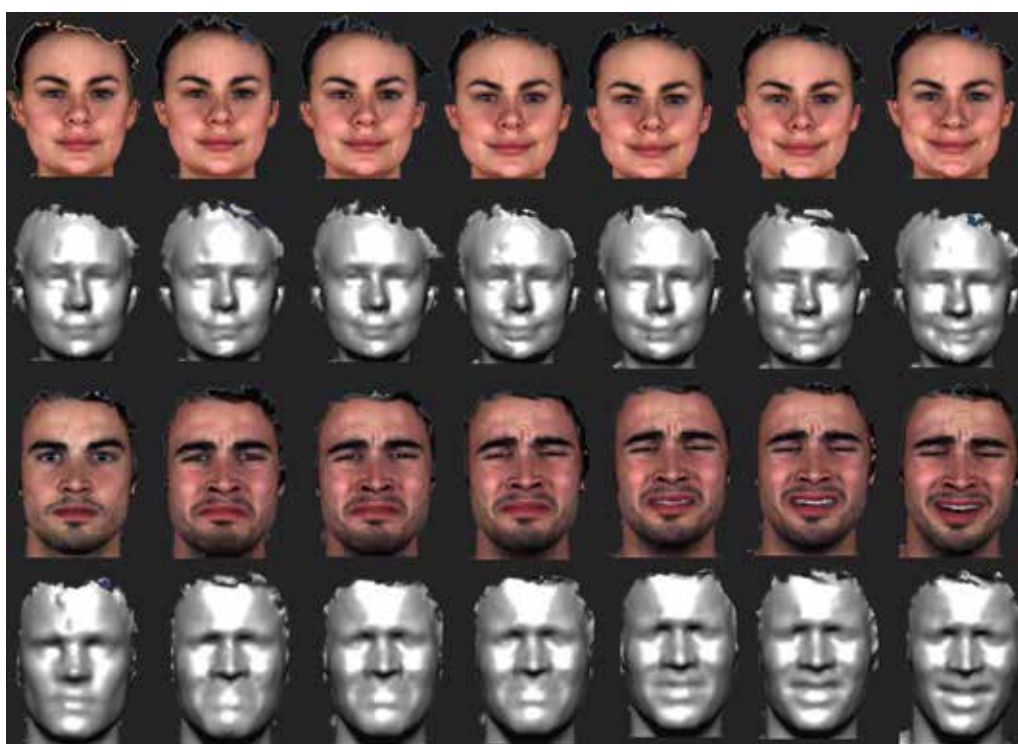


Fig. 18. Two examples from the ADSIPmark1 database showing happiness expression in normal intensity level (top two rows) and sadness expression in extreme intensity level (bottom two rows).

shows a couple of examples from the ADSIPmark1 database. That database is being gradually expanded. The new acquisitions are captured at 60 fps. Furthermore, some additional facial articulations with synchronised audio recording are captured, with each subject reading a number of predefined special phrases typically used for the assessment of neurological patients (Quan et al., 2010a; Quan et al., 2010b). The final objective of the ADSIP database is to contain 3-D dynamic facial data of over 100 control subjects and additional 100 subjects with different facial disfunctions. A couple of examples of this currently extended ADSIP database are shown in Figure 19.



Fig. 19. Examples from the extended ADSIP database with anger expression in normal intensity level (top two rows with two different views) and fear expression in extreme intensity level (bottom two rows with two different views).

4.5 Database validation

Facial expressions are very subjective in nature. In other words, some of expressions are difficult to interpret and classify even for human observers who are normally considered as the “best classifier” for facial expressions. In order to validate the quality of a 3-D facial expression database, human observers have to be involved to see whether a particular facial expression, performed by a subject in response to a given instruction, is executed in a way which is consistent with the human perception of that expression. Use of human observers in validation of a facial database enables the assumed ground truth to be benchmarked, as it

provides a performance target for facial expression algorithms. It is expected that the performance of the best automatic facial expression recognition systems should be comparable with the human observers (Black & Yacoob, 1997; Wang & Yin, 2007).

Using the video clips recorded simultaneously with dynamic 3-D facial scan, the first part of the ADSIP database was assessed by 10 invited observers. They were the staff and students at the University of Central Lancashire. A bespoke computer program was designed to present the recorded video clips and collect the confidence ratings given by each of the observers. Participants were shown one video clip at a time and were asked to enter their confidence ratings against seven categories for expressions: anger, disgust, fear, happiness, pain, sadness and surprise, with the confidence ratings selected from the range of 0 to 100% for each category. To reflect possible confusions from observers about an expression for a given video clip, ratings could be distributed over the various expression categories as long as scores added up to 100%. Table 1 presents the confidence scores for each expression averaged over all video clips scored by the all observes. It can be seen that happiness expressions were given near perfect confidence scores, and anger, pain and fear were the worst rated with fear scored below 50%. Also, the 'normal' intensity level was somewhat better rated than 'mild', and 'extreme' was also somewhat better than 'normal'. Table 2 shows the confidence confusion matrix for the seven expressions. It can be seen that the observers were again very confident about recognising the happiness expression whereas the fear expression was often confused with the surprise expression (Frowd et al., 2009).

Intensity	Anger (%)	Disgust (%)	Fear (%)	Happiness (%)	Sadness (%)	Surprise (%)	Pain (%)	Mean (%)
Mild	47.5	51.5	43.3	90.3	72.9	72.9	57.4	57.9
Normal	56.6	78.3	41.5	94.3	75.6	75.6	62.0	65.7
Extreme	61.4	80.7	48.4	96.0	74.0	74.0	75.7	70.3
Mean (%)	55.2	70.2	44.4	93.5	74.2	74.2	65.0	64.6

Table 1. Mean confidence scores for seven expressions.

Input/Output	Anger (%)	Disgust (%)	Fear (%)	Happiness (%)	Sadness (%)	Surprise (%)	Pain (%)
Anger	55.39	26.03	5.19	0.00	5.13	5.31	2.94
Disgust	7.70	68.86	5.22	0.00	8.47	4.59	5.16
Fear	3.80	9.02	46.90	0.00	7.13	23.90	9.26
Happiness	0.27	0.98	0.71	92.95	1.15	2.35	1.59
Sadness	4.07	5.87	3.63	0.71	74.15	3.22	8.33
Surprise	0.60	7.54	21.84	1.04	2.46	64.64	1.88
Pain	4.94	9.45	9.46	2.30	18.96	3.85	51.04

Table 2. Confidence confusion matrix for the human observers.

5. Evaluation of expression recognition using BU-3DFE database

In order to characterise the performance of the SSV based representation for facial expression recognition, its effectiveness is demonstrated here in two ways. At first, the

distribution of the low-dimensional SSV-based features, extracted from the faces with various expressions is visualised, thereby showing the potential of the SSV-based features for facial expression analysis and recognition. Secondly, standard classification methods were used with the SSV-based features to quantify their discriminative characteristics.

5.1 Visual illustration of expression separability

Since it is difficult to visualise SSV-based features in a space with more than three dimensions, only the first three elements of the SSV are used to reveal its clustering characteristics and discriminative powers. An SSM was built using 450 BU-3DFE faces which were randomly selected from the database. The SSV-based features extracted from another 450 BU-3DFE faces which cover 18 individuals with six universal expressions were chosen for the visual illustration in 3-D shape space. It can be found that the SSV-based feature exhibits good expression separability even in a low-dimensional space, especially for those expression such as “anger vs. disgust”, “disgust vs. surprise” and “fear vs. happiness”. Examples of the expressions are given in Figure 20, where the SSV-based features representing these expressions are seen to form relatively well defined clusters in the low-dimensional shape space. Although some parts of the clusters slightly intersect with each other, the clusters can be identified easily.

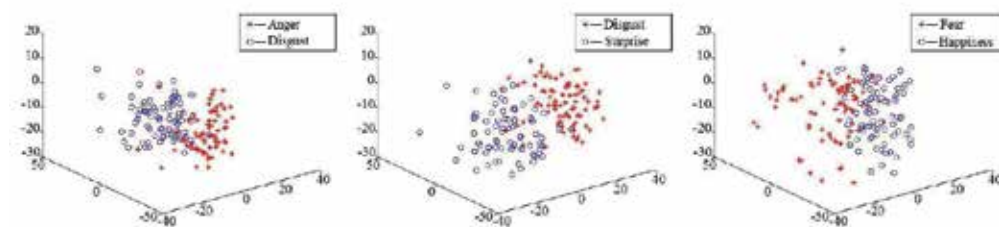


Fig. 20. Visualisation of expression separability in the low-dimensional SSV feature space: (a) anger vs. disgust, (b) disgust vs. surprise, and (c) fear vs. happiness.

5.2 Expression recognition

To better evaluate the discriminative characteristics of the vectors in the shape space, quantitative results of facial expression recognition are shown in this section. Several standard classification methods were employed in the experiments. They are linear discriminant analysis (LDA), quadratic discriminant classifier (QDC), and nearest neighbour classifier (NNC) (Duda, 2001; Nabney, 2004). 900 faces from the BU-3DFE database were used for testing, which were divided into six subsets with each subset containing 150 randomly selected faces. During the experiment, one of the subsets was selected as the test subset while the remaining subsets were used for learning. Such experiment was repeated six times, with the different subset selected as the test subset each time. Table 3 shows the averaged results for expression recognition achieved by the three classifiers. It can be seen that the LDA achieved the highest recognition rate of 81.89%.

Table 4 shows the confusion matrix of the LDA classifier. It can be seen that the anger, happiness, sadness and surprise expressions are all classified with above 80% accuracy, whereas the fear expression is only classified correctly for around 73%. This is consistent with the validation results for the ADSIP database discussed in Section 4.5, which showed that the fear expression is often confused with other expressions by the human observers.

Classifiers	LDA	QDC	NNC
Recognition rate	81.89%	80.11%	78.53%

Table 3. Average recognition rate of faces from BU-3DFE database using the SSV-based features.

Input/Output	Anger (%)	Disgust (%)	Fear (%)	Happiness (%)	Sadness (%)	Surprise (%)
Anger	82.64	3.48	4.17	3.47	4.86	1.39
Disgust	7.64	78.47	3.48	5.56	2.08	2.78
Fear	4.27	3.57	72.59	12.50	5.66	1.39
Happiness	2.78	5.56	8.33	83.33	0.00	0.00
Sadness	4.17	3.47	11.11	0.00	81.25	0.00
Surprise	0.00	0.00	4.17	2.78	0.00	93.06

Table 4. Confusion matrix of the LDA classifier for BU-3DFE database.

6. Evaluation of expression recognition using ADSIP database

The performance evaluation using the ADSIP database is an extension of that based on the BU-3DFE database, which aims to check the stability of the proposed methodology with a statistical shape model built from one database and tested using a different database. Furthermore, it is also used to assess the performance of the proposed algorithm against human observers.

As explained in Section 5.2, the statistical shape model was built based on 450 faces randomly selected from the BU-3DFE database. One hundred static 3-D faces were randomly chosen as the testing faces from the 210 dynamic 3-D facial sequences in the ADSIP database, with each one occurring approximately at the maximum of the expression. The selected test set contained 10 subjects with six different facial expressions at various intensities. For each testing face represented by its SSV, the three classifiers LDA, QDC and NNC, were used for the facial expression classification.

Table 5 shows the average results obtained from expression recognition using the three classifiers. It can be seen that the LDA again achieved the highest recognition rate of 72.84%. Although this result is worse than the result obtained using the test faces from the BU-3DFE database, it is around 2.5% higher than the mean recognition rate achieved by the human observers for those 'extreme' expressions in the same ADSIP database shown in Table 1, and it is much better than the human observers' mean recognition rates for the other two intensity levels of expressions.

Classifiers	LDA	QDC	NNC
Recognition rate	72.84%	69.31%	70.54%

Table 5. Average recognition rate of the SSV-based feature from ADSIP database.

The confusion matrix of the LDA is shown in Table 6. It can be seen that the happiness is the most recognisable expression and followed by the surprise expression. Fear and sadness are the expressions which are often confused with others.

Input/Output	Anger (%)	Disgust (%)	Fear (%)	Happiness (%)	Sadness (%)	Surprise (%)
Anger	73.08	11.54	0.00	1.92	13.46	0.00
Disgust	11.54	71.15	9.62	1.92	1.92	3.85
Fear	1.92	11.54	55.77	17.31	13.46	0.00
Happiness	2.08	0.00	10.83	87.08	0.00	0.00
Sadness	22.92	0.00	10.42	0.00	66.67	0.00
Surprise	0.00	6.25	8.33	0.00	2.08	83.33

Table 6. Confusion matrix of the LDA classifier for ADSIP database.

7. Discussion and conclusions

This chapter has introduced concepts related to automatic facial expression recognition. Although these have included description of general issues relevant to such problem, the main emphasis has been on a review of the recent developments in the corresponding processing pipeline including: data acquisition, face normalisation, feature extraction and subsequent classification. The available facial expression databases were also introduced to provide complete information about available options for algorithm validation. To make these ideas clearer one of the algorithm, using shape space vector (SSV) calculated for 3D static facial data, was described in detail. The algorithm consists of model building and model fitting. In the model building stage, a statistical shape model (SSM) is constructed from a set of training data. In the model fitting stage, the built SSM is aligned and matched to the input faces through iterative estimation of pose and shape parameters, and such estimated shape parameters (embedded in the SSV) are used as a feature vector for facial expression recognition.

In order to examine the facial expression recognition performance offered by the SSV representation method, two recently developed 3-D facial expression databases were used. The evaluation results based on the BU-3DFE database show that the SSV based representation method is capable of simulation and interpretation of 3-D human facial expressions. The test performed with the ADSIP database indicates that the method can be used with data collected using different acquisition protocols. More significantly, the recognition rates for different facial expressions obtained using the SSV based representation are shown to be similar to those obtained using the human observers.

The temporal information plays a vital role in the facial expression recognition. An analysis of a dynamic facial sequence instead of a single image is likely to improve the accuracy and robustness of the facial expression recognition systems (Sun & Yin, 2008). To reflect this new trend in facial data analysis the chapter includes short information about available 3D dynamic datasets.

8. References

- Aizawa, K. & Huang, T. S. (1995). Model-based image coding: advanced video coding techniques for very low bit-rate applications. *Proceedings of IEEE*, Vol.83(2), pp. 259-271.

- Battocchi, A.; Pianesi F. & Goren-Bar, D. (2005). DaFEx: database of facial expressions. Intelligent Technologies for Interactive Entertainment, *Lecture Notes in Computer Science*, Vol.3814(2005), pp. 303-306.
- Bartlett, M. S.; Littlewort, G.; Fasel, I. & Movellan J. R. (2003). Real time face detection and facial expression recognition: development and applications to human computer interaction. *CVPR workshop for HCI*.
- Bernardini, F. & Rushmeier, H. E. (2002). The 3D model acquisition pipeline. *Computer Graphics Forum*, Vol.21(2), pp. 149-172.
- Besl, P. J. & McKay, N. D. (1992). A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14(2), pp. 239-256.
- Bishop, C., M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Black, M. J.; Fleet, D. J. & Yacoob, Y. (1998). A framework for modeling appearance change in image sequence. *6th International Conference on Computer Vision*.
- Black, M. J. & Yacoob, Y. (1997). Recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision*, Vol.25(1), pp. 23-48.
- Blake, A. & Isard, M. (1998). *Active Contours*. Springer-Verlag Berlin and Heidelberg.
- Blanz, V. & Vetter, T. (2003). Face recognition based on fitting a 3d morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25(9), pp. 1063-1074.
- Blanz, V. & Vetter, T. (1999). A morphable model for the synthesis of 3-D faces. *SIGGRAPH*, pp. 187-194.
- Bookstein, F. L. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.11(6), pp. 567-585.
- Brahnam, S.; Chuang, C.; Shih, F. Y. & Slack M. R. (2006). Machine recognition and representation of neonatal facial displays of acute pain. *Artificial Intelligence in Medicine*, Vol.36(3), pp. 211-222.
- Breiman, L. (2001). Random forests. *Machine Learning*, Vol. 45(1), pp. 5-32.
- Cohen, I.; Sebe, N.; Garg, A.; Chen, L. & Huang T. S. (2003). Facial expression recognition from video sequences: temporal and static modelling. *Computer Vision and Image Understanding*, Vol. 91, pp. 160-187.
- Cootes, T. F.; Taylor, C. J.; Cooper, D. H. & Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, Vol.61(1), pp. 38-59.
- Cootes, T. F. & Taylor, C. J. (1996). Data driven refinement of active shape model search. *British Machine Vision Conference*, pp. 383-392.
- Cootes, T. F.; Edwards, G. J. & Taylor, C. J. (1998). Active appearance models. *European Conference on Computer Vision*, pp. 484-498.
- Curless, B. (2000). From range scans to 3D models. *ACM SIGGRAPH Computer Graphics*, Vol.33(4), pp. 39-41.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. J. Murray, London
- Dimensional Imaging (2010). <http://www.di3d.com>.
- Duda, R. O. (2001). *Pattern Classification*, second edition, John Wiley & Sons Inc.
- Edwards, G. J.; Cootes, T. F. & Taylor, C. J. (1998). Face recognition using active appearance models. *5th European Conference on Computer Vision*, pp. 581-595.

- Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Personality and Social Psychology*, Vol.17(2), pp. 124-129.
- Ekman, P. & Friesen, W. V. (1978). *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press.
- Eisert, P. & Girod, B. (1998). Analyzing facial expressions for virtual conferencing. *IEEE Computer Graphics & Applications*, Vol.18(5), pp. 70-78.
- Essa, I. A. & Pentland, A. P. (1997). Coding, analysis, interpretation and recognition of facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, pp. 757-763.
- FaceGen (2003). <http://www.facegen.com>. Singular Inversions.
- Fasel, B. & Luetttin, J. (2003). Automatic facial expression analysis: a survey. *Pattern Recognition*, Vol.36(1), pp. 259-275.
- Feldmar, J. & Ayache, N. (1996). Rigid, affine and locally affine registration of free-form surfaces. *International Journal of Computer Vision*, Vol.18(2), pp. 99-119.
- Fisher, R. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, Vol.7, pp. 179-188.
- Flanelli, G.; Yao, A.; Noel, P.-L.; Gall, J. & Van Gool, L. (2010). Hough forest-based facial expression recognition from video sequences. *International Workshop on Sign, Gesture and Activity (SGA) 2010, in conjunction with ECCV 2010*.
- Frangi, A.; Ruechert, D.; Schnabel, J. & Niessen, W. (2002). Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. *IEEE Transaction on Medical Imaging*, Vol.21(9), pp. 1151-1166.
- Freund, Y. & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Science*, Vol.55, pp. 119-139.
- Frowd, C. D.; Matuszewski, B. J.; Shark, L.-K. & Quan, W. (2009). Towards a comprehensive 3D dynamic facial expression database. *9th WSEAS International Conference on Multimedia, Internet and Video Technology (MIV'09)*, Budapest, Hungary.
- Gross, R.; Matthews I.; Cohn, J.; Kanade, T. & Baker, S. (2010). Multi-PIE. *Image and Vision Computing*, Vol.28(5), pp. 807 - 813.
- Hager, J. C.; Ekman, P. & Friesen, W. V. (2002). *Facial Action Coding System*. Salt Lake City, UT.
- Heisele, B.; Serre, T.; Pontil, M. & Poggio, T. (2001). Component-based face detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 657-662.
- Hoch, M.; Fleischmann, G. & Girod, B. (1994). Modeling and animation of facial expression based on B-splines. *Visual Computer*, pp. 87-95.
- Hong, H.; Neven, H. & Von der Malsburg, C. (1998). Online facial expression recognition based on personalized galleries. *2nd International Conference on Automatic Face and Gesture Recognition*, pp. 354-359.
- Hong, T.; Lee, Y.-B.; Kim, Y.-G. & Hagbae, K. (2006). Facial expression recognition using active appearance model. *Advance in Neural Network, Lecture Notes in Computer Science*, Vol.3972, pp. 69-76.
- Huang, T. M.; Kecman, V. & Kopriva, I. (2006). *Kernel Based Algorithms for Mining Huge Data Sets, Supervised, Semi-supervised, and Unsupervised Learning*. Springer-Verlag, Berlin.

- Isgro, F.; Trucco, E.; Kauff, P. & Schreer, O. (2004). Three-dimensional image processing in the future of immersive media. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.14(3), pp. 288-303.
- Ji, Y. & Idrissi, K. (2009). Facial expression recognition by automatic facial parts position detection with boosted-LBP. *5th International Conference on Signal Image Technology and Internet Based Systems*, Marrakech, Morocco, pp. 28-35.
- Johnson, H. & Christensen, G. (2002). Consistent landmark and intensity based image registration. *IEEE Transactions on Medical Imaging*, Vol.21(5), pp. 450-461.
- Jones, M. & Viola, P. (2003). Face recognition using boosted local feature. *International Conference on Computer Vision*.
- Kanade, T.; Cohn, J. F. & Tian, Y. (2000). Comprehensive database for facial expression analysis. *IEEE International Conference on Automatic Face and Gesture Recognition*.
- Kobayashi, H. & Hara, F. (1997). Facial interaction between animated 3D face robot and human beings. *International Conference on Systems, Man, Cybernetics*, pp. 3732-3737.
- Lanitis, A.; Taylor, C. J. & Cootes, T. F. (1997). Automatic interpretation and coding of face images using flexible models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19(7), pp. 743-756.
- Liao, S.; Fan, W.; Chung, A. & Yeung, D.-Y. (2006). Facial expression recognition using advanced local binary patterns, Tsallis entropies and global appearance feature. *IEEE International Conference on Image Processing*, pp. 665-668.
- Lisetti, C. L. & Schiano, D. J. (2000). Automatic facial expression interpretation: where human-computer interaction, artificial intelligence and cognitive science intersect. *Pragmatic and Cognition*, Vol.8(1), pp. 185-235.
- Littlewort, G.; Bartlett, M. S.; Fasel, I.; Susskind, J. & Movellan, J. (2005). Dynamics of facial expression extracted automatically from video. *Image and Vision Computing*, Vol.24, pp. 615-625.
- Lyons, M. J.; Budynek, J. & Akamatsu, S. (1999). Automatic classification of single facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.12, pp. 1357-1362.
- Lucey, P.; Cohn, J., F.; Kanade, T.; Saragih, J.; Ambadar, Z. & Matthews, I. (2010). The extended Cohn-Kanade Dataset (CK+): a complete dataset for action unit and emotion-specified expression. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, San Francisco, pages 94-101.
- Maalej, A.; Amor, B. B.; Daoudi, M.; Srivastava, A. & Berretti, S. (2010). Local 3D shape analysis for facial expression recognition. *20th International Conference on Pattern Recognition ICPR*, pp. 4129-4132.
- Md. Sohail, A. S. & Bhattacharya, P. (2007). Classification of facial expression using k-nearest neighbour classifier. *Lecture Note in Computer Science*, Vol.4418(2007), pp. 555-566.
- Morishima, S. & Harashima, H. (1993). Facial expression synthesis based on natural voice for virtual face-to-face communication with machine. *Virtual Reality Annual International Symposium*.
- Niese, R.; Al-Hamadi, A. & Michaelis, B. (2007). A novel method for 3D face detection and normalization. *Journal of Multimedia*, Vol.2(5), pp. 1-12.
- Nabney, I. T. (2004). *NETLAB: Algorithms for Pattern Recognition*, Springer-Verlag London.

- Nusseck, M.; Cunningham, D. W.; Wallraven, C. & Bulthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision*, Vol.8, pp. 1-23.
- Ojala, T.; Pietikainen, M. & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24(7), pp. 971-987.
- Padgett, C.; Cottrell, G. & Adolphs, R. (1996). Categorical perception in facial emotion classification. *18th Annual Cognitive Science Conference*, pp. 249-253.
- Pantic, M. & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: the state of art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22(12), pp. 1424-1445.
- Pantic, M.; Valstar, M. F.; Rademaker, R. & Maat, L. (2005). Web-based database for facial expression analysis. *Proc. IEEE International Conference on Multimedia and Expo (ICME'05)*, Amsterdam, Netherlands.
- Park, H. & Park, J. (2004). Analysis and recognition of facial expression based on point-wise motion energy. *Lecture Notes in Computer Sciences*, Vol.3212(2004), pp. 700-708.
- Pearson, D. E. (1995). Developments in model-based video coding. *Proceedings of IEEE*, Vol.83(6), pp. 892-906.
- Pentland, A. (2000). Looking at people: sensing for ubiquitous and wearable computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22(1), pp. 107-119.
- Pilz, K.; Thornton, I. M. & Bulthoff, H. H. (2006). A search advantage for faces learned in motion. *Experimental Brain Research*, Vol.171(4), pp. 436-447.
- Pollak, S. D. & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology*, Vol.38(5), pp. 784-791.
- Quan, W. ; Matuszewski, B. J. ; Shark, L.-K. & Ait-Boudaoud, D. (2007a). Low dimensional surface parameterisation with applications in biometrics. *IEEE 4th International Conference on Biomedical Visualisation (MediViz07)*, pp. 15-20.
- Quan, W.; Matuszewski, B. J.; Shark, L.-K. & Ait-Boudaoud, D. (2007b). 3-D facial expression representation using B-spline statistical shape model, *BMVC Vision, Video and Graphics Workshop*, Warwick.
- Quan, W. ; Matuszewski, B. J. ; Shark, L.-K. & Ait-Boudaoud, D. (2008). 3-D facial expression representation using statistical shape models. *BMVA Symposium on 3D Video Analysis, Display and Applications*, Royal Academy of Engineering, London.
- Quan, W. (2009a). 3-D facial expression representation using statistical shape models. PhD Thesis, University of Central Lancashire.
- Quan, W.; Matuszewski, B. J.; Shark, L.-K. & Ait-Boudaoud, D. (2009b). 3-D facial expression biometrics using statistical shape models. *Special Issue on Recent Advances in Biometric Systems: A Signal Processing Perspective, EURASIP Journal on Advances in Signal Processing*, Vol.2009, Article ID 261542, pp. 1-17.
- Quan, W.; Matuszewski, B. J. & Shark L.-K. (2010a). A statistical shape model for deformable surface registration. *International Conference on Computer Vision Theory and Applications*, Angers, France.
- Quan, W.; Matuszewski, B. J. & Shark L.-K. (2010b). Improved 3-D facial representation through statistical shape model. *IEEE International Conference on Image Processing (2010 ICIP)*, pp. 2433-2436.

- Ramanathan, S.; Kassim, A.; Venkatesh, Y. V. & Wah, W. S. (2006). Human facial expression recognition using a 3D morphable model. *IEEE International Conference on Image Processing*.
- Ruechert, D.; Frangi, A. F. & Schnabel, J. A. (2003). Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, Vol.22(8), pp. 1014-1025.
- Saxena, A.; Anand, A. & Mukerjee, A. (2004). Robust facial expression recognition using spatially localized geometric model. *International Conference on Systemic, Cybernetics and Informatics*, pp. 124-129.
- Steffens, J.; Elagin, E. & Neven, H. (1998). PresonSpotter – fast and robust system for human detection tracking and recognition. *2nd International Conference on Face and Gesture Recognition*, pp. 516-521.
- Sim, T.; Baker, S. & Bsat, M. (2002). The CMU pose, illumination, and expression (PIE) database. *5th IEEE International Conference on Automatic Face and Gesture Recognition*.
- Sun, Y. & Yin, L. (2008). Facial expression recognition based on 3D dynamic range model sequences. *10th European Conference on Computer Vision (ECCV08)*.
- Suwa, M.; Sugie, N. & Fujimora, K. (1978). A preliminary note on pattern recognition of human emotional expression. *4th International Joint Conference on Pattern Recognition*, pp. 408-410.
- Trier, O. D.; Taxt, T. & Jain, A. (1997). Recognition of digits in hydrographic maps: binary vs. topographic analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19(4), pp. 399-404.
- Umeyama, S. (1991). Least-square estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.13(4), pp. 376-380.
- Vural, E.; Cetin, M.; Littlewort, G.; Bartlett, M. & Movellan, J. (2007). Drowsy driver detection through facial movement analysis. *Lecture Note in Computer Science*, No.4796, pp. 6-18.
- van Dam, A. (2000). Beyond WIMP, *IEEE Computer Graphics and Applications*, Vol.20(1), pp. 50-51.
- Vrtovec, T.; Tomažević, D.; Likar, B.; Travnik, L. & Pernuš, F. (2004). Automated Construction of 3D statistical shape models. *Image Analysis and Stereology*, Vol.23, pp. 111-120.
- Wallhoff, F. (2006). Facial expressions and emotion database. www.mmk.ei.tum.de/~waf/fgnet/feedtum.html, Technical University Munich.
- Wang, J.; Yin, L.; Wei, X. & Sun, Y. (2006a). 3D facial expression recognition based on primitive surface feature distribution. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2006)*, pp. 17-22.
- Wang, Y.; Pan, G.; Wu, Z. & Wang, Y. (2006b). Exploring facial expression effects in 3D face recognition using partial ICP. *Lecture Notes in Computer Sciences*, Vol.3851(2006), pp. 581-590.
- Wang, J. & Yin, L. (2007). Static topographic modeling for facial expression recognition and analysis. *Computer Vision and Image Understanding*, Vol.108(1-2), pp. 19-34.
- Whitehill, J.; Littlewort, G.; Fasel, I.; Bartlett, M. & Movellan, J. (2009). Towards practical smile detection. *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol.31(11), pp. 2106-2111.

- Wimmer, M.; MacDonald, B. A.; Jayamuni, D. & Yadav, A. (2008). Facial expression recognition for human-robot interaction: a prototype. *2nd International Conference on Robot Vision*, pp. 139-152.
- Yin, L. ; Loi, J. & Xiong, W. (2004). Facial expression representation and recognition based on texture augmentation and topographic masking. *12th Annual ACM International Conference on Multimedia*, pp. 236-239.
- Yin, L.; Wei, X.; Sun, Y.; Wang, J. & Rosato, M. (2006). A 3D facial expression database for facial behavior research. *7th International Conference on Automatic Face and Gesture Recognition (FG2006)*, IEEE Computer Society TC PAMI, pp. 211-216.
- Yin, L.; Chen, X.; Sun, Y.; Worm T. & Reale, M. (2008). A High-Resolution 3D Dynamic Facial Expression Database. *8th International Conference on Automatic Face and Gesture Recognition (FGR08)*, IEEE Computer Society TC PAMI. Amsterdam, The Netherlands.
- Yousefi, S.; Nguyen, M. P.; Kehtarnavaz, N. & Cao, Y. (2010). Facial expression recognition based on diffeomorphic matching, *17th IEEE International Conference on Image Processing*, pp. 4549-4552.
- Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, Vol.13(2), pp. 119-152.
- Zhang, S. & Huang, P. (2006). High-resolution, real-time 3-D shape measurement. *Optical Engineering*, Vol.45(12), pp. 123601-1-8.

Implications of Adult Facial Aging on Biometrics

Midori Albert, Amrutha Sethuram and Karl Ricanek
*University of North Carolina Wilmington,
Institute for Interdisciplinary Studies in Identity Sciences
USA*

1. Introduction

1.1 Statement of the problem

Single features of the human face, facial components, as well as the human face taken as a whole may be viewed as a biometric tool for purposes of individual human identification. Using computer technology, automated face recognition (FR) systems have been created to match individual faces from print and digital photographs and video to faces of the same individual whose image is stored among many others in a computer database. One of the primary problems that arise with FR systems is how to contend with the passage of time and the resultant effects of facial aging. Enrolled, i.e. gallery, images become difficult to match against a query, i.e. probe, image as the time span between the gallery and the probe increases. Several researchers have highlighted this temporal performance degradation (Pentland & Choudhury, 2000; Phillips et al., 2000; Ricanek et al., 2006). To meet this challenge, two approaches have been taken: (1) research what is known about adult age-related craniofacial morphological changes to better understand how to (2) synthetically age individuals from a facial image, or rather, develop workable artificial age progression techniques to anticipate how an individual's face may appear years down the road. In addition to probing the question of how a person's face may age across the adult lifespan, other facial appearance questions have arisen, such as how might a person's face change simply over the course of one day? Herein we present an overview of face aging, its implications on biometrics, specifically FR systems, and provide an example of our work via relaying results of experiments designed to measure facial age changes and artificially synthesize facial images, all in an effort to improve FR systems and ensure that facial features, when used as a biometric tool, are reliable, consistent, and accurate with replicable results.

1.2 Overview

As adults age, myriad changes occur diachronically within the craniofacial complex. Notable soft tissue modifications may be seen across each decade of adult life that passes. As well, subtle hard tissue or bony changes slightly alter the overall shape of the human face, mainly in the dentoalveolar region. These age-related changes affect the accuracy and efficacy of biometric techniques concerning face recognition, namely computer automated face recognition (FR) systems. At the University of North Carolina Wilmington (UNCW), the Institute for Interdisciplinary Studies in Identity Sciences (IISIS), is composed of a team of researchers (formerly the Face Aging Group), dedicated to studying a wide variety of topics

related to face aging, facial age progression or synthetic face aging, and facial expression, with the goal of improving computer automated face recognition (FR) technology. IISIS comprises faculty and professional researchers, a post-doctoral scholar, graduate and undergraduate students and high school interns working on multiple aspects of human identification using facial features. In this chapter, we provide an overview of intrinsic and extrinsic factors that affect human adult craniofacial aging, and the ways in which these changes have been considered in research devoted to developing effective FR systems. We include an example of our work; we present results of an experiment conducted to evaluate FR systems and discuss synthetic aging techniques. We conclude with future goals.

2. Intrinsic and extrinsic factors causing facial aging

2.1 Intrinsic face aging factors

Intrinsic aging is caused by internal biological factors whereas extrinsic aging is caused by environmental influences. Intrinsic factors causing facial aging are mainly due to the natural changes that occur as soft tissues lose their elasticity, muscle tone, and volume (Coleman & Grover, 2006; Isse, 2008) as well as the bony shape modifications resulting from the lifelong and ongoing process of bone remodeling (Sadick et al., 2009; Zimblet et al., 2001). Affecting these changes are an individual's biological sex, ethnicity, and idiosyncratic features (i.e., features purely unique to the individual such as hyperdynamic facial expressions).

2.2 Extrinsic face aging factors

Extrinsic factors influencing facial aging may be due to lifestyle such as diet, drug use, and or smoking (although this is debated, (O'Hare et al., 1999; Taister et al., 2000)) but the main cause of skin aging is exposure to solar ultraviolet rays known as photoaging (Victor, 2003).

3. Body composition

Body composition is known to change as people age in adulthood. Muscle fibers atrophy and fat is accumulated. Appendages become thinner while abdominal girth expands due to a general weakening of abdominal muscles (Overfield, 1995). The ways in which muscle loss and weight gain affect the aging face are not entirely understood as there is a paucity of published literature specifically addressing weight gain and loss in and around the face. However, it is often anecdotally noted that weight gain in the face typically results in a more youthful appearance. This is due to facial folds and creases being buoyed and filled out by facial fat. Indeed several websites contain commentary that after weight loss of around 30 pounds, rhytides and wrinkles increased (Peertrainer., 2009). Generally, fat stored in the face shifts position as age advances. Fat from the upper cheeks diminishes and the face may take on a more hollow appearance. Fat that droops toward the jawline creates the appearance of jowls (Lipotherapy, 2009).

4. Ways in which the adult craniofacial complex changes with age

Adult age-related craniofacial changes comprise soft tissue and hard tissue changes. Beginning as early as the 20's, the soft tissues of the face age; however, the changes are most noticeable between the 40's and 50's and continue to become more pronounced between the 50's and 60's (Albert et al., 2007; Taylor, 2001). Bony changes that occur in adolescence and young adulthood are most likely a continuation of skeletal growth and development, while changes due to remodeling occur from young adulthood on into senescence. Some typical

aging skin changes are as shown in Figure 1. Soft tissue changes are presented first, followed by bony changes.



Figure 1. Typical aging skin changes. 1 = thinning hair and receding hairline; 2 = forehead rhytids and ptosis; 3 = glabellar rhytidosis; 4 = brow ptosis; 5 = temple rhytidosis and ptosis; 6 = upper lid redundancy and ptosis; 7 = lateral canthal rhytidosis; 8 = nasal root rhytidosis; 9 = lower lid redundancy and rhytidosis; 10 = lower lid fat pseudoherniation; 11 = malar bag formation; 12 = cheek rhytidosis; 13 = preauricular rhytidosis; 14 = nasal tip ptosis and dependency; 15 = cheek sagging and fat atrophy changes; 16 = deepening nasolabial crease; 17 = facial rhytidosis and sagging; 18 = perioral rhytidosis; 19 = upper lip flattening and lengthening; 20 = thinning and atrophy of vermillion (red lip); 21 = chin pad ptosis and retraction; 22 = jowl formation; 23 = cervical rhytidosis; 24 = submental fat accumulation; 25 = platysmal banding; 26 = rhytidosis and mid-neck hollowing; 27 = submaxillary gland ptosis.

Fig. 1. Typical aging skin changes

4.1 Soft tissue age changes: Upper third of the face

In the upper third of the face, decreasing skin elasticity, and the effects of gravity and repetitive periorbital muscle contractions, result in a tendency for brows to droop. Upper eyelid skin elasticity fades which causes an excess of unsupported skin that leads to folds; the folds in the eyelids create a tired or "old" facial appearance (Sadick et al., 2009). The ways in which these age-related changes in the periorbital region vary among individuals is of keen interest in biometric studies, namely face recognition. Studies investigating ethnic and sex differences in the periorbital region are of prime interest where FR research and development is concerned. (Kunjur et al., 2006) provide some clues as to how eyebrows and eyelids vary in three different ethnic populations: European, Indian, and Chinese. From a sample of 26 individuals, aged 18 to 25 years, a series of measurements were taken from photographs: eyebrow thickness, eyebrow length, arch, upper eyelid length and height, and of the palpebral fissure (the crease) the width, height, axis, and intercanthal (inner corner of the eye to the outer corner of the eye) distance. Results indicated that for the eyebrows, there was a significant difference between Indian and Chinese males; but, for the measurements of the upper eyelids, there were significant differences between all three populations in both females and males. For the palpebral fissure, there was a significant difference between Indian and European males, with

Indian males displaying wider palpebral fissures. The position of the axis of the palpebral fissure was statistically significant between European and Indian females. Kunjur et al. (2006) remark that previous research on facial features of individuals of North American European, Chinese, Turkish, African-American, and Asian-American descent all showed differences. How this may affect adult aging and the appearance of aged faces as it ties to face recognition is an area researchers associated with IISIS are targeting. Beyond the upper face, the mid-facial region and lower facial region display age-related soft tissue changes as well.

4.2 Soft tissue age changes: Middle third of the face

In the mid-facial region, fat deposition in the cheeks begins to fade with advancing age. The associated ligaments holding the malar fat pad in place weaken and a nasolabial fold develops (Sadick et al., 2009). While nasolabial lines begin to form in the 20's to 30's, the folds increase in depth in the 40's to 50's and continue to deepen into the 60's and beyond (Albert et al., 2007). There is evidence that the nose moves forward and downward with advancing age (Bishara et al., 1998; Sarnas & Solow, 1980; West & McNamara, 1999). Beyond the changes in the midface, in the cheeks and nose, and the upper face, eyebrows and eyelids, mentioned above, there are age-related changes in the lower face that may potentially significantly alter an individual's appearance after a number of years, mainly tied to the perioral region, the area around the mouth.

4.3 Soft tissue age changes: Lower third of the face

In the lower face, vertical rhytides tend to form above the vermilion border due to skin thinning with age (Sadick et al., 2009) and an oromental groove, or lines forming at the corners of the mouth, develop mainly in the 40's and deepen as age increases (Albert et al., 2007). A buccomandibular crease may arise; this is crease that forms at the side of the cheek above the lower jaw and is oriented vertically. Jowls, along with a sagging chin (i.e., fat deposition below the jawline), tend to occur in the 50's and become more pronounced in the 60's and beyond. Jowls form as a result of several mechanisms which include fat and volume loss, fat shifting and moving due to gravity, and changes in the mandible (Reece & Rohrich, 2008). Lips elongate with age and become thinner. Sforza et al. (2010) collected data on lip dimensions from 532 male and 386 female individuals aged 4 to 73 years. Linear dimensions included the width of the mouth, philtrum, and vermilion, and height of the upper, lower, and both lips. Lip volumes were also calculated for upper, lower and both lips. Regarding the findings for adults, vermilion areas and heights of the lower and total lips decreased with advancing age. Age-related changes occurring in the orolabial region likewise showed a significant sex-effect. Width of the mouth, width of the philtrum, lip height and lip volume all showed statistically significant sex differences, with males yielding larger measures than females as age advanced. However, the vermilion height to mouth width ratio was larger in females compared to males. Pecora et al. (2008), in their longitudinal study of craniofacial age changes, found that upper lip length increased an average of 3.2 mm for individuals first measured in their late teens then re-measured again in their 40's. Upper lip length increased about 1.4mm between the 40's and late 50's. Further, Pecora et al. (2008) noted that consistent with upper lip length increases, significant upper lip thinning occurred. Upper lips thinned an average of 3.6 mm from the late teens to 40's, and about 1.4 mm from the 40's to late 50's. Upper lip lengthening and thinning have also been confirmed by other researchers, such as Iblher et al. (2008). Given the phenotypic variation in lip length and volume—and numerous other soft tissue facial features—seen within and between populations of varying ethnic origins, the extent to which

facial feature changes affect an individual's facial appearance over the course of his or her adult lifespan is unknown. Biometric studies aimed at improving FR systems take this into consideration. Also noteworthy, but perhaps less relevant to face recognition, are age-related changes in the ear.

4.4 Soft tissue changes: Ear

Sforza et al. (2009) examined ear width, length, relevant ratios, ear areas and angles relative to the midline of the face, and asymmetry in 497 male and 346 female individuals aged 4 to 73 years. The researchers report a significant age effect; linear distances, areas and symmetry increased from childhood into senescence. They found that for all ear dimensions in adults, there was a statistically significant sex difference with males yielding measures larger than females.

5. Bony age changes in the skull

Craniofacial skeletal changes related to the adult aging process, or more specifically, bony remodeling which occurs throughout life, yields alterations in the morphology of the face. Shifts in craniofacial dimensions have been found along the horizontal, vertical, and sagittal planes, and in the dentoalveolar region (refer to Figure 2). However, many of these changes are often small, ranging from 1.1 mm to 1.60 mm, for example.

5.1 Bony age changes: Horizontal

Generally, the horizontal changes that occur include increases in overall head circumference, the length of the head (anterior to posterior), bizygomatic breadth (cheek bone to cheek bone), and the width of the head (Behrents, 1985). Horizontal size increases have been found to occur all throughout adult life, from the 20's to the 80's (Israel, 1977).

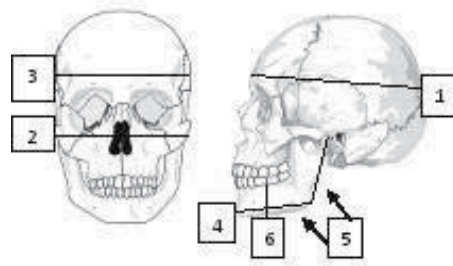


Fig. 2. The adult skull: 1. Head Length 2. Bizygomatic breadth 3. Head breadth 4. Mandibular length 5. Mandibular angle 6. Mandibular body height

5.2 Bony age changes: Vertical

Age changes in the vertical plane relate to anterior face height increases. Several researchers found that anterior face height tends to increase more in the lower part of the face than the upper face (Bondevi, 1995; Forsberg et al., 1991; West & McNamara, 1999). In the sagittal plane, slight anterior-posterior cranial thickness increases have been found to be highly statistically significant (Doual et al., 1997), although its direct contribution, if any, to soft tissue facial aging has yet to be reported and or studied.

5.3 Bony age changes: Dentoalveolar region

In the dentoalveolar region, changes from the middle 20's to middle 40's include a continued eruptive movement of the teeth likely tied to maxillary backward displacement, or retrusion (Akgul & Toygar, 2002; Forsberg et al., 1991). Mandibular length has also been shown to increase, though very slightly from the early 20's to early 30's (Bondevi, 1995). In study by Pecora et al. (2008), mandibular length was found to increase from their young age category (mean age 17 years) to their middle age category (mean age 47 years). This increase may be due in part to the young age group's mean age of 17 years, a stage in life where bone growth may not yet be completed because findings to the contrary also exist in the published literature. Indeed, Shaw et al. (2010), studying various dimensions of the mandible across three adult age groups—young (20-40 years), middle (41-64 years) and old (65 years and older)—found statistically significant decreases in mandibular length. For example, in males the mean mandibular length for the young age group (20-40 year olds) decreased from 90.4 mm to 83.1 mm in the middle age group (41-64 years), yet by old age (65 years and older), the mean mandibular length was 83.0 mm. In females, there was a similar pattern: the mean mandibular length decreased from 86.9 mm in the young age group (20-40 year olds) to 80.9 mm in the middle age group (41-64 years) and 80.8 mm in the old age group (65 years and older). Like many of the findings for both soft tissues and the bony craniofacial complex, major age changes seem to occur in the 40's and 50's. Shaw et al. (2010), measuring other dimensions of the mandible besides length across the adult lifespan, found significant increases in mandibular angle, particularly for the young to middle age groups. In females, mean angle measures in the young group aged 20-40 years significantly increased from 119.9° to 125.8° in the middle age group aged 41-64 years. Additionally, females also showed significant mandibular angle increases from the middle to old age groups (mean angle 125.8° for 41-64 year olds to 132.0° for 65 years and older). In males mean angle measurements increased from 112.3° to 124.1° to 125.6° from the young (20-40 year olds) to middle (41-64 year olds) to old age group (65 years and older), respectively. For males, increases were significant for the young to middle age groups, but not the middle to old age groups as seen in females. These increases in the mandibular angle correspond to the development of a softer and more oval appearance in the lower face as the jawline loses its definition, which may in turn be associated with the facial changes seen in jowl formation. Beyond increased mandibular angle and decreased mandibular length, Shaw et al. (2010) found statistically significant decreases in ramus height (the vertical portion of the mandible from the back corner of the jaw to the temporal-mandibular joint) as well as mandibular body height. The decreases in mandibular body height were not due to tooth loss; rather it seems generalized bone loss occurs as adult age progresses and may be tied to the finding of continued eruption of the teeth as noted earlier. Although changes in the bony face have been quantified, the degree to which the dimensional shifts discussed above affect the overall appearance of the aging face is not entirely clear. Research aimed at developing accurate age-progression techniques and face recognition (FR) systems as well as studies targeting specific facial features (e.g., eyebrow and eyelid shape) for use as a biometric tool in human facial identification, grapple with this limitation.

6. Adult age-related craniofacial morphological changes and the impact on biometrics, specifically face recognition (FR) systems

Inasmuch as a review of the literature on adult age-related craniofacial morphological changes in the soft tissues and bones confirms a certain pattern and general time line of changes,

predicting with accuracy and reliability exactly how a single individual's face may appear over the course of his or her adult life remains a challenge in biometric studies and identity sciences, particularly with regard to face recognition (FR) systems. To meet this challenge, the research team at IISIS endeavors to further advance the performance of synthetic age progression technologies. Synthetic aging is built in part upon the understanding of the general ways in which faces change with age while also taking into consideration the vast amount of individual human variation that is both idiosyncratic (unique to each person) and linked to sex differences, ethnic differences, and lifestyle—those intrinsic and extrinsic factors discussed at the beginning of this chapter.

6.1 Face recognition systems

Face recognition systems play a very important role in the modern day world. Some of the areas of application of face recognition include entertainment(video games, human computer interaction, virtual reality), identification(passports, driver's licenses, immigration control, voter registration etc), information security(medical records, internet/intranet access, application security) and law enforcement and surveillance (tracking a suspect, border control, video surveillance etc). These FR applications deal with facial images ranging from still controlled images to uncontrolled video sequences. No doubt, technical challenges posed by FR systems require a wide range of techniques involving image processing and image analysis.

Typical steps involved in a face recognition system is as shown in figure 3. Given an input image or video sequence to the FR system, the first step involved is the detection of a face. Various image segmentation algorithms are used to detect a face from the background. This is followed by feature extraction and face normalization. The last step is the actual recognition of a face, the output of which is usually an identification or a verification task. In the verification mode of authentication, the FR system verifies the claimed identity of the user by comparing his facial image to his enrolled image in the database. In the identification task, the system identifies a subject's face by matching the generated template to each of the templates enrolled in the gallery. There is ongoing research in the face recognition community to address the many challenges involved in each of these subtasks.

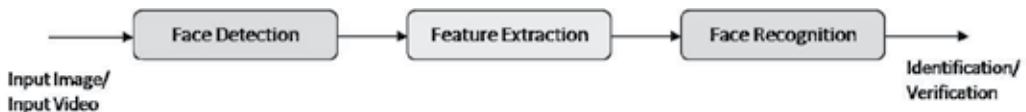


Fig. 3. Face recognition system: steps involved

Feature extraction methods in FR systems can be categorized into three types: (1) Generic methods - these methods are based on edges, lines and curves in an input image; (2) Template based methods - these methods are used to detect the actual facial features of the face such as the eyes, nose, mouth; (3) Structural matching methods - these methods take geometrical constraints on the features into consideration.

Based on the psychological study of how humans use holistic and local features for face recognition, FR techniques can be classified into the following three matching methods: (1) Holistic methods: These methods use the whole face as an input to a FR system. "Eigen faces", based on principal component analysis, is one of the widely used representations of such a method. Methods based on linear discriminant analysis (LDA) and probabilistic decision

based neural networks fall under this category. (2) Feature-based methods: Local features such as the eyes, nose and mouth are first extracted and information such as the location and local statistics are used for classification in these methods. Some of these methods are based on hidden markov model, convolution neural network and pure geometry. (3)Hybrid methods: These methods use a combination of both local features and the whole face region for face recognition. This is in line with the human perception methodology. Arguably, these methods should result in better recognition abilities. Some of the most promising methods are flexible appearance models based on shape-normalised techniques. Active shape models and active appearance models fall into this category.

Commercial face recognition systems are based on one of the FR techniques discussed above. Some of the commercial systems in use today are as listed in Table 1:

FACE RECOGNITION SDK	COMPANY	WEBSITE
FaceVACS	Cognitec Systems	http://www.cognitec-systems.de
Identix and Viisage	L-1 Identity Solutions	http://www.l1id.com
FTR-SDK	Pittpatt	http://www.pittpatt.com/products
FR security solutions	NEC	http://www.nec.com
Verilook SDK	Neurotechnology	http://www.neurotechnology.com
Toshiba face recognition	Toshiba	http://www.toshiba.com
FR Technology	Sagem	http://www.sagem-ds.com

Table 1. Commercial face Recognition Systems

6.2 Evaluation of face recognition systems

Due to the various techniques and theories that exist to address the problem of face recognition, evaluation of the algorithms is crucial. Some of the notable challenges faced by the FR systems are the variations in pose, illumination and expression(PIE). Some of the FR algorithms are developed to target a specific application. Thus the evaluation protocol must be based to evaluate the algorithms as closely as possible for the intended application. Over the past decade, several large publicly available face databases have been collected. Several appropriate testing protocols have been designed for evaluation purposes.

One of the early evaluation protocols were the series of FERET evaluations (FERET, URL). FERET (Face Recognition Technology) program was sponsored by the DoD Counterdrug Technology Development Program Office in 1993. This program aimed at providing a common baseline to evaluate FR algorithms. This was crucial in advancing FR research from theory to applied prototype system techniques. The goal of these evaluations was to develop automatic FR algorithms that could be used in applications such as security, intelligence and law enforcement. The collection and distribution of the FERET database was one of the valuable contributions of the program.

This was followed by the FRVT (Face Recognition Vendor Tests) (FRVT, URL) evaluations. FRVT provided independent government evaluations of commercially available and prototype face recognition technologies. These evaluations were designed to provide U.S. Government and law enforcement agencies with information to assist them in determining where and how facial recognition technology could be best deployed. FRVT 2000, FRVT 2002 and FRVT 2006 helped evaluate FR technology which had advanced from prototype systems to commercial systems. The performance of the systems were evaluated on real-life large-scale databases. FRVT 2006 was developed to assess if the goals of Face Recognition

Grand Challenge (FRGC) were met. In addition, FRVT results helped identify future research directions for the face recognition community.

More recently, the Multiple Biometric Grand Challenge (MBGC) (MBGC, URL) was formulated to build on the challenges and evaluation paradigm of FRGC and FRVT 2006. Some of the technology development areas emphasized by this program were (i) face recognition on still frontal, real-world-like high and low resolution imagery (ii) Iris recognition from video sequences and off-angle images (iii) Fusion of face and iris (at score and image levels) (iv) Unconstrained face recognition from still & video (v) Recognition from Near Infrared (NIR) & High Definition (HD) video streams taken through portals and (vi) Unconstrained face recognition from still & video.

Initiated in summer 2009, the Multi-biometric 2D Still-Face Recognition evaluation (MBE-STILL) was undertaken to respond to governmental and commercial requests to assess contemporary facial recognition (FR) implementations, to leverage massive operational corpora, to evaluate face recognition technologies in a proper one-to-many identification mode and to report parameters important to implementers and procurers. The first public report of the Multiple Biometrics Evaluation (MBE-STILL) was made available recently. Some of the results in the report included investigations on mode one-to-many search accuracy, dependence on population size, dependence on rank, impostor distribution stability, face ageing, influence of subject's weight etc.

Although the FR technology has improved considerably in the past decade, the problem of "facial aging" that affects the performance of the FR system is still considerable. The reports of the MBE concluded that for most verification algorithms, false non-match rates increase by roughly a factor of two over the eight year interval and that a more detailed statistical analysis was warranted. Thus it is clear that more work has to be done to address the problem of facial aging to improve the performance of FR systems. Some experiments conducted at the IISIS to evaluate the performance of a commercial FR engine on "aging" is discussed in the next section.

6.3 Experiment and results

To evaluate the performance of commercial face recognition systems on facial aging, a large set of longitudinal data for the same individual across adult lifespan is required. However, it has still been a challenge to gather such data. In this regard, IISIS is engaged in developing the MORPH database (Ricanek & Tesafaye, 2006) in an attempt to aid researchers involved in investigating the effects of facial aging. MORPH database is currently the largest publicly available longitudinal face database which is actively being used in over 30 countries. It is comprised of two datasets (albums). The first is a set of scanned photographs with up to 20 year age spans of 515 subjects. The second, album 2, is all digital photographs collected over several years that contains 55,606 longitudinal images. Both public albums include meta data for race, date of birth, date of acquisition, and gender. Eye coordinates for the sets are also available. These data can be requested at the group's website (FaceAgingGroup, URL) for research purposes. FGNET database (FG-Net, URL) is currently the only other publicly available database with access to longitudinal data.

To evaluate the performance of a FR system on longitudinal data, a commercial SDK was chosen. Typically, any FR system consists of a gallery and a probe. The gallery comprises a set of "known" individuals, while the probe consists of unknown facial images presented for recognition. A "match" score is obtained to match any probe image with the gallery image. To evaluate the effects of aging on the performance of the FR engine, the gallery set was made

to comprise the youngest images of individuals from the database. Older images of the same individuals formed the probe set. For any individual, match scores were obtained for each image in the probe against his/her image in the gallery. Some of the results of the match scores are shown in Figure 4. It can be seen from Figure 4 that the match score decreases when older images in the probe are matched to the younger image of the individual in the gallery. This is due to the effects of aging and variation in illumination, pose and expression in the probe images.

In addition, to evaluate the recognition rate of the SDK on a large dataset, a combined dataset of Morph-1 and FG-NET databases was used for this experiment. The gallery and the probes were so constructed such that they comprised images of a specific age group as shown in Table 2.

G/P	Age-Range(years)
Gallery (G)	18-30
Probe A (P_a)	18-30
Probe B (P_b)	31-40
Probe C (P_c)	41-50

Table 2. Age range for the gallery and probes for Experiment 1

The gallery data, G was constructed such that one youngest image of each person in the age range of 18-30 years available in the dataset was selected. The probe data P was constructed with aged images of the subjects in the Gallery. Three different probes P_a , P_b and P_c were generated to include aged images in different age-groups as shown in Table 2.

A quality rating of 80% or higher was used to include images in either the gallery or the probes. Input images were converted to 8-bit grayscale for all the experiments. Having set-up the data, the FR system was trained on the images in the gallery G . For every experiment with a Probe set matched with the Gallery, scores were obtained for each image in the probe compared with every image in the gallery. The match list obtained was then ranked with decreasing order of the scores generated.

In biometric systems, the cumulative match curve (CMC) is used as a measure of identification system performance. This measure judges the ranking capabilities of the system. The CMC curves were generated for different probes to measure the performance of the FR engine. This is as shown in Figure 5.

It is clear from Figure 5 that the performance of the algorithm decreases with increasing ages of the Probes. In other words, the FR system fails to identify subjects as they age when the gallery consists of younger images. Thus the engine is not robust to effects due to facial aging. Here at the IISIS, we try to address the problem of "Aging" by generating synthetic facial images to augment the gallery set. This is explained in the next section.





6.4 Synthetic facial aging

At IISIS, we have developed several algorithms for synthetic aging/de-aging a facial image. The techniques are built upon understanding of the fundamental principles of aging for youth and adults. A model-based approach is taken for the development of these algorithms. Active appearance models, a group of flexible deformable models, are used to generate our synthetic models. Figure 6 shows the general approach of our training algorithm.

The first step of the algorithm is to choose a set of clean (frontal pose, minimum variation in pose, illumination and minimum occlusion due to facial hair), high quality, training images with sufficient variation in gender, ethnicity and age range. These images are then annotated

p r o b e	Gallery			
		Age: 54 years	Age: 54 years	Age: 61 years
		0.9995	0.9993	0.8344
	Age: 53 years			




(a) Example 1: From Morph database

p r o b e	Gallery			
		Age: 46 years	Age: 47 years	Age: 48 years
		0.9994	0.8517	0.7652
	Age: 45 years			

(b) Example 2: From Morph database

p r o b e	Gallery				
		Age: 33 years	Age: 46 years	Age: 49 years	Age: 54 years
		0.4267	0.3433	0.183	0.0924
	Age: 30 years				

(c) Example 3: From FG-NET database

P r o b e	Gallery					
		Age: 26 years	Age: 41 years	Age: 50 years	Age: 54 years	
		0.9995	0.3015	0.1379	0.0479	0.0083
	Age: 24 years					

(d) Example 4: Individual collection

Fig. 4. Examples of match scores for longitudinal images

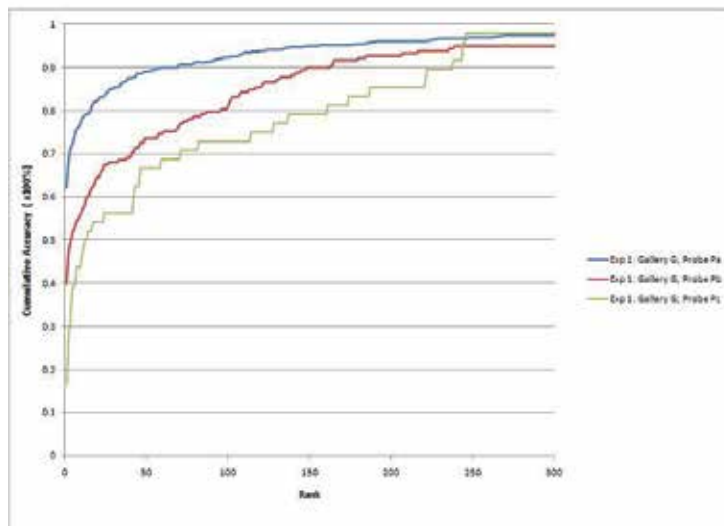


Fig. 5. CMC curves for the FR system

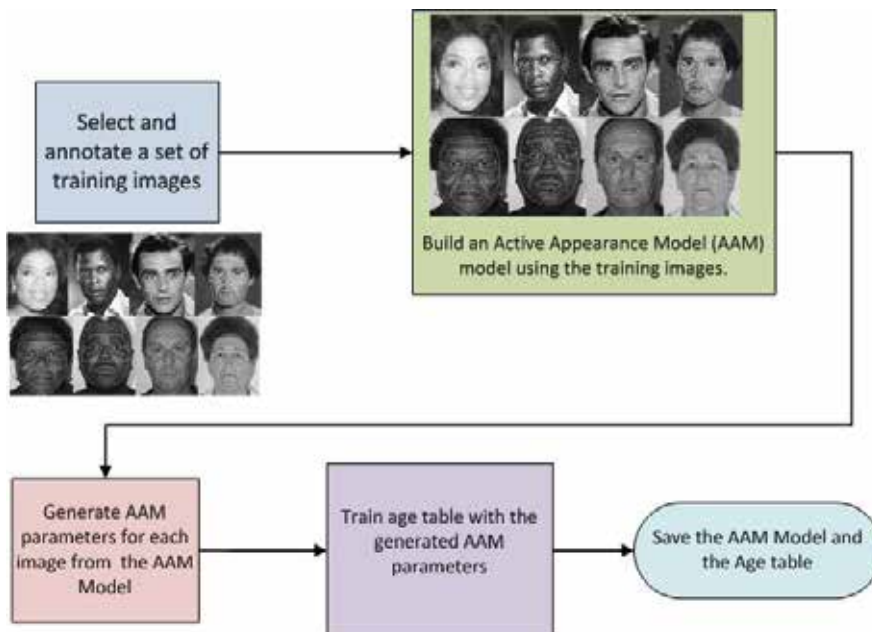


Fig. 6. Synthetic Facial Aging: Training phase

with a suitable annotation scheme as shown in Figure 7. An active appearance model (AAM) is then built from the set of training images and the corresponding AAM parameters are generated for each image. An "age table" is then trained using the generated AAM parameters. The AAM Model and the age table is then saved which completes the training phase of the algorithm as shown in Figure 6.



Fig. 7. An example of annotation scheme

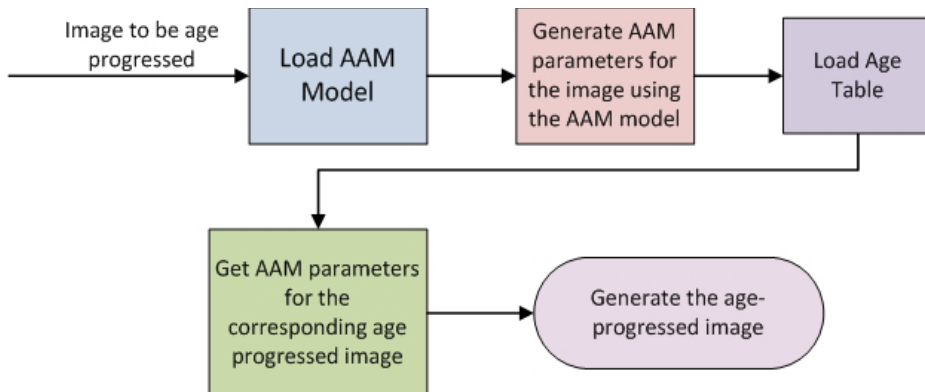
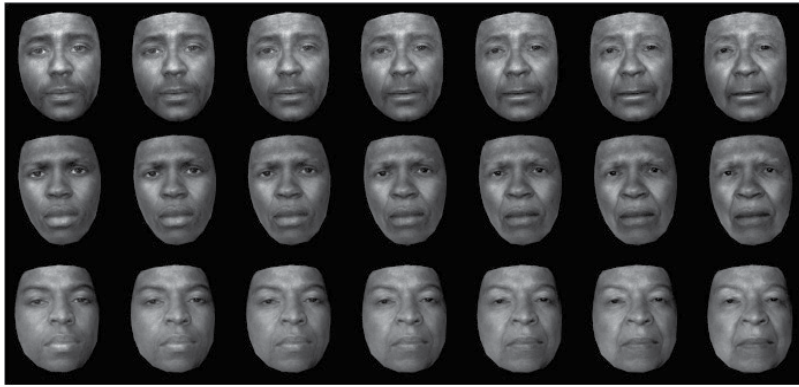


Fig. 8. Synthetic Facial Aging: Age progression technique

Figure 8 shows the technique for age-progression. For any input image to be age-progressed, the corresponding AAM parameters are generated from the AAM model. The age table is then loaded and the original parameters are shifted to the new age to which the image is to be progressed. The age-progressed image is then generated from the new set of shifted AAM parameters.

Some examples of age-progressed images using the technique described above is as shown in Figures 9 and 10. For comparison, the actual images at corresponding ages are included in Figure 10. These were generated using greyscale AAM models. Some recent images generated using a color AAM model are as shown in Figure 11. The AAM models and the synthetic aging technique is described in more detail in (Patterson et al., 2007a;b; 2009; Sethuram et al., 2009)



(a) Example 1: From Morph database



(b) Example 2: From FG-Net database

Fig. 9. Age progression using synthetic technique - Grayscale



Fig. 10. Age progression using synthetic technique - Grayscale

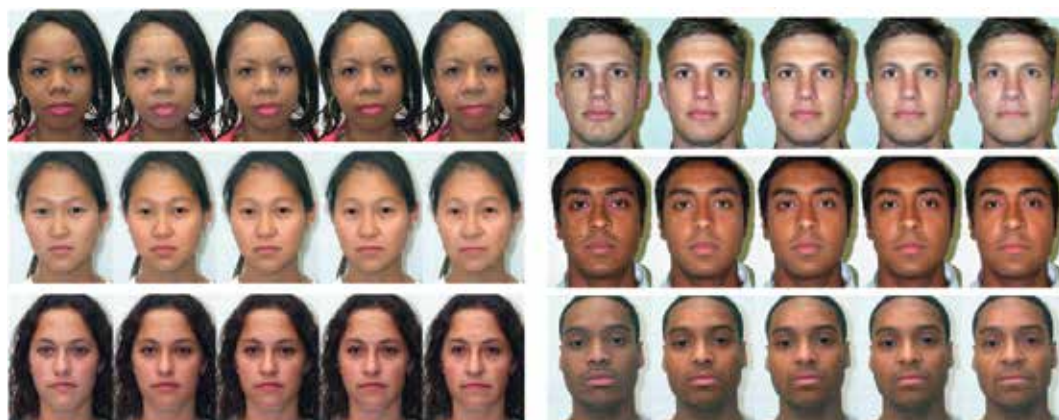


Fig. 11. Age progression using synthetic technique - Color

To improve the performance of the FR system, we propose that the gallery be augmented with the synthetic facial images generated by the above technique. Experiments are being conducted at IISIS to quantify the actual performance gain and is a work in progress. If any significant improvement is found, this work will be a major contribution to make the FR system robust to any aging effects.

7. New considerations: Age-related diurnal changes in facial appearance

Most recently, the work of face aging as it applies to FR systems has begun to consider diurnal changes in individual faces as an offshoot of more traditional diachronic face aging studies—studies that survey faces over many years of adult life. Diurnal changes are those alterations in a person’s face that occur over the course of a single day, from morning until evening. If using individual facial features or the entire facial appearance as a biometric tool for human identification, it is imperative to take into account the dynamic nature of human faces. Aside from considering facial expressions which change the look of a face, there are alterations in facial appearance which may be seen in print or digital photographs, as well as video that are due to pose and lighting—effects of image capture that alter how a person looks. These factors can modify facial texture and contour, create shadows or bright spots, and so on, further complicating computer automated face recognition (FR) capabilities. At IISIS, research is being expanded to study the natural changes in facial appearance in a single day, which can further enhance our understanding of diurnal changes in facial appearance.

7.1 Diurnal changes in face wrinkling

In a study by Tsukahara et al. (2004), image analysis of 38 Japanese females and males with a mean age of 34 years showed significant intensification of wrinkling in all areas of the face evaluated—forehead, corners of the eyes, and nasolabial grooves—in the afternoon when compared with the morning. Wrinkles were likely exacerbated later in the day due to a decrease in swelling of the face thought to occur in the morning since during sleep gravity is not pulling fluids downward; wrinkles may be plumped out in the morning. As the day progresses and a person is vertical, gravity may pull fluids downward from the face and into the appendages. As well, a person’s repeated facial movements or expressions may reinforce already existing lines, where the lines become intensified and deepen into wrinkles. How

noticeable these changes are across a single day is a critical question. If the intensification in wrinkling occurring from morning to late afternoon is discernable, the degree to which the change alters the appearance of an individual's face is meaningful. Preliminary results of an experiment conducted on a very small dataset of diurnal images did not show any significant changes in the match scores of the FR engine. Further research to quantify any diurnal changes in face wrinkling is thus warranted.

8. Conclusion

In this chapter, we discussed the many adult age-related craniofacial changes that occur in a human face. The performance of a commercial SDK was evaluated on longitudinal data. Synthetic age progression techniques were presented which could improve the performance of face recognition systems. Future research will be focussed on quantifying any performance gain in the FR system by augmenting the gallery with synthetic aged images of the face. Studies to quantify any diurnal changes that occur in a human face is another area of ongoing research.

9. References

- Akgul, A. & Toygar, T. (2002). Natural craniofacial changes in the third decade of life: A longitudinal study, *Am. J. Orthod. Dentofacial. Orthop* 122: 512–522.
- Albert, A., Ricanek, K. & Patterson, E. (2007). A review of the literature on the aging adult skull and face: Implications for forensic science research and applications., *For. Sci. Int* 172: 1–9.
- Behrents, R. (1985). Growth in the aging craniofacial skeleton, *center for human growth and development, University of Michigan, Ann Arbor, Michigan*. monograph 17.
- Bishara, S., Jakobsen, J., Hession, T. & Treder, J. (1998). Soft tissue profile changes from 5 to 45 years of age, *Am. J. Orthod. Dentofacial Orthop*. 114: 698–706.
- Bondevi, O. (1995). Growth changes in the cranial base and the face: a longitudinal cephalometric study of linear and angular changes in adult norwegians, *Eur. J. Orthod*. 17: 525–532.
- Coleman, S. & Grover, R. (2006). The anatomy of the aging face: Volume loss and changes in 3-dimensional topography, *Aesth. Surg. J*. 26: S4–S9.
- Doual, J., Ferri, J. & Laude, M. (1997). The influence of senescence on craniofacial and cervical morphology in humans, *Surg. Radiol. Anat*. 19: 175–183.
- FaceAgingGroup (URL).
URL: <http://www.faceaginggroup.com>
- FERET (URL).
URL: <http://face.nist.gov/frvt/feret/feret.htm/>
- FG-Net (URL).
URL: <http://www.fgnet.rsunit.com>
- Forsberg, C., Eliasson, S. & Westergren, H. (1991). Face height and tooth eruption in adults—A 20 year follow-up investigation, *Eur. J. Orthod*. 13: 249–254.
- FRVT (URL).
URL: <http://face.nist.gov/frvt/>
- Ibher, N., Kloepper, J., Penna, V., Bartholomae, J. & Stark, G. (2008). Changes in the upper lip—A photomorphometric and mri-based study (on a quest to find the right rejuvenation approach), *J. Plas. Recon. Aesth. Surg* 61: 1170–1176.

- Israel, H. (1977). The dichotomous pattern of craniofacial expansion during aging, *Am. J. Phys. Anthropol.* 47: 47–51.
- Isse, N. (2008). Silhouette sutures for treatment of facial aging: Facial rejuvenation, remodeling, and facial tissue support., *Clin. Plas. Surg* 35: 481–486.
- Kunjur, J., Sabesan, T. & Ilankovan, V. (2006). Anthropometric analysis of eyebrows and eyelids: An interracial study., *Brit. J. Oral Maxillofac. Surg* 44: 89–93.
- Lipotherapy (2009).
URL: www.queensparkclinic.co.uk/lipotherapy-fat-treatments-c56.html
- MBGC (URL).
URL: <http://face.nist.gov/mbgc/>
- O'Hare, P., Fleischer, A., D'Agostino, J., Feldman, S., Hinds, M. & Rasette, S. (1999). Tobacco smoking contributes little to facial wrinkling, *J. Eur. Acad. Dermatol. Venereol.* 12: 133–139.
- Overfield, T. (1995). Biological variation in health and illness: Race, age, and sex differences, in 2nd edition (ed.), *Biological Variation in Health and Illness: Race, Age, and Sex Differences*, pp. 71–72.
- Patterson, E., Sethuram, A., Albert, M. & Ricanek, K. (2007). Comparison of synthetic face aging to age progression by forensic sketch artist, *Proceedings of the seventh IASTED International Conference on Visualization, Imaging, and Image Processing* pp. 247–252.
- Patterson, E., Sethuram, A., M, A., Ricanek, K. & King, M. (2007). Aspects of age variation in facial morphology affecting biometrics, *1st IEEE Int. Conf. on Biometrics: Theory, Applications and Systems* pp. 1–6.
- Patterson, E., Sethuram, A., Ricanek, K. & Bingham, F. (2009). Improvements in active appearance model based synthetic age progression for adult aging, *3rd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems* pp. 104–108.
- Pecora, N., Baccetti, T. & McNamara, J. (2008). The aging craniofacial complex: A longitudinal cephalometric study from late adolescence to late adulthood, *Am. J. Orthod. Orthop* 134: 496–505.
- Peertrainer. (2009). What to do about wrinkly facial skin after weight loss.
URL: www.peertrainer.com/LoungeCommunity/Thread.aspx?Forum1..
- Pentland, A. & Choudhury, T. (2000). Face recognition for smart environments, *Computer* 33(2): 50–55.
- Phillips, P., Martin, A., Wilson, C. & Przybocki, M. (2000). An introduction to evaluating biometric systems, *Computer* 33(2): 56–63.
- Reece, E, M. & Rohrich, R. (2008). The aesthetic jaw line: Management of the aging jowl., *Aesth. Surg. J.* 28: 668–674.
- Ricanek, K., Boone, E. & Patterson, E. (2006). Craniofacial aging impacts on the eigenface biometric, *Proceedings of the sixth IASTED International Conference on Visualization, Imaging, and Image Processing* pp. 541–004.
- Ricanek, K. & Tesafaye, T. (2006). Morph: A longitudinal image database of normal adult age-progression, *7th Int. Conf. on Automatic and Face and Gesture Recognition* pp. 341–345.
URL: www.faceaginggroup.com
- Sadick, N., Karcher, C. & Palmisano, L. (2009). Cosmetic dermatology of the aging face., *Clin. Dermatol* 27(No.): S3–S12.
- Sarnas, K. & Solow, B. (1980). Early adult changes in the skeletal and soft tissue profile, *Eur. J. Orthod.* 2: 1–12.

- Sethuram, A., Patterson, E., Ricanek, K. & Rawls, A. (2009). Improvements and performance evaluation concerning synthetic age progression and face recognition affected by adult aging, *Proceedings of the 3rd IEEE International Conference on Biometrics* 5558(5): 62–71.
- Sforza, C., Grandi, G., Binelli, M., Dolci, C., De Menezes, M. & Ferrario, V. (2010). Age and sex-related changes in three-dimensional lip morphology., *For. Sci. Int.* 200: 182.e1–182.e7.
- Sforza, C., Grandi, G., Binelli, M., Tommasi, D., Rosati, R. & Ferrario, V. (2009). Age and sex-related changes in the normal human ear, *For. Sci. Int.* 187: 110.e1–110.e7.
- Shaw, R., Katzel, E., Koltz, P., Kahn, D., Giroto, J. & Langtein, H. (2010). Aging of the mandible and its aesthetic implications., *Plast. Reconstr. Surg.* 125: 332–342.
- Taister, M., Holliday, S. & Borman, H. (2000). Comments on facial aging in law enforcement investigation, *For. Sci. Comm.* 2.
- Taylor, K. (2001). Age progression: aging, *Forensic Art and Illustration.*, CRC Press, Boca Raton, pp. 251–281.
- Tsukahara, K. and Moriwaki, S., Hotta, M., Fujimura, T. & Kitahara, T. (2004). A study of diurnal variation in wrinkles on the human face, *Arch. Dermatol. Res.* 296.: 169–174.
- Victor, S. (2003). You under the sun, *Victoria* 17: 46–48.
- West, K. & McNamara, J. (1999). Changes in the craniofacial complex from adolescence to midadulthood: a cephalometric study, *Am. J. Orthod. Dentofacial Orthop.* 115: 521–532.
- Zimble, M., Kokosa, M. & Thomas, J. (2001). Anatomy and pathophysiology of facial aging, *Facial Plastic Surg. Clin. N. Am.* 9: 179–187.

Iris Recognition on Low Computational Power Mobile Devices

Huiqi Lu, Chris R. Chatwin and Rupert C.D. Young
*Industrial Informatics and Manufacturing System Research Centre,
 Richmond Building, School of Engineering,
 University of Sussex, Brighton, BN1 9QT
 United Kingdom*

1. Introduction

The growth in the use of smart mobile phones and PDAs, see Figure 1, for business transactions incorporating high performance cameras has created the possibility for developing camera and image processing applications for identity verification using handheld devices. The objective is to create a reliable, portable way of identifying and authenticating individuals. Another objective is to investigate the low power computing technology and establish a real-time biometrics software application. Improved, optimized and application specific image processing techniques are required to adapt to the currently available computing power on mobile phones.

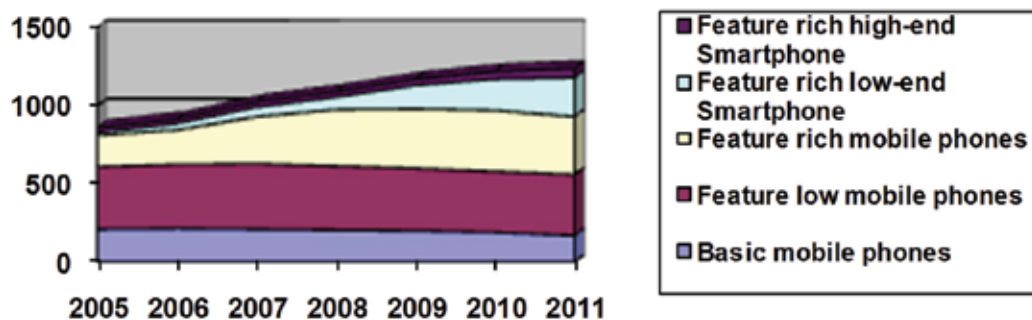


Fig. 1. Worldwide Mobile Phone Sales (in million), by technologies (Nokia Forum)

1.1 Biometrics technology and Iris recognition

Biometrics is the technique of using physiological or behavioural characteristics to identify or verify identities (Lu et al., 2007). The first evidence of using biometrics can be tracked back to Ancient Egypt, when a man called Khasehem used body measurements to distinguish people. His method prevented frauds when people were collecting food (Ashbourn, 2004). In 14th century China, children's palm prints were stamped on paper when they were born. These palm prints were stored and used by the Chinese Court for criminal identification purposes. In the 19th century a German, Franz Joseph Gall, founded

the discipline of phrenology, whereby specific cranial shapes and features were thought to align with certain mental characteristics. A couple of years later, an Italian physician Cesare Lombroso further developed the concept of phrenology with specific regard to criminal behaviour (Ashbourn, 2000). In 1897, Sir Richard Edward Henry of Scotland Yard developed a finger printing authentication system, later adopted by the British police (London Metropolitan Police, 2005).

Hundreds of years since the first traceable use of biometrics, the biometric technologies in use today include DNA profiling, fingerprinting, measurement of hand geometry, face recognition, iris recognition, voice verification and signature verification. These technologies are widely used in criminal investigations, applicant background checks, biometric visas and passports, border control, and preventing social services fraud.

Biometric technologies are popular because biometric characteristics are unique to the individual and cannot be shared, forgotten or stolen, as can keys, cards, passwords and usernames. Of course, a determined attacker will go to the great lengths to mimic the verification system by synthesizing a fake biometric input that is a close match to the authentic biometric signature (Vijaya Kumar et al., 2004). Applications demanding very high levels of security need a multilayered approach, in which passwords and multi-biometric technologies are employed.

Biometrics	Author	Error Handling	False Rate
Voice	Monrose	Discretization	20%
Signature	Monrose	Feature coding	28%
Fingerprint	Charles	Reed-Solomon code	30%
Face Recognition	Phillips	Discretization	43%
Iris Recognition	Daugman	Concentrated coding	0.47%

Table 1. Accuracy Comparison among Biometrics Technologies (Ma et al, 2004)

As shown in Table 1, of all the different biometrics technologies that are used for human authentication and are readily accessible for public users, iris recognition is the most accurate. This high reliability for personal identification is due to iris's randomly distributed features and its stability throughout life. According to Daugman (1993) the complex nature of the iris pattern has the highest variability in biometrics, 244 degrees-of-freedom.

The iris is an opaque sheet of tissue with a circular hole (the pupil) in it. It operates as an aperture stop in the optical pathway for the pupil. The iris lies just in front of the lens, a position that permits oblique rays to pass through the pupil to the peripheral retina, while limiting the size of the incident ray bundle (Clyde, 1999). The prenatal morphogenesis of the iris is completed during the eighth month of gestation. It consists of pectinate ligaments adhering into a tangled mesh causing striations, ciliary processes, crypts, rings, furrows, a corona, sometimes freckles, vasculature, and other features (Daugman, 1993). During the first year of a human life a layer of chromatophore cells often changes the colour of the iris, but the available clinical evidence indicates that the trabecular pattern itself is stable throughout the lifespan (Adler, F. H., 1965). At three years old, with the exception of people with eye diseases, for instance coloboma, ectopic pupils and subsidiary pupil, the iris is fully developed and is essentially stable over a person's life (Clyde, 1999).

Iris recognition is a mix of technologies from several fields, including computer vision, pattern recognition, statistical inference, and optics. The concept of automated iris

recognition was first proposed by Flom and Safir in 1987 (Flom, L., and Safir, A., 1987). Perhaps the best-known and most thoroughly tested algorithm is that of Daugman. Currently, the most widely used systems for iris recognition are employing John Daugman's algorithm (Daugman, 1993). Since the development of Daugman's algorithm, Wildes *et al.* (1996-1997), Boles and Boashash (1998), Lim *et al.* (2001), Sanchez-Reillo *et al.* (2001), Tissue *et al.* (2002), and Ma *et al.* (2004) have made significant contributions to the study of iris recognition.

There are several advantages in using iris recognition compared with other biometric technologies. These include the constancy of features over time and the method of measurement without contact measurement tools. Moreover, because it is the internal part of an eye, iris features cannot be easily changed by surgery.

Iris recognition provides high accuracy identification compared to other biometric technologies, which is why it is an attractive technology. As most mobile devices have at least one embedded camera, iris recognition is examined as a means of personnel identification using mobile computing devices.

1.2 Iris recognition solutions that are used on handheld Iris recognition devices

IRIS ID (formerly LG Electronics) and L-1 Identity are the two most well-known companies in the world for manufacturing iris identification machines. Both these iris recognition machines implement software solutions from a third party: Iridian Technology (<http://l1idsol.net>). The Iridain Technology is committed to providing software solutions for iris identification based on the iris recognition algorithm developed by Daugman. Iridian and Secure Metrics (<http://www.securimetrics.com>) are a part of the L-1 Identity Group. A review is given below of these companies and their biometrics related products:

2. IRIS ID, formerly LG Iris (<http://www.irisid.com/>)

LG Electronics Iris Technology Division (<http://www.lgiris.com/index.php>) of LG Electronics Incorporation was established in the United States in 2002, providing global management responsibility and overall direction of LG's iris recognition technology products. LG has released three main products up to now: 'LG IrisAccess 2200', 'LG IrisAccess 3000' and 'LG IrisAccess 4000'. The LG IrisAccess 2200 machine introduced in 1999 utilized conventional camera technology with advanced lens design and special optics to capture the intricate detail found in the iris. In 2001, LG IrisAccess 3000 was released, providing a platform incorporating more system security features, improved speed and enhanced user interfaces. LG IrisAccess 4000 is the third generation of iris recognition machine deployed. The latest generation offers more application versatility and integration flexibility. LG Electronics are using their product as a door entry system which can access significant computing power due to its static nature.

LG's Iris Technology Division later changed into IRIS ID Ltd. Their latest portable products released are iCAM TD100 (http://www.irisid.com/ps/products/icam_td100.htm) and iCam H100 (not released to the public until November 2010). Both are based on the Linux operating system and Daugman's algorithm 2Pi, multiband IR is used for illumination on both devices. Both machines require less than 2 seconds for complete two iris capture, and less than 8 seconds for a complete transaction, which includes face and iris (IRIS ID Systems Document #IRISIDTD100-0110).

3. L-1 Identity solutions (<http://www.l1id.com>)

L-1 Identity is a world-leading company providing multiple biometric identity solutions. They work closely with the US government for public safety enhancement. They have released two main products: *PIER*TM, a portable Iris enrolment and recognition Device as well as *HIIDE*TM Series handheld interagency identity detection equipment.

Their iris recognition solution is a major achievement for the SecuriMetrics's devices. Their iris recognition algorithm is also based on Professor John Daugman's algorithm, the same as IRIS ID. According to a SecuriMetrics's report ('Offender Identification System-Fast and Accurate Offender Identification', SecuriMetrics Incorporated, CA 94553, USA), their iris recognition technology was chosen by the US Department of Defense and other government agencies because 'Iris recognition is the fastest, most accurate and scalable of all biometric technologies'.

Due to the similarity of their ambition to provide an iris recognition solution for portable devices, more information about the SecuriMetrics's products is give below:

- *PIER*TM 2.3: A hand-held device that allows an operator to identify individuals using the highly unique patterns and textures of the human iris. By linking to a computer, the device can match an unknown individual against a database of millions.
- *HIIDE*TM Series 5: This is the first handheld recognition device in the world that combines iris recognition, fingerprint recognition and face recognition solutions. This latest release uses an Inter ® Atom TM Processor (1.6 GHz, 512KB cache, 533 MHz FSB) with 2GB DDR2 SDRAM. It supports 3G and 4G network.

4. Case study: A mobile biometrics system

4.1 System architecture

The main goal and the initial purpose of building this MBI system is to identify and/or verify individuals using a mobile phone. Hence, iris images are captured using a camera embedded in the mobile phone and are processed and identified locally. To deal with large-volumes of iris images that can be used by banks and mobile phone operators, an extended MBI system was developed. This extension allows the MBI system to run on benchmark workstations (server stations). Figure 2 demonstrates the components and functional design of the extended MBI system.

As seen from Figure 2, there are three kinds of medium used for storing template iris images. The first one is a SIM card. Using this medium, the mobile phone user's private iris codes are stored in a SIM card and the iris image identification can be carried out on the mobile phone. The benefit of this is that the iris information is retrievable locally and is reusable on different mobile phones. There is always a copy of the iris code stored in the SIM card and in the database of a mobile network operator. Users do not need to register their iris information again if they use another mobile phone with the old SIM card. The disadvantage of using a SIM card is the lack of privacy and security of personal information. The second kind of medium is the local memory of a mobile phone. This is similar to using the SIM card but in practice uses a slightly shorter time when executing the MBI system. By using this medium, users do not need to worry about private iris information being shared or stolen by network operators, however, iris information can still be lost or stolen if the phone is lost or stolen. Both the first and second mediums suit users using mobile phones for personal iris image recognition, e.g. as a login authentication.

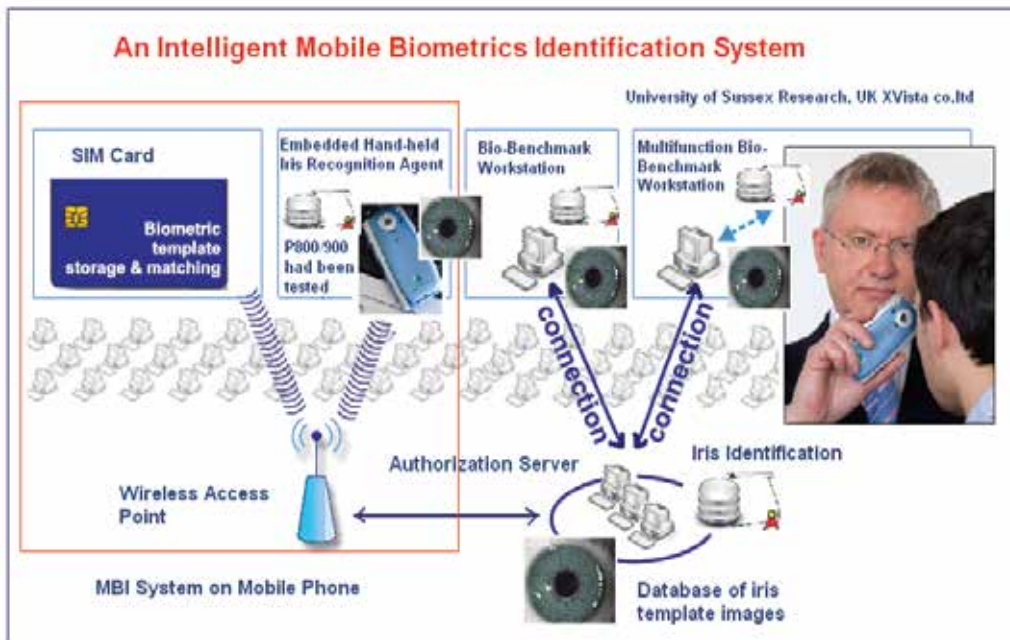


Fig. 2. Component and Functional Design of the MBI System

The third kind of medium is the database of a benchmark workstation. Comparatively, this method is specially designed to allow enterprise users access to a large-scale iris template database to authenticate iris images. However, this medium is currently not psychologically acceptable to some users.

As introduced in the abstract, the MBI system can collaborate seamlessly with other applications on mobile phones. In our work, we combined the MBI system with a mobile two-dimensional barcode reader as a biometrics mobile payment (BMP) system. The aim of the BMP system is to provide a secure means for authorizing users therefore preventing payment risks for both user (customer) and acquirer (financial institutions). The BMP system consists of two main parts: a two dimensional barcode reader and a MBI system.

Users firstly need to use the barcode reader to scan the barcodes of products or bills that they would like to purchase or pay. There are two means of getting information on products: locally from the database of the mobile phone or by logging into a server station. The local database in the mobile phone can provide personal preference information of the user, such as the list of favorite products and the account information of frequently scanned bills. This approach allows users to keep their iris information private. However it is not possible to save the entire history of what a user has been purchasing in a mobile phone database, due to the memory limitation on mobile phones. There is no such limitation when getting product information via the network; however, the network speed can be constrained, hence the purchase procedure may require more time.

Once the product information has been received, a user can decide whether or not to purchase the goods. If a user gives a confirmation for purchase, the MBI system will automatically start up and provide an authorization process to establish the user's identity.

There are two ways to get a user authorized. The first way is to recognize the user's iris

using the template images stored in a local database, for example on a mobile phone or PDA. By using this means, a user can send confirmation of a specific certificate of authorization that is generated on the mobile phone to a financial institution through the wireless network. Once a financial institution receives this certificate, they will decode the certificate and confirm the payment with both the merchant and the user; this confirms the user's identity and authorizes the payment. The second method is to send the iris code to the financial institution for transaction authorization; this method can help the financial institution reduce the risk of accepting fake certificates; however this method is currently not psychologically accepted by some users.

Once a payment confirmation is given to both a user (customer) and a merchant, the purchase by mobile phone is completed.

After explaining the working procedure and system design of the MBI and BMP system, the next section illustrates its agent oriented software architecture.

When designing the architecture of the MBI system, technologies such as distributed agent and multi-agent systems were employed. These technologies are the complementary approaches to the object paradigm and provide a means of designing and implementing complex distributed software (Garcia *et al.*, 2004). Agent technologies (Chen *et al.*, 2006) represent the most important new paradigm for software development (Luck, 2004), and have already been used in a diverse range of applications in manufacturing (Shen *et al.*, 1998; Wada *et al.*, 2002; Van Dyke Parunak *et al.*, 2001), process control system, electronic commerce (Sandholm, 2002), network management, transportation systems (Chen, 2002), information management (Stathis *et al.*, 2002), scientific computing (Boloni *et al.*, 2000), health care (Huang *et al.*, 1995) and entertainment (Noda and Stone 2003). The MBI system applies a combination of agent technologies, e.g. scientific computing, network management and information management.

By using these agent technologies, seven distributed agents are generated: image acquisition agent, data matrix decoding agent, iris image processing agent, iris image identification agent, wireless network agent, database management agent and communication agent. The organisational networks in the MBI system and the BMP system are shown in Figure 4 and Figure 5.

Image Acquiring Agent: This agent acquires images from live video inputs via cameras integrated into mobile phones to capture data matrix images or iris images.

Data Matrix Decoding Agent: With an image acquired using the image acquiring agent, the data matrix decoding agent will implement our algorithm to provide the decoded result of the data matrix.

Iris Image Processing Agent: With an eye image acquired using the image acquiring agent, the iris image processing agent can find the inner and outer boundaries and extract the iris information into an iris code.

Iris Image Identification Agent: The iris image identification agent compares an iris code generated by the iris image processing agent with template arrays stored in a local database using the Hamming Distance. If the similarity level between the user iris code and the best matched template code is higher than a threshold, the user iris code is considered to be correctly identified; otherwise a message 'no match' is displayed.

Database Management Agent: This agent is used to manage the local database of the MBI system. Its functionality includes register, update, check and remove user's personal iris code or other relevant information such as a person's name, address, banking details, etc.

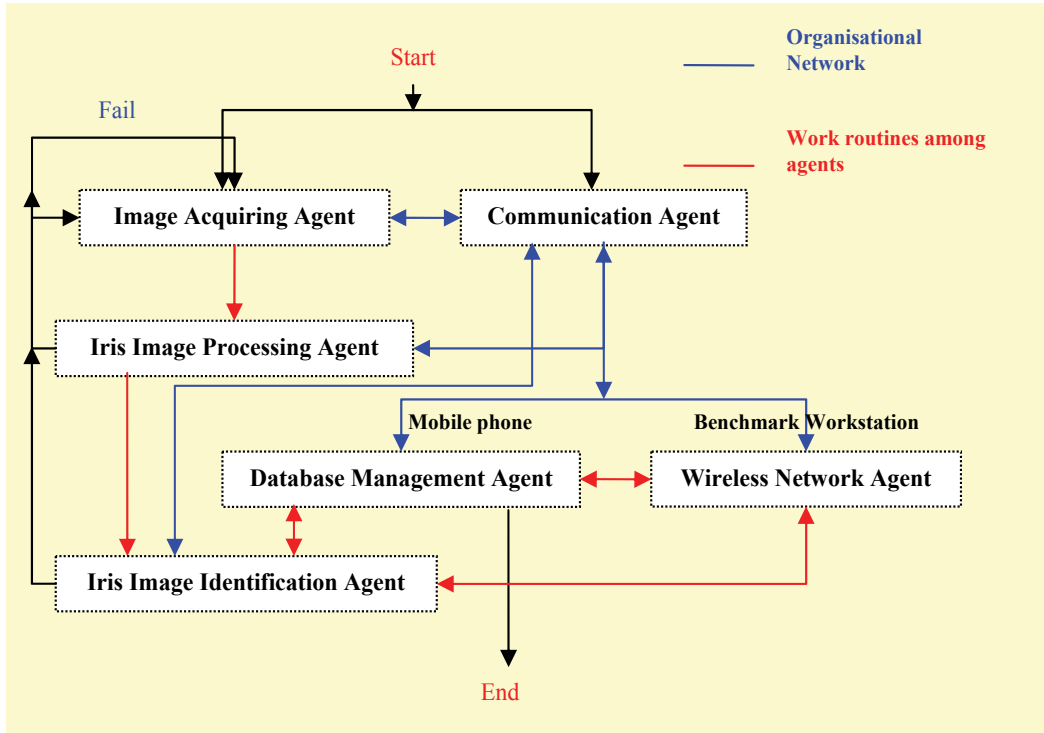


Fig. 4. Distributed Agents and Agents' Organisational Network in MBI System

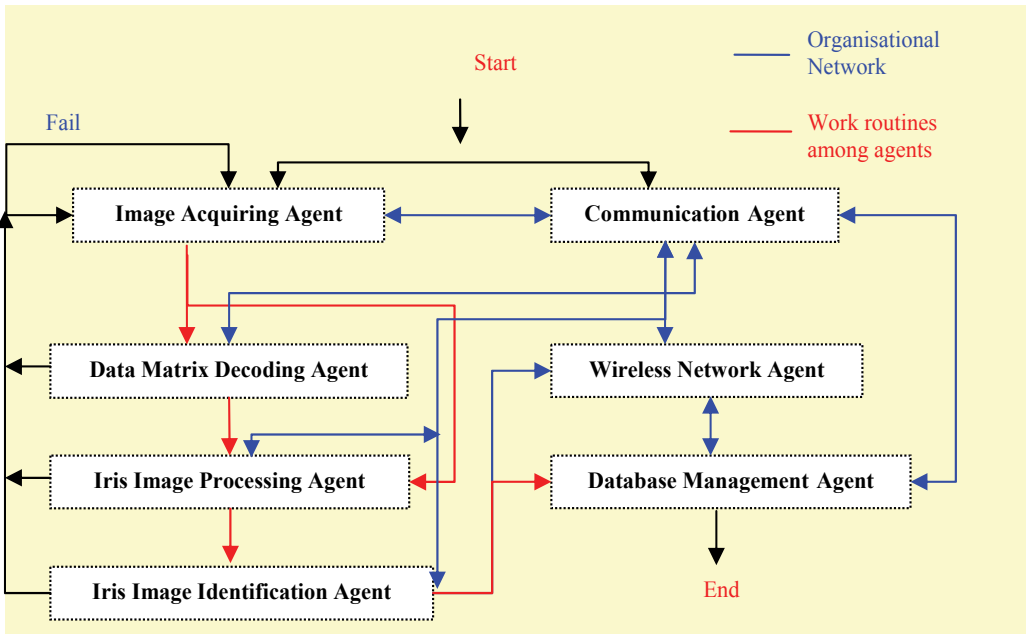


Fig. 5. Distributed Agents and Agents' Organisational Network in BMP System

Wireless Network Agent: This agent is used to transfer information between mobile phones and wireless network servers (also used for the benchmark workstation). Currently this agent can only enable the information transformation between mobile phones and network servers having fixed IP addresses due to security considerations. All internet information required by users needs to be downloaded by a server station and then filtered and redirected to the mobile users. This arrangement enhances the security of the MBI system.

Communication Agent: There is no core agent in the network. All communication such as requirements or negotiations should be sent via the communication agent. The communication agent creates links between the six other distributed agents and orchestrates all the actions required for both the data matrix decoding and iris image recognition.

4.2 Optics design

An assumption is made in developing the iris image processing that every eye image is taken from an environment with the same illumination and with the same distance between eye and camera. With this precondition, it is possible to set up parameters, such as boundary searching areas, and test the performance of the image processing algorithm. To achieve the same illumination, an 'eye cup' was produced with a luminous diode attached, as shown in the picture below:



Fig. 6. An 'Eye Cup' with a Luminous Diode links to the Sony Ericsson P800 Mobile Phone

This standardized illumination environment plays a crucial role in the design of the image processing algorithms. Several optical experiments were performed to test the 'eye cup' to ensure that the live video inputs taken with this 'eye cup' always focuses on the iris area.

4.3 Iris recognition algorithm

The procedure for iris image processing, designed to be run on mobile phones, follows the steps: loading an input image of eye, iris image localization and extraction of the iris characteristics into an iris code (patent no: GB0703399.6-PCT/GB2007/000591). This procedure is shown in Figure 7.

For the same illumination the same eye will give the same pupil response, hence, the relationship between the radius of the pupil and the radius of the outer iris boundary will not change. The variability due to the illumination response of the pupil is removed during iris localization, by using the 'eye cup'.

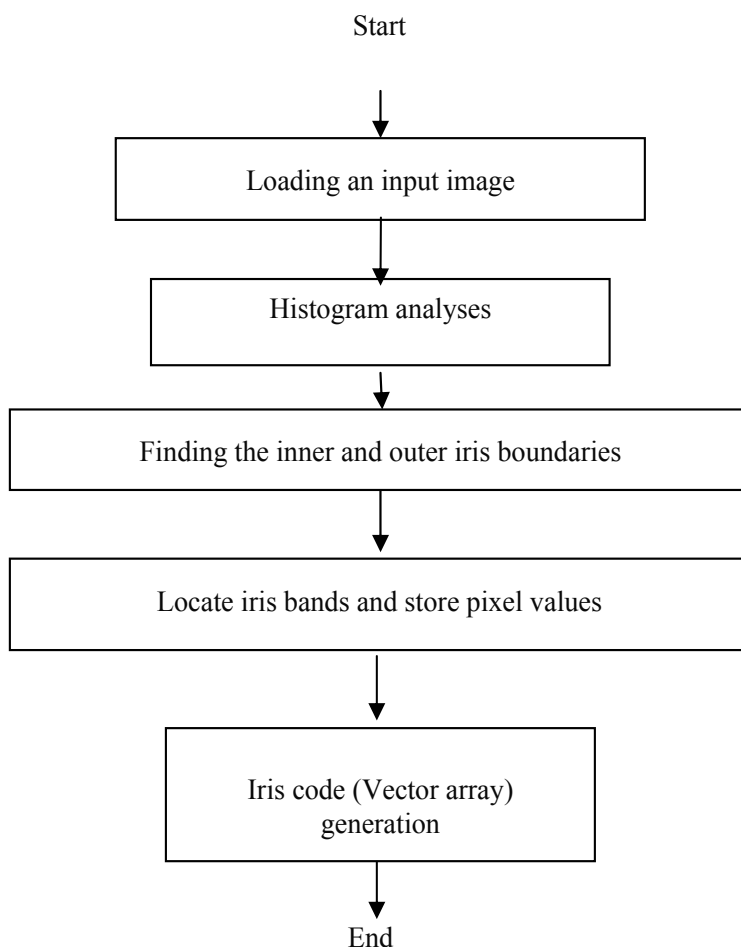


Fig. 7. The Procedures of Iris Image Processing

There are two sub-steps during the iris image localizations. Firstly, histogram analysis is applied for finding the average pixel values of the pupil area, iris area and sclera area from the input image, as well as the minimum pixel value and maximum pixel value of the iris area. Secondly, the inner and outer iris boundaries are found using the pixel-oriented method with the association of the different average pixel values and maximum/minimum values that are found from the histogram analysis. Figure 8 shows the calculated pupil boundary.

In these steps, failure is allowed in each step. Once a failure occurs, the iris image processing agent will communicate with the image acquiring agent via a communication agent to acquire a new image. There is a setting for the 'maximum number of images' in the MBI system, the communication agent monitors if the number of failures reaches this pre-set threshold. If this threshold is reached, a 'fail' notice for iris image decoding will appear on the main panel GUI of the MBI system. For more details of the iris recognition algorithms, please refer to (Lu, 2008).

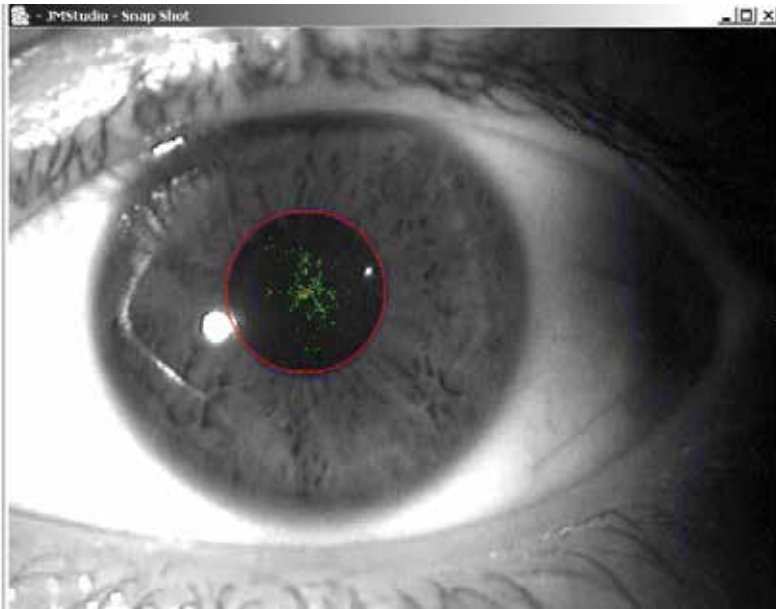


Fig. 8. Pupil Centre and Pupil Boundary

The area between the inner and outer iris boundaries is the iris area. There are also two sub-steps in generating an iris code. First, a number of (in this algorithm the number is eight) iris bands will be located on the iris area. Second, a radial averaging filter is applied to transform the band-data into a one dimensional signal, as shown in Figure 9. For the cases of eight iris bands located on the iris area, eight one dimensional signals will be generated after filtering.

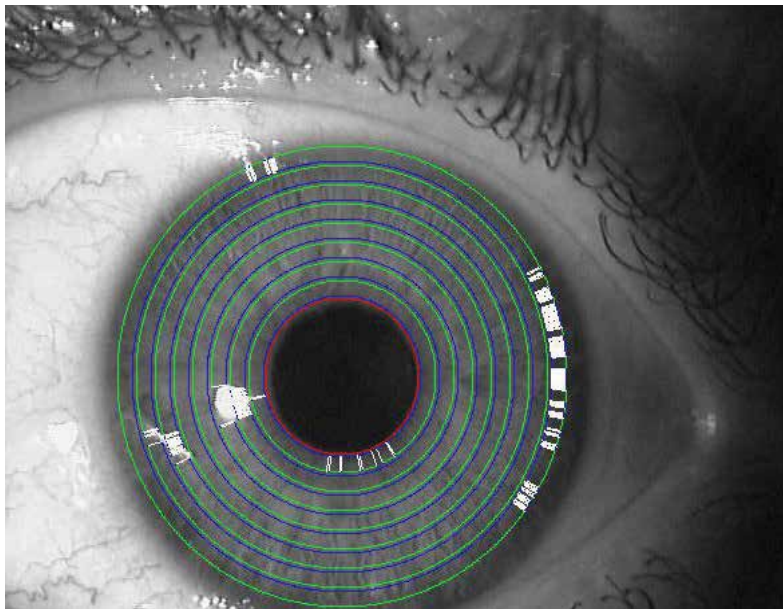


Fig. 9. An Example of the 8-Circle Iris Bands

5. Implementation

Based on the notion that many of the latest mobile phones are running on the Symbian operating system and using Java or the C++ programming language for programming developments, the biometric mobile payment system is developed for mobile phones running Symbian OS using *Java 2 Micro Edition* (J2ME) with *Mobile Information Device Profile* (MIDP). By using J2ME, the iris identification program is compatible with benchmark workstations with the benefits of the platform independency of Java codes.

All programming developments indicated in this paper have had been tested on a Sony Ericsson P800 mobile phone. It was the most advanced Symbian OS mobile phone at the time this project was started, supporting personal Java and MIDP 2.0 and equipped with a VGA sized camera. The Sony Ericsson P800 is embedded with an ARM 9 processor running at 156MHz. This processor power must be shared among common mobile phone functions. In the following section, the design of the GUIs and their functions are described in detail. Moreover, how distributed agents are launched via the GUIs and how these agents communicate with each other is also outlined.

The different GUIs are described in a sequence following the working procedures of the BMP system. There are two main graphical user interfaces (GUIs) and six preference menu GUIs that were developed using Symbian UIQ. After starting the BMP system, the main panel of the mobile two-dimensional barcode reader is the first interface to be displayed, as shown in Figure 10 and Figure 11.

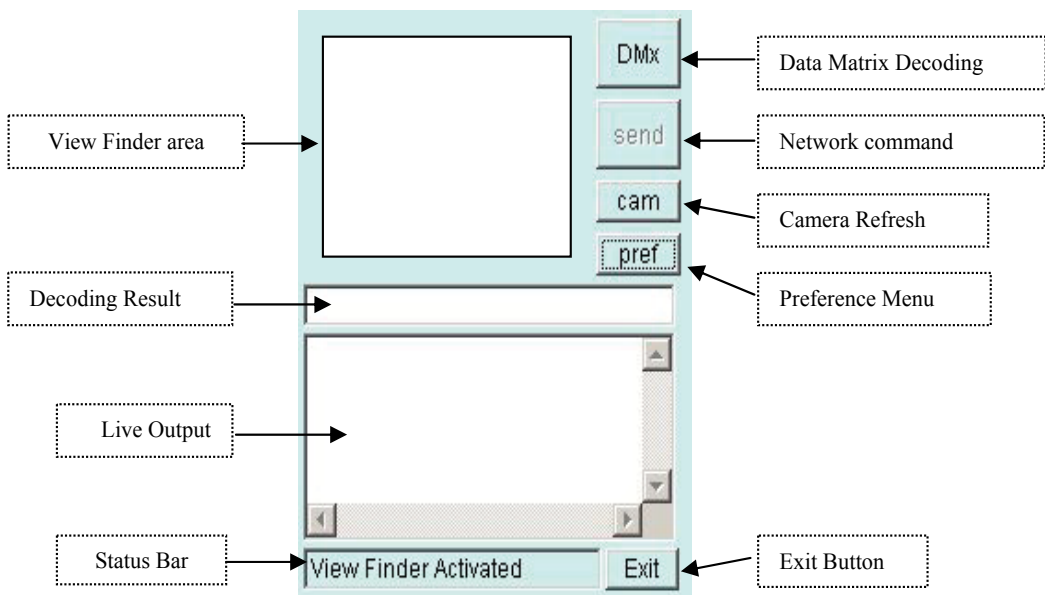


Fig. 10. Main panel of Mobile Two-Dimensional Barcode Reader

This main panel consists of three display windows and five command buttons. To use this barcode reader, users must position the mobile phone at a distance from the symbol so that the image formed on the mobile phone's *View finder area* is in focus and approximately centred and straight. The View finder area window is managed by the image acquiring

agent and the barcode image processing agent for displaying a data matrix image taken by the camera on the mobile phone. After pressing the *DMx* button, the image acquiring agent will start the image acquisition and send an image to the data matrix decoding agent. The *view finder area* will be switched off during processing to optimize the available CPU power. When the code has been successfully decoded, information related to the barcode, such as price, amounts of product left, will be displayed in real-time in the *Live Output Window*. By using the *check* button (as shown in Figure 14(a)), users can access the local database to get information that relates to the barcode last scanned. The *send* button is used to send the contents in the decoding result window to the server. This button is designed for the mobile commerce purposes and not to be used without a connection to a wireless network. According to the tests on the Sony Ericsson P800 mobile phone, an average decoding time, using a standard VGA size image, is 60ms to 100ms.

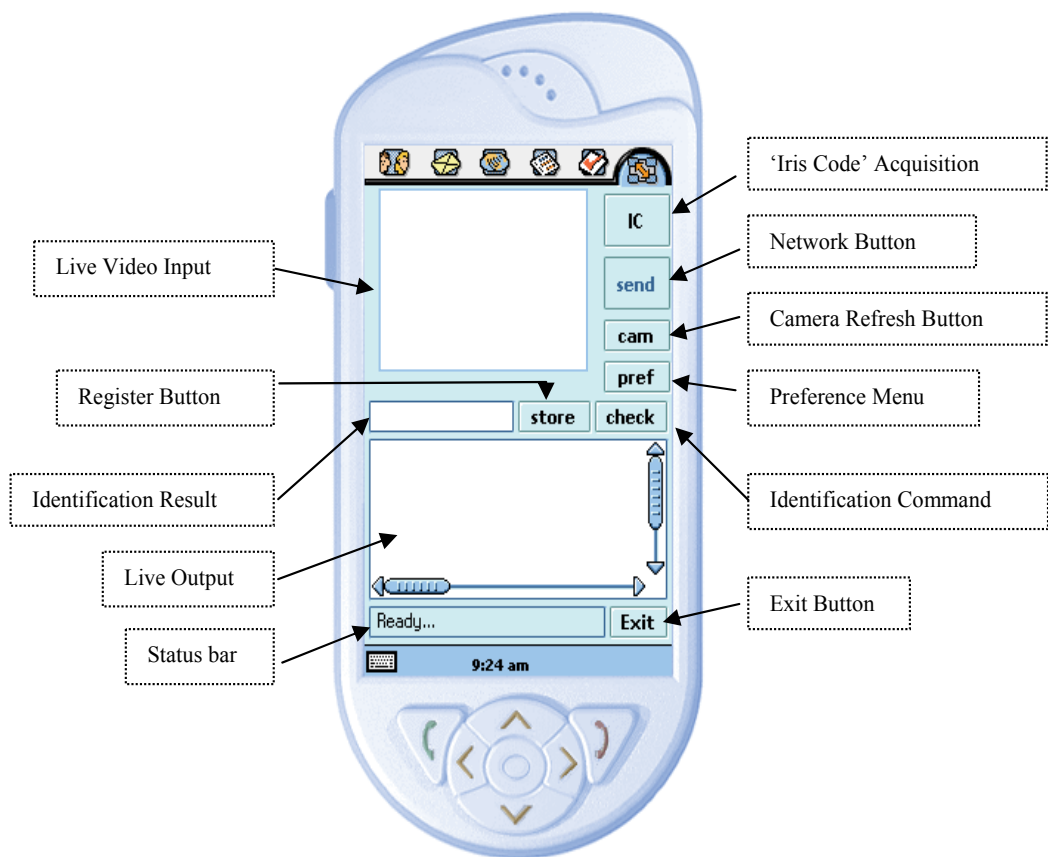


Fig. 11. Main Panel for the Mobile Biometrics Identification System

After users confirm purchase, a GUI for mobile biometrics identification will appear, as shown in Figure 10. There are four display windows and seven command buttons in this Panel. The *Live Video Input Window* is monitored by the communication agent and managed by the image acquiring agent and the iris image processing agent when displaying the live video input from the camera or displaying an iris image with its iris boundaries located (ie,

when the iris image processing succeeds). The *Identification Result Window* is managed by the iris image identification agent and used to display an identification result in a single line. It is also used to name a new iris code with a registration procedure. The *Live Output Window* works as a real time messenger interface to display live messages received from each agent. The messages include the status of image processing progress (success or failure), the full list of iris comparison results and the iris code generation time, etc. The *IC Button*, short for 'Iris Code', is used to activate the image processing agent to achieve a valid iris code. After the IC button is pressed, the live video input window will be temporally closed to save computational power. It will be reopened after the iris image processing agent has produced an output, which is either a valid iris code or a notice of failure. When an iris code is successfully generated, the pupil boundary and the iris bands are shown on the iris image. Therefore users can manually decide whether the iris image processing agent produced a correct pupil boundary or not. If a user does not agree with the localization of a pupil boundary, s/he needs to press the '*cam button*' on the main panel and press the IC button again to restart the image processing procedures. Alternatively, if a user agrees with the localization, the iris code can be compared with iris templates in a database for iris identification (by pressing the '*Check button*') or be registered to a database as a template (by pressing the '*Store button*'). Different from the two dimensional barcode reader, the *Send button* in the MBI system is a network button used to activate the wireless network agent and send the iris code generated most recently to a server station for registration or removal purposes (the IP address and the Port number of the benchmark workstation are set in the *Network Preference Menu*). After pressing the *check button*, the last calculated iris code will be compared with the template codes stored in the database of a benchmark workstation or the mobile phone. This button is only activated after a 'success' notice is received from the image processing agent and the communication agent. The identification result will be displayed in the *Identification Result window*. The matching rate that is the total number of pixels matching the iris code's template pixels will be displayed on the Live Output window. The *store button* is used to store the iris code that successfully decoded most recently into a database, either located in the benchmark workstation or in the mobile phone. In order to register an iris code, a name must be typed in the identification result window. This name needs to be less than 20 characters and have a different name from other data that has been stored in the database. Any preceding or trailing spaces in this name will be removed automatically by the database management agent during registration. Associated with the panels of the two dimensional barcode reader and MBI system, there are in total six preference menus which can be set by users. These menus can be accessed by pressing the *pref button* on the main panel.

As shown in Figure 12(a), this menu provides the facilities for setting the image colour, image size and the maximum number of images that can be acquired during a biometrics identification process and the minimum matching rate (details are given in the latter part of this section). The image setting menu for barcode reading, as shown in Figure 12 (b) enables contrast and luminosity tuning which is required for barcode reading different coloured data matrix.

Figure 13 (a) illustrates the GUI for the database options. This GUI can help a user review or remove information, such as iris information and product information, on the local database on a mobile phone. Figure 13 (b) illustrates the GUI of the log monitor, as indicated by its name, this GUI is designed to help users monitor all the attempts and the history of attempts for accessing information from the BMP system on a mobile phone.

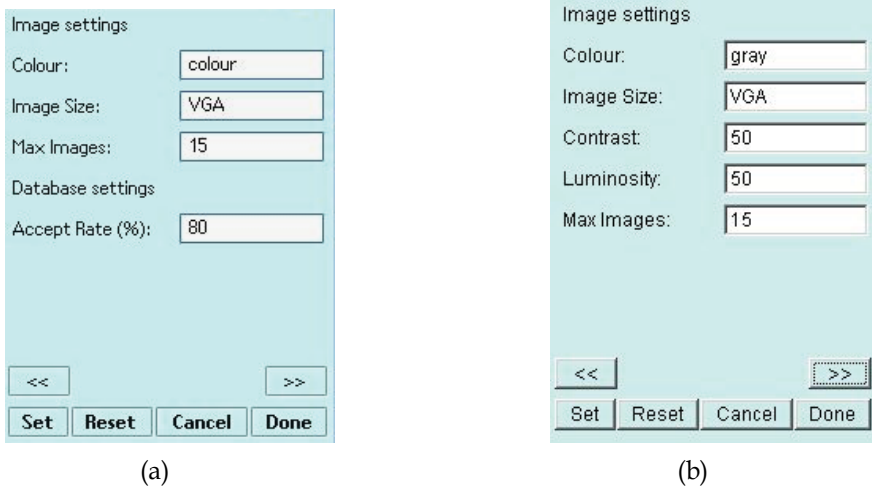


Fig. 12. Image Setting Menus of BMP System: (a) image setting and accept rate setting menu for biometrics identification; (b) image setting menu for barcode reading.

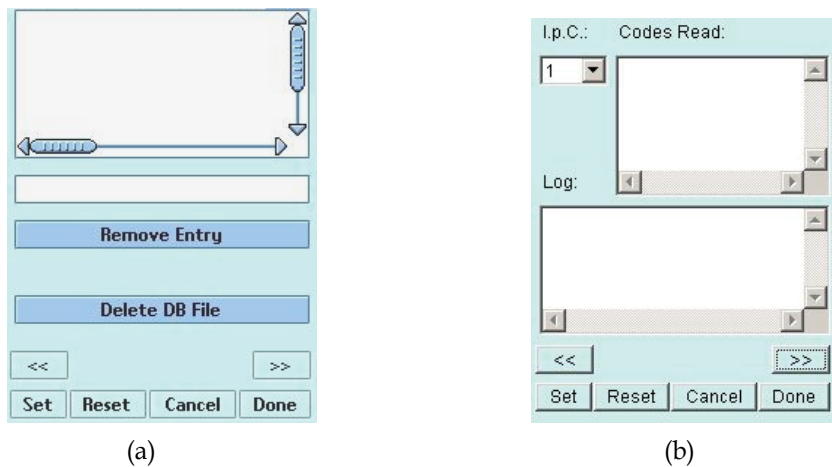


Fig. 13. Graphical User Interfaces: (a) Database Options (b) Log Monitor

The GUIs shown in Figure 14 (a) and (b) illustrate the network menu and the tool options. These two GUIs help mobile phone users transmit product information and iris codes between server stations and mobile phones.

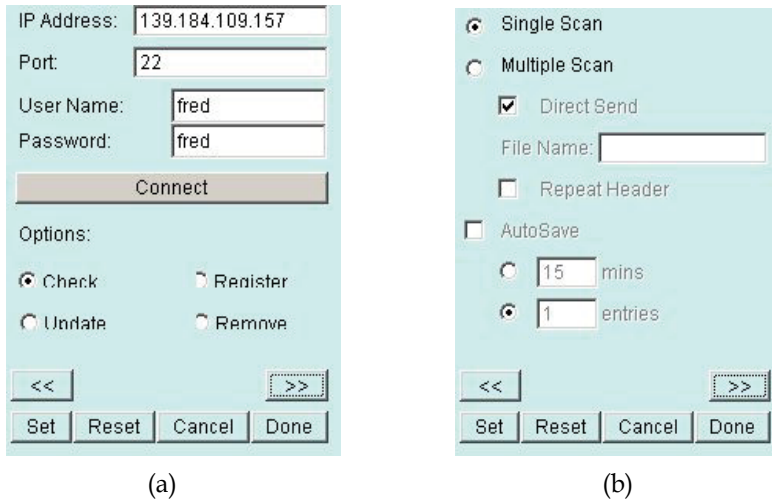


Fig. 14. Graphical User Interfaces: (a) Network Menu; (b) Tool Options.

6. System performance test

In order to provide a statistical test result on the accuracy of the iris recognition algorithm that we developed to run on a mobile phone, 25,200 image-correlation tests (12,000 auto correlation tests on images of authentic and 13,200 correlation tests on images of imposters) were carried out using the iris images captured by the Chinese Academy of Science Institute of Automation (CASIA). The test results are used to compare with results from John Daugman, who developed an internationally used and tested iris recognition algorithm. In Daugman's (1993) theory, if authentic and imposters can be successfully separated using a Hamming distance (HD) threshold during statistical tests this recognition algorithm can recognize iris patterns reliably.

The HD between two Boolean codes is defined as the number of differences between those two codes when compared bit by bit, defined as follows:

$$HD = \frac{1}{n_{pixel}} \sum_{j=1}^N C_A(j) \oplus C_B(j) \quad (1)$$

where C_A and C_B are the coefficients of two iris images and n_{pixel} is the size of the iris code. It is a Boolean Exclusive-OR (XOR) operation. Therefore the matching rate which is an XNOR operation can be written as:

$$\text{Matching Rate} = 1 - HD \quad (2)$$

When two iris codes are matched, the greater the value of the HD, the bigger the difference is between the two. A result of 0 would represent a perfect match between these two iris codes. To reject all imposters, the decision threshold of Daugman's algorithm is 38%, which is equal to a 62% matching rate.

During the statistical independence tests of our algorithm, the mean value for authentic' HD results is 0.1743, the standard deviation is 0.0797, the minimum value is 0 and the maximum

is 0.4200. When it comes to impostor's HD results, the mean value is 0.3748, the standard deviation is 0.0537, the minimum value is 0.2000 and the maximum value is 0.5400.

The distribution of impostor's and the distribution of authentic's have an overlap area between 0.2000 (the minimum value of impostor's HD) to 0.4200 (the maximum value of authentic's HD). However, by setting a threshold of 0.2 on Hamming distance all impostors can be recognised and by setting a threshold of 0.42 on Hamming distance, all authentic's can be recognised. The results indicate a success in separating authentic's and impostors by setting a threshold between 0.2000 and 0.4200 (matching rate between 58% to 80%) for the iris recognition algorithm of the MBI system.

According to 25,200 test results, several conclusions can be given on the accuracy and stability of the MBI system for different security levels that applications require. First, the false negative rates are found to be between 0.13% to 4.3% with no change in sensitivity resulting from changes of threshold. This result shows that the MBI system mostly does not reject authentic users hence providing good performance with a low inconvenience level for authentic users. Second, the false positive rates are found to be between 0% to 8%. According to table 2, from about 65% threshold, 0.13% false negative rate is achieved which leads to the rejection of almost all the impostors.

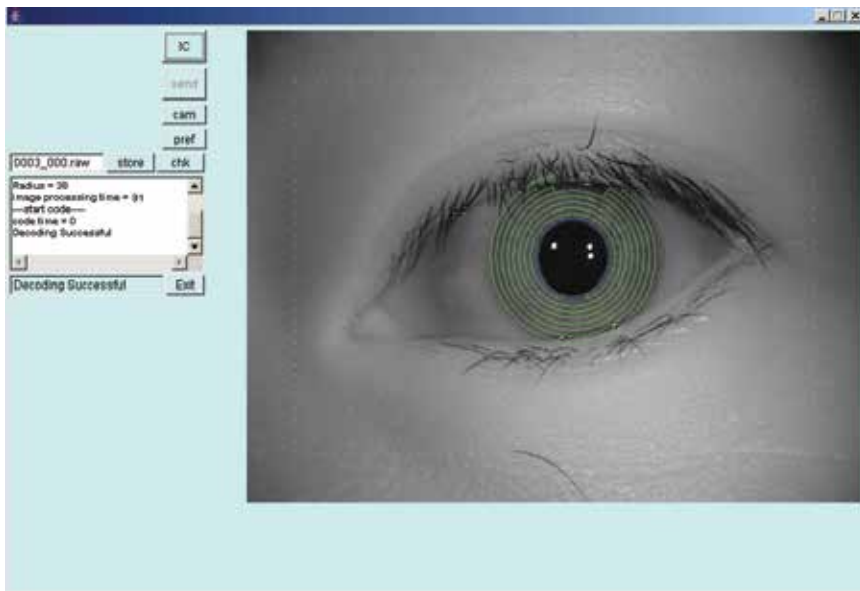
The false negative rate is also used in the paper of Wildes (1996), Lim et al (2001), Tissue et al (2002), and Daugman (2007) for measuring the security of a recognition system. Among them, Daugman's algorithm provides 0 in 984 million FAR in cross comparison tests. Comparatively to Daugman's result, the MBI system has a 0.13% FAR rate when setting a threshold at 65%.

When it comes to true rates (true positive and true negative), the true negative rate increases greatly from a threshold of 50% depending on the threshold increment and reaches the value of 100% for the true negative rate for a threshold higher than 65%; The true positive rate decreases from a threshold of 60% depending upon the threshold increment. For both true rates, the closer to 100%, the better performance the system will provide.

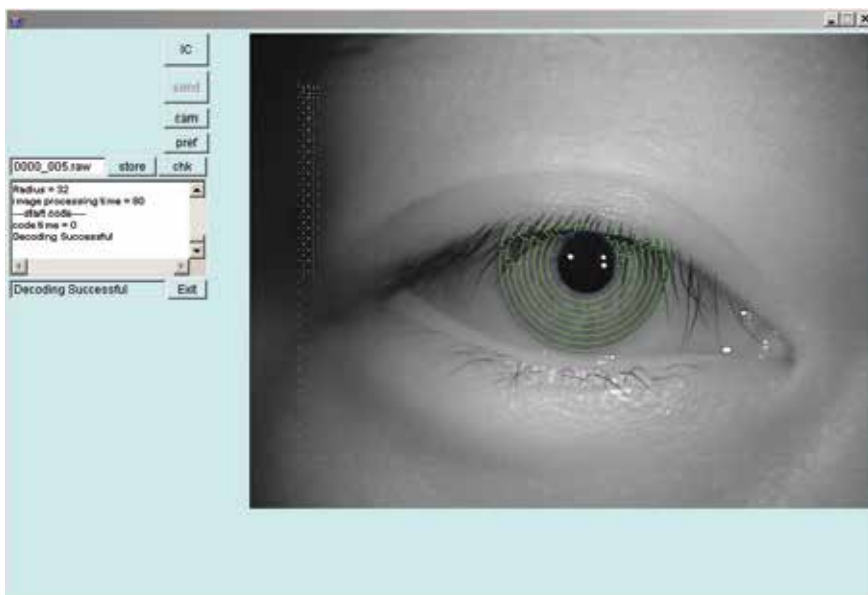
For demonstration and analysis purposes, three groups of false and true rates were generated when using three different threshold settings: 60%, 70% and 80% (these thresholds roughly indicate the min, medium and max value in the range of [58%, 80%]) are used. As shown in Table 2, these results can be used to analyse the security robustness of the MBI system.

Matching Rate	False Positive Rate	True Negative Rate	True Positive Rate	False Negative Rate
60%	8.3328882%	91.2605042%	97.12973594%	0.133539875%
70%	0.000000000%	100.000000000%	50.63145809%	2.29688585%
80%	0.000000000%	100.000000000%	7.347876005%	4.310667165

Table 2. True and False Rates in the Statistical Independence Test



(a)



(b)

Fig. 15. Performance of Iris Identifications on the MBI System: (a) and (b): identification results on the main panel GUIs of the MBI system.

By analysing the correlation test results, the authentic-group and the impostor-group can be successfully separated with a false acceptance rate of almost zero, by setting a decision threshold at 65% of matching pixels. Figure 15 provides two sample results from these tests to illustrate that the iris recognition algorithm itself produces reliable identification results.

As shown in Figure 15, the Live Output Window on the main panel GUI of the MBI system works as a real time messenger interface to display live messages to show the status of image processing progress (success or failure), the iris comparison results and the iris code generation time, etc; and the Identification Result Window is used to display the identification result. According to this feedback, the image processing time when implementing the iris image processing agent is 80 ms in (a) and the 91 ms for the image in (b). A note of '---start code---' is given when a successful result is achieved during iris image processing. Both 'code time' (iris code comparison time) shown in the Live Output Window are 0 ms. A message of 'Decoding Successful' is given when the image of an eye is identified. An identification result (the name of the template that the latest acquired image matched with) of '0000_005' is shown in the Identification Result Window in (a) and an identification result of '0003_000' is shown in (b). Moreover, these two typical examples demonstrate how the MBI system works based on different degrees of eye openness. As shown in Figure 15(a), one third of the iris image in (a) is covered by the eyelid and a small amount of noise comes from the eyelash. In Figure 15(b) a smaller amount of noise comes from the eyelid and the eyelash. All the noise in (a) and (b) has been removed when implementing the iris image processing agent (in other words, before implementing the iris image identification agent).

After demonstrating the performance of the BMP system (the two dimensional barcode reader and the MBI system), the execute time of BMP system is another issue that needs to be explored.

The procedure of barcode decoding and barcode information searching can be achieved within 100 ms if the product information has already been stored in the local database in the mobile phone. If product information is not stored in the mobile phone memory and an internet search is required, the execution time of the BMP system then depends on the data transmission speed between the server station and the mobile phone.

Once product information is received from server or found from the local database and the user decides to purchase, a 10ms pause will be given when switching user interfaces. This pause allows the communication agent to communicate with the image acquiring agent and iris image processing agent to activate the user identification process. After this pause, the biometric identification GUI, as shown in Figure 10, will appear on the screen. The 'Live Output Window' in this GUI will keep providing updated user information and information relating to a purchase, such as payment confirmation from a financial institution.

The execution time for local iris recognition is within 200 ms when compared to 100 iris code templates in the local database. If the user prefers to be authorized by a bank or other financial institution, code transformation times may vary but the iris code identification time will be less than 50ms, as shown in Table 3.

Based on the execution time of the mobile biometric identification and mobile two dimensional barcode reader, this BMP system can decode a barcode and produce the information related to this barcode and get a payment authorised within one second if there is no stand by time for data transfer via a network connection.

No. of Iris Templates	Database (ms=milliseconds)			
	Local		Server	
	IC generation	Identification	IC Generation	Identification
100	200 ms	-	<100 ms	50 ms
1000	500 ms	-	<100 ms	500 ms

Table 3. The Execution Time of the MBI System

7. Conclusions and future works

The growth in the creation and maintenance of secure identities for mobile devices has created challenges for individuals, society and businesses particularly in mobile added value services (mobile banking, mobile check-in, mobile ticket, et. al) and government security services. Although many obstacles remain, the growth in wireless technology, and the improvement of mobile devices will stimulate growth in the mobile biometrics market. The new product enhancements will result in an expansion of m-Identification and m-Commerce applications worldwide.

When it comes to enhancing the work being done in the MBI case study, there are two prospective areas that would benefit from further work.

The first aspect is the improvement of the performance of the MBI system. A test has been carried out on using three iris images at a time to produce lower false positive and false negative rates when applying the MBI system for iris recognition. According to the test, the false rates are about one tenth of those using one image with the same iris recognition algorithm that was implemented in the MBI system. However, using three images at a time for iris recognition requires higher computational power and a longer standby by time. The future work of this aspect addresses both the code modification of MBI system and the improvement of technologies applied on portable computing devices.

Secondly, an improvement can be made to the BMP system by employing a mobile shopping agent. In this blueprint, the shopping agent can be developed to enhance the intelligence of the BMP system to suit users' preferences. With such a shopping agent users scanning the barcodes of interesting products can be provided with multiple options on where best to buy or how to pay for the item.

The main difficulty of employing such a mobile shopping agent is not the algorithm development but the cooperation with the mobile network provider. In the design of a wireless network agent for security, mobile phone users need to connect to a network server with a fixed IP address. All the information required from the internet needs transformation and filtering through a network server. Therefore the functionality of the BMP system is unfortunately partly constrained by the security concerns. Many advanced mobile phones can explore the internet without using the BMP system, but the security level of the mobile phone is not yet ready for mobile shopping. There are some world leading anti-virus developers, such as McAfee who are currently committing on the software solutions for

anti-virus on mobile phones; however such software requires computation power and memory, which can make a mobile phone 'pause forever' when exploring the internet. In a world challenged to find new ways to authenticate identity and privileges when processing people and information, all with increased levels of security, the future of iris recognition technology on portable computing devices looks bright.

8. Acknowledgement

This is a commercial project delivered by the *iims* Research Center, University of Sussex for, xVista Ltd, United Kingdom. The authors are grateful to Chinese Academy of Science Institute of Automation (CASIA) for use of CASIA Iris Image Database, on which some results are based.

9. References

- Adler, F. H., *Physiology of the Eye: Clinical Application*. The C.V. Mosby Company, 4th edition, 1965.
- Ashbourn, Julian., *Biometrics: Advanced Identity Verification:the Complete Guide*, Longon: Springer-Verlag, 2000.
- Ashbourn, Julian., *Practical Biometrics: from Aspiration to Implementation*, London: Springer-Verlag, 2004.
- Boles, W., and Boashash, B. "A human identification technique using images of the iris and wavelet transform". *IEEE Trans. Signal Processing* 46: (1998): 1185-1188.
- Boloni, L., Marinescu, D.C., Rice, J.R., Tsompanopoulou, P. and Vavalis, E.A. "Agent based scientific simulation and modelling." *Concurrency Practice and Experience*. Vol.12(9), p. 845-861.(2000)
- Chen, W.S.E. and Hu, C.L. (2002) "A mobile agent-based active network architecture for intelligent network control." *Information Sciences*. Vol. 141.
- Chen, B., Cheng, H.H. and Palen, J. (2006) "Moble-C: a Mobile Agent Platform for Mobile C/C++ Agents." *Journal of Software-Practice and Experience*. Vol. 36, p. 1711-1733.
- Clyde, W., Oyster, *The Human Eye: Structure and function*. Sunderland, U.S.A.: Sinauer Publisher, 1999.
- Daugman, J., "Probing the Uniqueness and Randomness of Iris Codes: Results from 200 billion iris pair comparisons" *Proc. IEEE* 9411(2007): 1927-1935.
- Flom, L., and Safir, A., "Iris Recognition System." U.S. Patent 4 641 394, 1987.
- Kiniry, J., Zimmerman D. "A hands-on look at Java mobile agents". *IEEE Computer Computing* (July-August 1997): 21-30.
- Garcia, A.F., de Lucena, C.J.P. and Cowan, D.D. (2004) "Agents in Object-Oriented Software Engineering." *Journal of Software-Practice and Experience*. Vol. 34, p. 489-521.
- Huang, J., Jennings, N.R. and Fox, J. (1995) "Agent-based approach to health care management." *Applied Artificial Intelligence*. Vol. 9(4), p. 401-420.
- Lim, S., Kee, K., Byeon, O., and Kim, T. "Efficient iris recognition through improvement of feature vector and classifier," *ETRI J.* 23.2 (2001): 1-70.

- London Metropolitan Police, United Kingdom.
<http://www.met.police.uk/so/100years/henry.htm>. Accessed on 15 June 2005.
- Lu, H., Claret-Tournier, F., Chatwin, C., and Young, R., "M-Commerce Secured Using Web-enabled Mobile Biometric Agents." *Proc. 2nd IEEE/ WIC/ACM International Conf. on Web Intelligence and Intelligent Agent Technology*, Silicon Valley, USA (2007).
- Lu, H., 'Recognition Algorithms for Biometrics on Portable Computing Devices', D. Phil Thesis, University of Sussex, UK, 2008.
- Luck, M., "Challenges for agent-based computing." *Autonomous Agents and Multi-Agent Systems*. Vol. 9(3), p. 199-201, (2004).
- Ma, L., Tan, T., Wang, Y., and Zhang, D., "Efficient iris recognition by characterizing key local variations". *IEEE Trans. Image Processing* 13.6 (2004).
- Noda I. and Stone P. (2003) "The RoboCup Soccer Server and CMUnited clients: Implemented infrastructure for MAS research." *Conf. of Autonomous Agents and Multi-Agent Systems*. Vol. 7(1-2), p101-120.
- Nokia Forum, "Delivering Native Symbian OS Applications Over the Air-New Opportunities to Drive ARPU." version 1.0,
http://www.s60.com/pics/pdf/symbian_whitepaper.pdf
- Paul May (2001). *Mobile Commerce*, Cambridge University Press, Cambridge, United Kingdom.
- Sanchez-Reillo, R., and Sanchez-Avila, C. "Iris recognition with low template size". *Proc. Int. Conf. Audio and Video-Based Biometric Person Authentication* (2001): 324-329.
- Sandholm, T. (e-Mediator) (2002) "A next generation electronic commerce server." , *Computational Intelligence*. Vol. 18(4), p. 656-676.
- Shen, W., Xue, D. and Norrie, D.H. (1998) "An agent-based manufacturing enterprise infrastructure for distributed integrated intelligent manufacturing systems." *Proc. of the 3rd Int. Conf. on the Practical Applications of Agents and Multi-Agent System.*, UK, p. 533-548.
- Stathis, K., de Bruijn, O. and Macedo, S. (2002) "Agent-based information management for connected local communities." *Living memory Interacting with Computers*. Vol. 14(6), p. 663-688.
- Tissue, C., Martin, L., Torres, L., and Robert, M. "Person identification technique using human iris recognition." *Proc. Vision Interface* (2002): 294-299.
- Van Dyke Parunak, H., Baker A.D. and Clark, S.J. (2001) "The AARIA agent architecture: From manufacturing requirements to agent based system design.", *Integrated Computer-Aided Engineering*. Vol. 8(1), p. 45-58.
- Vijaya Kumar, B.V.K., Marios S., et al., "Biometric Verification with Correlation Filters." *Journal of Applied Optics* 43.2 (2004): 391-402.
- Wada, H. and Okada, S. (2002) "An autonomous agent approach for manufacturing execution control systems." *Integrated Computer-Aided Engineering*. Vol. 9(3), p. 251-262.
- Wilders, R. "Iris recognition: an emerging biometric technology," *Proc. IEEE* 85 (1997): 1348-1363.

Wildes, R., Asmuth, J., Green, G., Hsu, S., Kolczynski, R., Matey, J., and McBride, S. "A machine-vision system for iris recognition". *Mach. Vis. Applic.* 9 (1996): 1-8.

Biometric Data Mining Applied to On-line Recognition Systems

José Alberto Hernández-Aguilar¹, Crispin Zavala¹, Ocotlán Díaz¹,
Gennadiy Burlak², Alberto Ochoa³ and Julio César Ponce⁴

¹FCAeI-UAEM,

²CIICAp. Universidad Autónoma del Estado de Morelos

³Universidad Autónoma de Ciudad Juárez

⁴Universidad Autónoma de Aguascalientes

^{1,2,3,4}México

1. Introduction

Data mining has become an increasingly popular activity in all areas of research, from business to science, biometrics being no exception. **Data mining** is the computer-intensive activity of exploring large data sets with the purpose of discovering, within a subset of data, some relationship of patterns or hypothesis that may be worthy of further study (Hernández-Aguilar et al., 2008; Amaratunga & Cabrera, 2004). According to a widely accepted definition, knowledge discovery in databases (KDD), more widely known as data mining, is a non-trivial process of identifying valid, novel, potentially useful and understandable patterns in data (Fayyad et al., 1996).

1.1 Basic definitions

But, what is biometric data mining? What does it study? First of all, let us clarify the meaning of biometric. According to Dunstone & Yager, 2009, there is a considerable amount of inconsistency among the terminology used within the biometric research and industrial communities. The best effort to date is ISO/IEC 17975-1, Information technology – Biometric performance testing and reporting. The following definitions of biometric and biometrics are consistent with this document:

Biometrics is the automatic identification of an individual based on his or her physiological or behavioural characteristics.

Biometric: A measure of a biological or behavioural characteristic used for recognition. There are four requirements for a biometric attribute: every person must have it, it should be sufficiently different for every person, it should remain constant over time, and it must measurable quantitatively.

Let us now define **Biometric data mining (BDM)**. BDM is the application of knowledge discovery techniques to biometric information with the purpose to identify underlying patterns. A principal objective of many data mining problems in biometrics research is to uncover characteristics of subsets of cases that are substantially different from the rest of the cases (Amaratunga & Cabrera, 2004). Consider the following:

Case study 1. Structure activity databases in the pharmaceutical industry are datasets prepared with the objective of studying the relationship between the biological activities of a series of compounds and their chemical properties. The primary goal is to identify ranges of values of $x=(x_1, \dots, x_n)$ associated with higher likelihood of in vivo activity.

Case study 2. An epidemiological database of several women of Indian Heritage, collected by the US National Institute of Diabetes and Digestive Kidney Diseases, studied the relationship between the incidence of diabetes among them using several predictors such as age, plasma concentration level, serum insulin level, diastolic blood pressure and body mass index (Blake and Merz, 1998). The objective was to identify the characteristics of the subjects associated with high incidence of diabetes.

Case Study 3. The selection of a small panel of proteins from the thousands of m/z points in mass-spectra, by means of selected variables, and their respective roles and interactions with the purpose to generate a classifier, that makes biological sense, for cancer diagnosis (Hilario et al., 2004).

Case Study 4. Several sports use anthropometric (biometric) analysis based on Data Mining such as Water-polo, Fencing, Synchronized Diving, Tumbling, Synchronized Gymnastics, Badminton, Archery and Curling (Ochoa et al., 2009). A study on people of Asian-Canadian descent has been performed in Quebec regarding their practice of Curling. The results have shown that the particular anthropometric characteristics of equilibrium, stamina and speed of this ethnic group minimize its possibility to play Ice Hockey, but are of high value to Curling.

Case Study 5. In (Ochoa et al., 2009) the biometric characteristics that Professional Wrestler Idols must have are discussed. Study reveals that age, height and weight always play a very significant role in wrestling, which should of course not be surprising. Besides, hidden patterns found by intelligent agents are related to the size of circuit, match records and cultural distances (ethnicity), as well as the selection of a good wrestler with specific attributes. A Wrestler with features similar to Scott Steel was selected as the most popular by the majority of societies because the attributes he has are adequate for other wrestlers.

In the aforementioned cases the relationship between a response variable Y and a set of predictor variables is studied. As shown in figure 1, the heart of the process is the learning step or automatic construction of a predictive model by generalizing from the training data. Here is where data mining finds applicability.

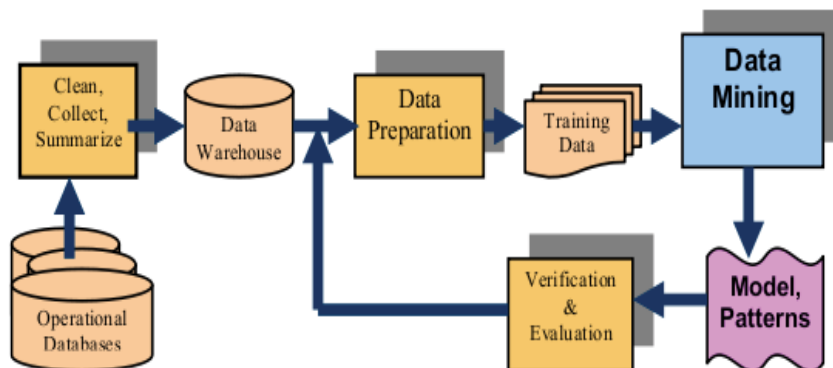


Fig. 1. The KDD (Knowledge Discovery Process) and data mining process (Han & Kamber, 2002)

Data mining is the process of searching through a large volume of data in an effort to discover patterns, trends, and relationships. Data mining is an umbrella term, and refers to a wide variety of processes and algorithms for knowledge discovery. The potential value of these techniques, applied to biometrics, is that it can automatically uncover hidden trends within a system, allowing researchers and system integrators to identify, diagnose and correct problems (Dunstone & Yager, 2009).

For example: Identifying a negative correlation between age and template quality would indicate that elderly people are more likely to have poor quality enrolments than young people. Another example is using a decision tree to classify specific behaviour of users during the enrolment process, see Figure 2. The resulting tree defines two goat populations (i.e. having difficulty matching against their own enrolments): children and adults who wore glasses.

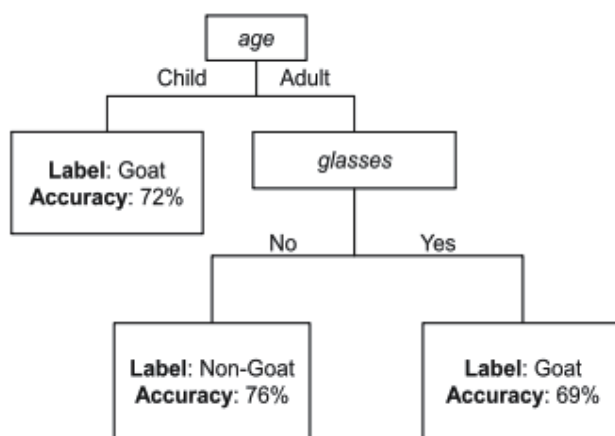


Fig. 2. In this case, 72% of children and 69% of adults who wear glasses exhibit goat-like behaviour (Dunstone & Yager, 2009).

According to (Van Der Ploeg, 2005), the information of the body answers the questions: who you are, how you are, and how you are going to be treated in various situations. - All this is information that is going to be processed through the networks (i.e. the Internet), databases, and algorithms of the information society. This author suggests that once translations of bodily characteristics are transformed into electronic data, they can become amenable to forms of analysis and categorization in ways not possible before. She suggests three levels of categorization:

- Authentication (1:1 comparisons). Classification of people as legitimate and truthful or illegitimate or fraudulent.
- Identification (1:n). Classifies and categorizes people according to the type and purpose of the database against which the biometric signal is checked. People may be identified as: someone with a criminal record, an illegal immigrant, a bad credit risk subject, or an employee with legitimate access to a high security facility.
- Aggregate level. Bringing together biometric information with other types of information, to generate, through a process of cross matching and data mining, more or less detailed and specific profiles that will be used to predict behaviour, assess dangerousness, or label as a risk.

Level three is used for preventing terrorist attacks, whereby suspicious electronic patterns of behaviour are tracked and recorded. According to (Neil, 2008) in the U.S., computer-driven searches look through an individual's web history, credit card transactions and personal background, allowing authorities to flag suspect behaviour. This author cites the example of a Muslim chemistry graduate who takes an ill-paid job at a farm-supplies store. What does this signify? Is he merely earning some cash, or using the job as cover to get his hands on a supply of potassium nitrate (used in fertiliser, and explosives)? What if his credit-card records show purchases of timing devices? Data mining allows analysts access to this information, but it is left to their judgement to decide whether or not it constitutes the beginning of a criminal plot, or just some innocent individual's "eccentric but legal behaviour".

1.2 Advantages and disadvantages of BDM

Main **advantages** of BDM include the following: most of the current authentication systems include a type of biometric measure like photo id, fingerprints and even iris recognition systems. Biometric characteristics are unique for each person. The cost of computers and biometric devices is decreasing and computer power is growing exponentially. The cost of building and operation of large data bases is decreasing. Every year more and more devices and systems are designed to support biometric characteristics. It is now possible to use several biometric characteristics (multi biometry) to support authentication. Most modern computer systems allow biometric identification upon logging in.

Disadvantages. A major disadvantage we can name is definitely the privacy concern for inappropriate use of biometric information in remote authentication settings, where biometric measurements are collected at remote sites and some weak points certainly exist. Dishonest entities such as servers that impersonate a user or perform data mining to gather information could be the source of successful attacks. Furthermore, the communication channel could also be compromised and biometric data could be intercepted or modified (Vetro et al., 2009).

To perform proper BDM, large data sets are mandatory. According to (Turban et al., 2004) to perform proper data mining a clear understanding of the problem is required. Identifying trends and patterns inside large volumes of data is a time consuming activity. Technical knowledge about biometric devices and systems is necessary. Performance of biometric devices can be reduced by dust or unintentionally by user misuse.

2. Application of biometric data mining to on-line recognition systems

Humans are believed to have several unique characteristics such as fingerprints, hand geometry, and iris (Tapiador & Singuenza, 2005), and nowadays almost all computerized systems involve an identity authentication process before a user can access requested services (Iglesias et al., 2008); for instance : secure access to buildings, logging into a computer system, laptop or cell phone, or logging in with the purpose of making e-commerce transactions on ATM machines, or to use on-line assessment systems (Hernandez-Aguilar et al., 2010). The applications of biometric identification range from forensics and law enforcements to novel biometrics-based access to personal information that protects user privacy and mitigates fraud. Biometric systems recognize users based on two types of characteristics: 1) behavioural (i.e. voice, signature, keystroke dynamics or haptics) or 2) physiological characteristics (i.e. fingerprints, iris pattern, face image or hand geometry). We will concentrate our analysis on

those technologies that make applying biometric data mining to on-line recognition systems possible. In Table 1 a summary of these techniques and their reported overall accuracy is shown.

Technology	Overall accuracy	Reported on
Behavioural		
Keystroke patterns/dynamics	95 %	(Revett et al. 2005)
Haptics and virtual reality	95 - 97%	(Iglesias et al. 2008)
Mouse movements	95.00%	(Kaklauskas et al. 2008)
Stylometry	82.00% or less	(Shalhoub et al. 2010)
On-line behaviour	56.0%-98.0%	(Yan & Padmanabhan, 2010)
Physiological		
Fingerprint	98.00 - 100%	(Umamaheswari et al. 2007)
Face	90.00%	(Lovell & Chen, 2008)

Table 1. State of the art Biometric data mining technologies for on-line recognition systems

Converting physiological characteristics of users into techniques for biometric identification has been an active research for several years, and combined with data mining techniques, produces biometric data mining, a hot topic of discussion in the scientific community.

2.1 Keystroke patterns/keystroke dynamics

Research focused on Keystroke patterns, in terms of Keyboard Duration and keyboard latency. Evidence from (Revett et al., 2005) indicated that when two individuals entered the same login details, their typing patterns would be sufficiently unique as to provide a characteristic signature that could be used to differentiate one from another. Keystroke dynamics is a cost effective means of enhancing computer access security, and has been successfully employed as a means of identifying legitimate/illegitimate login attempts based on typing style of the login entry.

2.2 Biometric behavior through haptics and virtual reality

Haptics refers to the science of sensing and manipulating through touch in real and virtual environments. Haptics technology allows users to interact via sense of touch by applying forces, vibrations and/or motions to users. Examples are vibrating phones, gaming controllers, force-feedback control knobs in cars and the wiimote controller. Data directly generated by the user that interacts with the system is recorded and used for authentication purposes. Therefore, Haptics can be seen as a mechanism to extract behavioural features that characterize a biometric profile for authentication. (Iglesias et al., 2008) applied non-linear transformations to the original feature space to produce Euclidean 3D spaces preserving the similarity structure of the samples, which were represented with Virtual Reality (VR) techniques. These new spaces were analysed using visual data mining to know how certain features (i.e. position, pressure and torque) contain more meaningful information that can characterize a biometric profile when signing in.

2.3 Mouse movements

User identification using mouse movement parameters has been discussed by different researchers (Eusebi et al., 2008; Brodley & Pusara, 2004; Weiss et al., 2007). In (Kaklauskas et al., 2008) a Web-based biometric Mouse decision support system for user's emotional and labour productivity is discussed, and reported as able to analyse data from a biometric mouse – designed for the same authors- and e-self reports. They mixed different biometric parameters, including physiological (skin conductance, amplitude of hand trembles, and skin temperature), physiological (self-reports) and behavioural (mouse pressure, speed of mouse pointer movement, acceleration of mouse, etc.) and made a correlation between the user's emotional state and labour productivity. The possibilities of the biometric mouse are remarkable; it is able to measure the temperature and humidity of a user's palm and his/her intensity of pressing. These parameters could be used to identify suspicious behaviour and single out impostors.

2.4 Stylometry

Stylometry is a discipline that determines authorship of literary works through the use of statistical analysis and machine learning. When someone authors a literary work, document, or email they leave behind certain attributes to their writing style that can be analysed and used to determine other works by the same author. The rise of the Internet has opened new uses for stylometry in the area of e-mail, social networking and various types of digital content. In (Shalhoub et al., 2010) a study to identify if some stylometric tools (i.e. C# tool testing) can correctly assign authorship of electronic mail to its original author is presented, and has produced results which indicate moderate accuracy - suggesting that none of the tools evaluated is capable of correct author identification.

2.5 On-line behavior

Research in biometrics suggests that the time period a specific trait is monitored over (i.e. observing speech or hand writing long enough) is useful for identification, most notably the research in Web usage mining (Adomavicius & Tuzhilin, 2001; Pirolli, 2007) which suggests that user behaviour is not random and there is often a purpose that translates into revealed on-line behaviour. In (Yang & Padmanabhan, 2010) a data mining analysis of the effect of observation time period on user identification based on on-line behaviour is presented, and the identification of unique behavioural characteristics that can possibly serve as identifiers is discussed. The quality of data can be measured by the features created from his/her behaviour. For user identification from on-line behaviour, quantity is a measure of how much user data is observed. In this research the authors use aggregation to describe the processes of observing and collecting data over long periods of time. Specifically they use aggregation over multiple web sessions. Results suggest that at the user-centric level it is possible to build reasonably accurate models identifying the user by observing enough data, at least for some users.

2.6 Fingerprint

Fingerprinting is the first biometric science used worldwide for the validation and verification of an entry into specific tasks and is one of the most popular techniques to perform biometric recognition. However, Fingerprint classification and recognition is still an open and very challenging problem in real world applications. In (Umamaheswari et al.,

2007) fingerprint classification and recognition using data mining techniques is discussed, and the proposed method involves various stages like image enhancement, line detector based feature extraction and neuronal network classification using Learning Vector Quantization and Kohonen networks. Optimization of neuronal parameters and recognition of images is done by a genetic algorithm K nearest Neighbour. The exact image is recognized from the classified database using a crisp and fuzzy K Nearest Neighbour algorithm. The resulting system is one of the most reliable methods of personal verification (98 to 100%), and can be used for authority access verification, ATM verification and other civilian applications.

2.7 Face recognition

According to (Lovell & Chen, 2010), there are several applications of data mining for face recognition: 1) Person recognition and location services on a planetary wide sensor net, 2) Recognizing faces in a crowd from video surveillance, 3) Searching for video or images of selected persons in multimedia databases, 4) Forensic examination of multiple video streams to detect movements of certain persons, 5) Automatic annotation and labelling of video streams to provide added value for digital interactive television. All of these applications are subject of intensive research around the world and require on-line processing.

2.8 Architecture of BDM systems

The underlying architecture for a Biometric Data Mining System consists of a client-server application (Burlak et al., 2005). On the server side resides the biometric information derived from in site or on-line enrolments, as well as all of the information and algorithms required to perform the data mining process. What's worth noting is that this information and algorithms can be requested as a web service, but if this approach is used then computational cost will increase. On the client side there may be important practical applications that might mitigate on-line fraud and identity theft, along with client-side software from a trusted third party that will track client-side activities to build users identification models (Yang & Padmanabhan, 2010). Such models may be used to provide behavioural authentication services on behalf of the user. For instance, when the user makes a large on-line brokerage transaction, the financial institution may, in real time, request the client-side software for a user score, and if the user is identified as who he claims to be, the firm may proceed with the transaction. The challenge consists on designing a system with accurate user identification models. This requires a deeper understanding of the factors that can result in better or worse identification accuracy.

3. Privacy concern

Some research on Internet privacy has examined various aspects of privacy regulation and user privacy concerns. Nowadays, the Internet has heightened a variety of users' concerns regarding privacy. The concept of privacy is "the claim for individuals to determine when, how and to what extent information about themselves can be communicated to others" (Westin, 1967). Other concepts refer to the degree to which a website is safe and user information is protected. This dimension holds an important position. Users perceive significant risks in the virtual environment of e-services and e-commerce stemming from the possibility of improper use of their financial and personal data. There is good reason

for the Anglo-American public to be resistant to national identification cards. Yet the British and American governments seem increasingly willing to neglect privacy in pursuit of personal data (Hansen, 2009). According to (Lovell & Chen, 2008) Privacy concerns that have hindered public acceptance of these technologies in the past are now yielding to society's need for increased security while maintaining a free society.

3.1 Perception of on-line users

Information security is increasingly recognized as a vital element for ensuring wide participation in Internet use. The successful use of the Internet in the Society depends on trust and confidence in our information infrastructures. Within this context, the real effect of the security problems is the inhibition of the development of the Internet use and of e-commerce as a whole.

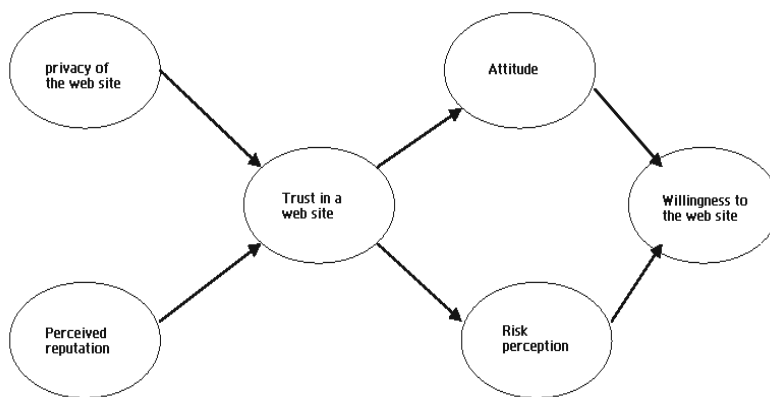


Fig. 3. Model of trust in a web site by the users

There are three variables that influence people's perception of Internet safety:

- The involvement level with respect to the category of products/services sold by the company.
- Demographics (age, gender, education level).
- Use of the Internet (frequency, familiarity).

Reputation is defined as the belief in that the web site is honest and concerned about its users. Users of Internet can favour sites that represent merchants or organizations with which they are familiar from traditional channels. This is due to the fact that this familiarity increases any positive affect as well as any positive cognition on the part of the user (Jarvenpaa et al., 2000).

Specifically, the user's preoccupation or concern with privacy includes the proliferation of databases, the great volumes of personal data being collected, and the possibility of privacy violations and loss of control in the process of collecting, accessing, and utilizing this information (Hiller & Cohen, 2001). Generally speaking, the more standardized the biometric technology, the more interoperability between different systems and databases is attained, and the more ubiquitous and pervasive the categorization of people can become (Van der Ploeg, 2005).

In fact, the rapid progress in the development of communication technologies, biometrics, sensor technologies and data storage and analysis capabilities are perceived as causing

constant pressure on the fundamental right to privacy for both economic and security reasons (Pavone & Pereira, 2009).

The rapid technology changes, accelerated general acceptance of the Internet, Social Networks, E-Commerce, and the development of more sophisticated methods of collecting, analysing, and using personal information have made privacy a major socio-political issue in a lot of countries.

In many social networking sites, users are responsible for deciding what information to disclose and whether or not to protect any of that information with privacy settings. From the time they join the community, users are challenged to create a mental model of their on-line audience and desired levels of privacy, and then determine how to best match the disclosures and accessibility of their personal information (Strater & Lipford, 2008). Some times the users have personal experience of privacy intrusions, usually in the form of unwanted contact from an unknown person.

Security as well as privacy is a need of society and its members. Designing security technologies without keeping privacy requirements in mind may result in systems which create additional risks to society and where side-effects are difficult to control.

Privacy issues have increasingly attracted the attention of the media, politicians, government agencies, businesses, etc. In addition, the public has become increasingly sensitised to the protection of their personal information.

The perception may also be due, at least in part, to the level of familiarity associated with this type of sites, and technology use. Only when the users have noticeable and disturbing events, such as a privacy intrusion, do users modify the privacy level. Based on the perceived intrusiveness many users use the privacy controls of that feature. One challenge is that users learn what to disclose and what to protect over time, both through the social norms of the community and through their own experiences.

4. Prototype BDM system applied for on-line assessments

Virtual proctoring involves using biometric technology to monitor students at remote locations. For virtual proctoring, using a layered approach depending on critical maturity of the test is recommended. With high stakes tests, video monitoring and a biometric measure such as iris scanning may be used. For medium stakes tests, a single biometrics measure may be acceptable (BSU, 2006). Despite most on-line assessments being located in the middle of both definitions, we consider the fact of high levels of cheating in remote assessments. On one hand, fingerprint recognition is a single biometric measure, the cheapest, fastest, most convenient and most reliable way to identify someone. And the tendency, due to scale, easiness and the existing foundation, is that the use of fingerprint recognition will only increase. Cars, cell phones, PDAs, personal computers and dozens of products and devices are using fingerprint recognition more and more (Tapiador & Singuenza, 2005). One current trend is to incorporate fingerprint scanners into personal computers, laptops, and mice. In addition, computer networks and large databases can be secured using fingerprint technology. This is a hot topic of discussion since the phenomenon of the Internet and the development of Intra nets has spawned new digital technologies such as E-commerce and online services. Besides, users are more willing to use fingerprint recognition than iris recognition, as they believe it is safer, health-wise. Unfortunately, fingerprint recognition is used merely for authentication, and then what? The student is free to use any media to cheat on the exam. To avoid that situation we considered the

possibility of using web cams. Web cams are inexpensive and most of the students are used to dealing with them, they are part of their common tools for work and chat. Naturally, some students reject the possibility to be monitored, and the percentage varies from country to country, but it is our intention to measure this figure as a part of our research. Based on the aforementioned, we propose the mixed use of video monitoring, by means of web cams, and fingerprint recognition to provide a secure on-line assessment environment.

4.1 Technical requirements

The Server Side

- Keep information of biometrics measures (fingerprints) and associated student information in data base.
- Scanning of finger prints (enrolment of students).
- Provide a recognition tool to determine validity of fingerprint and grant authorization to on-line assessment.
- Monitor remote students by means of web cams located in remote locations.
- Support the on-line assessments process.
- Provide security mechanisms to ensure confidentiality and validity of data: Encryption of data transmitted and received and log files.

The Client Side

- Scanning of finger prints.
- Enrolment of students (optional).
- Avoid unauthorized access to on-line assessment.
- Show the diagnosis of security.
- Provide capacity of students' monitoring using web cams during assessment process.
- Provide mechanisms for client set-up, students authentication (using fingerprint), and evaluation preferences.
- Support the evaluation process and show final results of evaluation.

4.2 Performance schema (n-Tier C-S system)

We separated the application in two main modules: the first one is in charge of the conduction the on-line assessment, and the second one in charge of fingerprint recognition and web cam monitoring in real time.

The server must be in listening mode waiting for Clients that require a service. In order to use fingerprint recognition, the first step is to enrol students –top, right side in Figure 4-, the student's fingerprint is saved and indexed in the Features Database. We highly recommend separating this from the Assessment System Database, using even separated servers, to improve overall system performance A Student Personnel ID is assigned in the features database which is used to link the students' personnel information with the fingerprint image.

If the Server is in listening mode and the student has been enrolled, the assessment process can start. The student enters the on-line assessment application, and when the system requires the username and password, the student uses the Mouse Id –superior right side of Fig. 4- to scan his/her fingerprint.

The fingerprint is verified in the Features Database, and if it is recognized as valid, the Server authorizes access to the on-line assessment application. If the print is invalid, an error

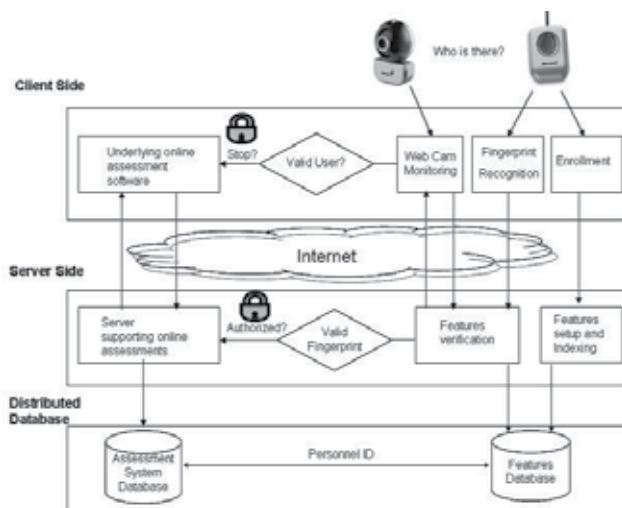


Fig. 4. Fingerprint recognition in real time in online assessments. The Client side (at top) and the Server side (at bottom)

message is sent to the Client instructing it to try again. If, on the other hand, if the student's fingerprint is valid, the user is authenticated into the system, the evaluation process starts and web cam transmission is initialized at the Client Side to conduct real time monitoring by means of multitasking. If someone else tries to get the control of the computer during the on-line assessment, the evaluation process is finished prematurely, and the results are sent to server side to be processed as they are. If the evaluation process is finished successfully, the assessment is processed at the Server Side, and the final results of evaluation and security status are shown at the Client Side. For a detailed review of the performance schema refer to (Hernández-Aguilar et al., 2008).

4.3 Methodology

For our experiment, we selected a random sample of students ($n=102$) from the José María Morelos y Pavón High School, located in Temixco, Morelos, México. We carried out two evaluations, a control evaluation (paper and pencil), and a second evaluation with our on-line assessment system with biometric recognition.

- **Tests design.** Tests were designed by the professors on August 5th and 6th 2007, of which one was implemented for the on-line assessment using our authoring tool. The tests consisted of 30 questions with similar level of complexity; we evaluated arithmetic, algebra, geometrics, and trigonometric subjects.
- **Setting up.** Computers were prepared and our on-line assessment client software and biometric devices installed, network connectivity was tested.
- **The traditional test.** The paper and pencil test was conducted on August 14th 2007.
- **Enrolment.** Students were enrolled into the system by taking their left-hand index fingerprint on August 15th 2007. We made sure the students were identified by the system after their enrolment.
- **The on-line assessment with biometric recognition test.** The test was conducted in the Computers Network Laboratory located at the High School facilities from August 16th to August 17th 2007. Each computer used in the experiment had a Microsoft Fingerprint

Reader attached, a web cam, and a broad band connection to our server as well as our proprietary client system. First of all, the students were instructed in the usage of the system, and were explained that a web cam was monitoring their activities. At that point, they authenticated by means of their fingerprint into our Server System and the computerized assessment started. The use of calculators and cellular phones was forbidden.

- **The Survey.** At the end of the exam we performed a survey to determine the students' profile and perceptions about our system's operation.
- **Statistical Analysis.** Data was processed using descriptive analysis, using relative numbers and percentages using Ccount gnu free software.

4.4 Preliminary results and discussion

In this test with an obtained FAR of 99.99% and FRR of 97.09%, only one female student could not be recognized despite several trials. Her fingerprint template cannot be understood by the system due to her fingerprints having stain-like shape. Similar cases are registered in (Michigan Org, 2007). We worked around this problem by providing her with a username and a strong password. The average grade in the paper and pencil test was 3.8 while the on-line average was 3.5. This difference can be explained with the fact that a small percentage of students must improve their computer skills. We noticed that video games and chat could improve students' skills and performance in on-line assessments. In general, students perceived our system as faster, easy to use and secure, fingerprint recognition playing an important role in this last point. However, they dislike time limited answers, and 13% dislike web cam monitoring. They felt under pressure, get nervous and dislike being monitored or watched. 20% noticed a way to cheat using a system like ours. We made an in-depth analysis and discovered that students with poor performance (low grades) are more willing and likely to cheat.

4.5 Future work

This work is by no means complete, and could be improved by a number of future research directions. First, we want to compare the results obtained with Weka as a DM engine with results obtained using other software tools like Matlab and SPSS. Secondly, we want to improve the human-computer interface and assessment methodology by analysing students' comments and users' feedback. Third, regarding biometric recognition, we want to improve facial recognition since at this point of our research we can only detect a student's presence or absence, while comparing face patterns automatically by means of photo Ids stored in our features databases is our final intention. Four, we want to include behavioural characteristics in our BDM analysis and verify its feasibility to identify remote users. Finally, we want to test the newest fingerprint scanners and bio-sensors included in new mice, keyboards and in some laptops and try to incorporate them to work within our system.

5. Conclusions

We consider the use of BDM to identify student cheating and impersonation in on-line assessments very important. We believe that Data mining can be successfully used to find out whether the student cheats in on-line tests as long as enough information is provided. Best results can be obtained from information received from remote places and during different intervals of time.

To perform a better DM analysis about student's cheating in on-line assessments we have to pay attention not only to the student's perceptions and behaviour, but to the professor's teaching style as well. The information received from such an analysis can result in very important findings in discussing and improving the testing environments as well as the institution's general philosophy. To study this process behavioural data mining can be performed.

Biometric data mining is a promising technology proposing to deal with the problem of "who is there?" in not only on-line assessments, but on-line recognition in general. This technology is generally very well accepted but must be improved to be perceived as unobtrusive.

6. Future of biometric data mining applied to on-line recognition systems.

Many techniques analyse Mouse Movement, Stylometry, and Keystroke Capture data sets using data mining techniques. Numerous algorithms and methodologies have been used during the last five years. In the future, data sets will be using PredictiveApriori (Ibk), while the Stylometry data set will additionally be analysed using simpleKmeans. All of these machine-learning methods have differences in applicability, meaning there is no one best method; rather, there are only optimal methods, depending on the particular data set. It is important to note that most of the algorithms in reality do not produce results which are 100% accurate. An observed trend is to mix several physiological and behavioural characteristics to perform a better identification of remote users, and consequently high quality data mining, and we will find more creative ways to mix those parameters and techniques in the near future.

For example, although highly accurate results were obtained using sophisticated learning methods on many datasets, some approaches were more successful than others. The most successful approaches have been shown in detail - and future researchers may find that they can improve the results found in the literature using similar techniques. Future researchers may be particularly interested in trying different approaches for their authentication experiments. In these experiments, a community of subjects could be authenticated against another community of subjects. However, for the use of the biometric information to identify an individual, it would be more efficient to attempt authentication based solely on the subject in question. Experiments that would lead to an adequate system for identifying individuals would require splitting the data set into separate data sets that hold only the within and between class records pertaining to each of the subjects.

In many applications of On-line Social Networking as Facebook, when the user moves the location, for example another country with government restrictions as Belarus, the user's authenticity is proven by answering questions about the person's contact photos, and analysing the time required to answer each question. Such an approach enables detection of an impostor. Once there are people recognized within one's social networking contacts, the user is given 7 pictures of which he must identify at least five, and is given a 15 seconds time frame to do so for each.

Authentication and identification tests will be run in order to produce a comparative study on Mouse Movement, Stylometry, and Keystroke Biometric data. In most cases, the identification tests will result in higher accuracies than the authentication tests. Reactive recommendations will be made for legitimately increasing the classification accuracy of the authentication experiments.

The k-nearest neighbour approach will be used with cross validation and 80% splits on the Stylometry and Keystroke Capture data sets, which showed high accuracy results for non-training data before. Decision rule and k-means clustering will be used on the Stylometry data set and make for interesting experiments, in that the decision rules may be useful to improve future researches and in that the clustering algorithm will show an average accuracy that is similar to the accuracies achieved using the k-nearest neighbour approach with cross validation on the modified Stylometry data. Many researches have extended previous studies by running additional experiments on the Mouse Movement, Stylometry, and Keystroke Biometric data, new and previously obtained, using data mining tools like Orange, Weka or SAS. The data mining algorithms with which the experiments will be conducted are widely used and provide an entry point for future researchers into the use of data mining with biometric data sets.

Biometric data mining will be used to identify emotions and will identify suspicious biometrics to fight against crime and terrorism - to see what is being done now refer to (Security Focus, 2006) and related web sites. Biometric technologies are coming of age due to the need to address heightened security concerns in the 21st century. Consequently, the performance and robustness of systems will increase significantly but more research effort is necessary. In the future, the use and reconstruction of 3D images, as well as the use of virtual reality will increase, and biometric data mining will analyse this data as well as physiological and behavioural characteristics simultaneously.

7. References

- Adomavicius, G. & Tuzhilin, A. (2001). Using data mining methods to build customer profiles. *IEEE computer*, Vol. 34 (2), 74-82
- Amaratunga, D. & Cabrera, C. (2004). Mining data to find subsets of high activity. *Journal of statistical planning and inference*, Journal of Statistical Planning and Inference. 122, 23-41, ISSN 0378-3758
- Blake, C.L. & Merz, C.J. (1998). UCI Repository of machine learning databases. <http://archive.ics.uci.edu/ml/>
- Burlak, G.; Hernandez, J.A. & Zamudio-Lara, A. (2005). The Application Of Online Testing For Educational Process In Client-Server System, *Proceedings of CONGRESS: IADIS International Conference*, pp. 389-392, ISBN 972-8924-04-6 , October 2005, Lisboa, Portugal
- Brodley, C. & Pusara, E. (2004). User Re-Authentication via Mouse Movements, *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, pp. 1-8, ISBN 1-58113-974-8, Washington, D.C., ACM, NY, USA
- BSU Ball State University (2006) Technology assessment reports on The use of Biometrics in education. <http://web.bsu.edu/dpmuller/biometrics/>
- Dunstone, T. & Yager, N. (2009). *Biometric System and Data Analysis. Design Evaluation and Datamining*. Springer Science+Business Media, LLC. ISBN-13: 978-0-387-77625-5. USA
- Eusebi, C.; Gliga, C.; John, D. & Maisonave, A. (2008). A Data Mining Study of Mouse Movement, Stylometry, and Keystroke Biometric Data, *Proceedings of Student-Faculty Research Day*, pp. B1.1-B1.6, Pace University, May 2008, NY, USA.
- Fayyad, U. et al. (1996). From Data mining to knowledge discovery: an overview. *Advances in Knowledge Discovery and Data mining*. MIT Press. 1-34.

- Han, J. & Kamber, M. (2002). *Data mining Concepts and Techniques*, Morgan Kaufman Publishers, ISBN 1-55860-489-8, CA, USA.
- Hansen, M. (2009). Concepts of Privacy-Enhancing Identity Management for Privacy-Enhancing Security Technologies, *Proceedings of PRISE Conference: Towards privacy enhancing security technologies – the next steps*, pp. 91-93, April 2008, Vienna.
- Hernández-Aguilar, J.A.; Burlak, G. & Lara, B. (2010). Design and Implementation of an Advanced Security Remote Assessment System for Universities Using Data Mining. Resúmen de Tesis Doctoral. *Revista Computación y Sistemas*, Vol.13, No.4, 463-473, ISSN 1405-5546
- Hernández-Aguilar, J.A.; Ochoa, A.; Andaverde, J. & Burlak, G. (2008). Biometrics in online assessments: A Study Case in High School Students, *Proceedings of 18th International Conference on Electronics, Communications and Computers Conielectcomp 2008*, pp. 111-116, ISBN 0-7695-3120-2 & ISBN 978-0-7695-3120, Cholula, February 2008, IEEE-Explore, Puebla, México
- Hilario, M.; Kalousis, A.; Prados, J. and Alain, P. (2004). Data mining for mass spectra-based cancer diagnosis and biomarker discovery. *Biosilico*, Vol.2, No.5, 214-222, ISSN 1741-8364.
- Hiller, J. & Cohen, R. (2001). *Internet Law and Policy*, Prentice-Hall, ISBN 9780455222639, Upper Saddle River, NJ.
- Iglesias, R.; Orozco, M. & Valdes, J. (2008). Characterizing Biometric Behaviour Through Haptics and Virtual Reality, *Proceedings of ICCST*, pp. 174-179, ISBN 978-1-4244-1817-6, Lillehammer, April 2008, IEEE Computer Society, Norway
- Jarvenpaa, S.; Tractinsky, N, and Vitale, M., (2000). Consumer trust in an Internet store. *Information Technology and Management*, Vol.1, No. 1-2, 45-71, ISSN 1741-5179.
- Kaklauskas, A.; Zavadskas, E.K.; Seniut, M. et al. (2008). Web-based biometric mouse decision support system for user's emotional and labour productivity analysis, *Proceedings of The 25th International Symposium on Automation and Robotics in Construction*, pp. 69-75, Lithuania, June 2008, Institute of Internet and Intelligent Technologies, Vilnius
- Lovell, B. C. & Chen, S. (2008). *Robust face recognition for data mining*. In John Wang, editor, *Encyclopedia of Data Warehousing and Mining, volume II*, Idea Group Reference, ISBN 1-59140-559-9, Hershey, PA.
- Michigan Org. (2007) Fingerprint Classification.
http://www.reachoutmichigan.org/funexperiments/agesubject/lessons/prints_ext.html
- Neil, R.L. (2008) Privacy, Biometrics and the war on terror.
http://www.acus.org/new_atlanticist/biometrics-civil-liberties-and-war-against-terror
- Ochoa, A.; Hernández, A.; Sánchez, J.; García, Y. et al. (2009). Identify an Adequate Anthropometry to Water Polo Using Social Data Mining, *Proceedings of International Conference on Electrical, Communications, and Computers*, 2009, pp. 144-147, ISBN ISBN: 978-0-7695-3587-6, Cholula, February 2009, IEEE Xplore, Puebla, México
- Pavone, V. and Pereira M. (2009). The privacy Vs security dilemma in a risk society. *Proceedings of PRISE Conference: "Towards privacy enhancing security technologies – the next steps"*, pp. 109-127, Vienna.
- Pirolli, P.L. (2007). *Information Foraging Theory: Adaptive Interaction with Information*, Oxford University Press, Cambridge, UK

- Revett, K.; De Magalhaes, S.T. & Santos, H. (2005). Data Mining a Keystroke Dynamics Based Biometrics Database Using Rough Sets, *Proceedings of Portuguese conference on Artificial intelligence, 2005. epia 2005*, pp. 188-191, ISBN 0-7803-9366-X, Harrow Sch. of Comput. Sci., Westminster Univ., Dec 2005, IEEE Xplore, London
- Security Focus. (2006). Biometric polygraph next for airport security? Biometric polygraph next for airport security? <http://www.securityfocus.com/brief/281>
- Shalhoub, G.; Simon, R.; Iyer, R. et al. (2010). Stylometry System – Use Cases and Feasibility Study, *Proceedings of Student-Faculty Research Day, CSIS*, pp. A3.1-A3.8, Pace University, May 2010, NY, USA.
- Strater, K. & Lipford, H. (2008). Strategies and Struggles with Privacy in an Online Social Networking Community, *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 1*, pp. 111-119, ISBN 978-1-906124-04-5, Liverpool, UK, 2007, British Computer Society, Swinton, UK
- Tapiador, M. & Singuenza, J.A. (2005). *Tecnologías biométricas aplicadas a la seguridad*, AlfaOmega Grupo Editor, S.A. de C.V., México, D.F.
- Turban, E.; Aronson, J.; Liang, T. & Sharda, R. (2004). *Decision Support and Business Intelligence Systems*, Pearson Education, Inc., ISBN 0-13-198660-0, New Jersey, USA
- Umamaheswari, K.; Sumathi, S.; Sivanandam, S.N. & Anburajan K.K. (2007). Neuro Genetic-Nearest Neighbor Based Data Mining Techniques for Fingerprint Classification and Recognition System. *ICGST-GVIP Journal*, Vol.7, No.3, November 2007, 39-50, ISSN 1687-3998
- Van der Ploeg, I. (2005). The Politics of Biometric Identification, In: *Biometric Technology & Ethics*, BITE, Policy Paper no. 2, 1-16, www.biteproject.org/documents/politics_of_biometric_identity%20.pdf
- Vetro, A.; Drapper, S.; Rane, S. & Yedidia, J. (2009). *Securing Biometric Data*, MERL, Technical Report TR-2009-002, Massachusetts, USA.
- Weiss, A.; Ramapanikar, A.; Shah, P.; Nobel, S. & Immohr, L. (2007). Mouse Movements Biometric Identification: A feasibility Study, *Proceedings of Student/Faculty Research Day*, Pace University, May 2007, NY, USA
- Westin, A. (1967). *Privacy and Freedom*, Atheneum, New York
- Yang, Y. & Padmanabhan, B. (2010). Toward user patterns for on-line security: Observation time and online user identification. *Decision Support Systems*, Vol. 48, 548-558, ISSN 0167-9236

Parallel Secure Computation Scheme for Biometric Security and Privacy in Standard-Based BioAPI Framework

Arun P. Kumara Krishnan, Bon K. Sy and Adam Ramirez
Queens College & Graduate Center/CUNY
Computer Science Department, SB A202
Flushing NY 11367

1. Introduction

A biometric system may be viewed as a composition of various segments; i.e., unit(s) capturing biometric samples, unit(s) preprocessing captured samples into formats from which biometric data can be extracted, or unit(s) matching biometric samples with enrolled signature(s). Various vendors may provide these components. Traditionally, once a unit is chosen from a vendor, the user is restricted in their choice of interfacing units. This restriction stems from the burden of building a custom interface and the pool of available vendors who offer units compatible with the chosen unit. The BioAPI consortium, an organization of various industry leaders, has recently developed the ISO/IEC 19784-1:2005 (also known as BioAPI 2.0) standard to address the interoperability issue. The BioAPI 2.0 framework provides a common interface for biometric systems to subscribe services of standard compliant units made available by various vendors. This allows one to pick and choose between vendors or swap units from different vendors with minimal overhead. Thus, if someone develops a more efficient/accurate matching unit, the existing matching unit can be essentially “hot swapped” with minimal interruption in service; no longer is one limited to continue using a less desirable system because of expensive migration costs. BioAPI provides an excellent framework for interoperability, but lacks provisions for the security or privacy of biometric data. To illustrate this point, consider a typical scenario:

A biometric based authentication or credentialing system requires an individual to enroll their biometric data (e.g., fingerprint) so that the biometric data can be stored for future comparison purposes. For authentication, user A presents their biometric data to be compared against enrolled biometric template(s). User A must present their biometric data to a remote matching engine for it to compare with the enrolled biometrics. In this case, user A is forced to reveal their biometric data to another party.

Various approaches exist to ensure privacy and security in biometric systems. These approaches fall into two general categories: lossy (Newton 2005) and lossless (Yao 82, Sy 2009). The problem with a lossy approach is that the anonymization of biometric data is data dependent and may not be extendable from one application to another. On the other hand, lossless approaches such as secure computation (Yao 82) are practical only on a limited set of functions characterized by linear arithmetic operations. At the time of this writing, secure

computation based on full privacy homomorphism (Gentry 2009) for computing privately any arbitrary function is not practical due to its prohibitively expensive cost in terms of computation time. Therefore, there remains a need to protect the privacy of biometric data at the same level as a lossless approach yet at the same time achieving real time results.

We have developed a unique secure computation privacy protection mechanism - referred to as SIPPA (Secure Information Processing with Privacy Assurance) - for privacy preserving biometric data retrieval and processing. SIPPA (Sy 2010) is essentially a hybrid approach where it is a lossless approach when the biometric data being compared is exactly the same, and it is a lossy approach when the biometric data being compared are not exactly the same. The basic idea behind SIPPA is to model data ($D1/D2$) as a linearized vector, and to transform the vector into a symmetric matrix. To compare the two data sets, we take advantage of the approximately consistent relationship between the norm deviation of the data and the norm deviation of the corresponding eigenvectors weighted by the eigenvalues. For privacy protection, neither party will share data, the corresponding matrix and the eigenvalues/eigenvectors with the other party. Instead, P1 and P2 will engage in a two-party secure computation to uncover the similarity between the weighted eigenvectors. Through the approximately consistent relationship between the data and eigenvectors, the similarity between the data can be derived, and used subsequently as a basis for the reconstruction of the source data D2. As we will discuss in the later section, SIPPA offers two very attractive properties:

1. The private data D2 can be reconstructed by P1 only if P2 sends helper data to P1 after judging D1 and D2 to be sufficiently similar. In other words, P2 has control over the sharing of his/her private data D2 - a desirable property often mentioned by privacy advocates (Solove et al 2006) as facilitating privacy by providing control over information.
2. The accuracy of the reconstruction depends on the level of closeness/similarity between D1 and D2. In other words, the ability of P1 to reconstruct accurately D2 decreases when P1 does not already know some D1 that bears similarity to D2 - typical scenario of biometrics where the biometric samples from the same individual are "sufficiently similar" but are seldom identical.

For SIPPA to be practically useful and to deliver real time computational results, a parallel processing architecture is necessary. We illustrate the design and implementation of a parallel processing architecture for SIPPA (PSIPPA, MPSIPPA), and the realization of PSIPPA (clients and servers) as service components in the BioAPI 2.0 framework. Our attempt to realize PSIPPA as a service component has two specific objectives. First, PSIPPA is exposed as a biometric service in the BioAPI framework. Second, PSIPPA presents itself as a case study to explore the possibility of extending the BioAPI framework to incorporate an innovative slice-based provisioning mechanism of biometric services.

2. Related work

The privacy of personal biometrics is generally protected by means of two approaches; namely, lossy data processing and lossless data processing. In security surveillance, protecting privacy through the lossy approach is not uncommon. Preserving privacy through the lossy approach is achieved by protecting the private content typically by perturbation, randomization or masking of the original data to the extent that it could still be useful for security purposes.

Many face de-identification techniques (Gross et al 2005) for privacy protection are based on lossy approach. The basic idea is to conduct lossy anonymization to mask away granular biometric face information not essential to its end goal on security identification or surveillance. For example, Newton et al. (2005) proposed a k-Same model based on the k-anonymity framework. The k-Same model takes the average of k face images in a face set and replaces these images with the average image; therefore, each face image presented to a face recognition/comparison system cannot be distinguished from at least k-1 faces in the gallery. In general, information leakage could be a significant risk when k is small and/or known unique aspect of an individual is not sufficiently anonymized. In other words, the degree of privacy protection based on lossy anonymization is data dependent and may not be extendable from one application to another that have different privacy requirements.

An approach towards lossless privacy protection is Secure Multi-party Computation (SMC). SMC protects computational privacy while preserving content (Du & Atallah 2001). In other words, the original content on personal biometrics can be retrieved and used in some secure computation scheme for deriving event information of interest, but the scope of derivation is limited to what is allowed by the private computational mechanism of the process.

Generally speaking, SMC deals with the problem in which two or more parties with private inputs would like to compute jointly some function of their inputs, but no party wishes to reveal its private input to other participants. For example, a physician (party P1) wants to compare a medical image of a patient containing personal biometrics with the medical image of another patient stored in a hospital database. The data custodian (e.g., party P2 who is the database administrator) in the hospital and the physician (party P1) may participate in a SMC protocol to jointly compute the output of a matching function that compares the personal biometrics in the medical images. In another example, a user (party P1) and the authentication server (party P2) may jointly compute the distance function based on the user voice sample and the enrolled voice template withheld by the authentication server for biometric verification. The Secure Multi-party Computation (SMC) problem was first introduced by Yao (1982) and extended by Goldreich et al (1987), and by many others (Goldwasser 1997).

Goldreich pointed out that solutions to specific problems should be developed and customized for efficiency reasons. Du and Atallah (2001a) presented a series of specific solutions to specific problems; e.g., privacy-preserving cooperative scientific computations, privacy-preserving database query, and privacy-preserving geometric computation. In their Privacy-Preserving Cooperative Scientific Computations (PPCSC) (Du & Atallah 2001b), they proposed a protocol between two parties to solve the problem $(A1+A2)x = b1 + b2$, where matrix A1 and vector b1 belong to party P1, matrix A2 and vector b2 belong to party P2. At the end of the protocol, both parties know the solution x while nobody knows the other party's private inputs.

In SIPPA, the private data exchange and information processing involves PPCSC. Specifically, biometric data is represented in the form of a matrix A_i ($i=1, 2$), and the (most significant) eigenvector weighted by the eigenvalue is represented by the vector b_i ($i=1, 2$). We will show in the later section that the solution vector x satisfying $(A1+A2)x = b1 + b2$ for PPCSC offers the boundary information for each party to estimate the distance between the eigenvectors of the data matrices of the two parties. This distance estimate then forms the basis for comparing the biometric data of both parties as well as the generation of helper data for reconstructing the source data. Further details about this will be discussed in a later section.

Our approach to tackle the problem of PPCSC is to employ homomorphic encryption and singular value decomposition (SVD) on the matrices of P1 and P2 to achieve privacy protection. This approach was discussed in our recent BTAS paper (Sy 2009a). For completeness, we will include the high level summary of the approach in a later section.

3. Secure Information Processing with Privacy Assurance (SIPPA)

In this research, the privacy model for biometrics can be formulated as below: Party P1 has some biometric data expressed in terms of a linearized vector D1. Party P2 has some linearized vector template D2 about a subject P3. The objective is for P1 to determine whether D1 and D2 are similar under the following conditions:

1. P1 and P2 do not reveal their private data to each other.
2. P1 and P2 both need to determine whether D1 and D2 are sufficiently similar.
3. If D1 and D2 are sufficiently similar, P2 will provide some helper data HD with a negligible overhead for P1 to reconstruct D2 using only D1 and HD.

Eigen-based approach (Turk & Pentland 1991) has been developed in the early 90s for people identification based on face biometrics. An important property of eigenface is to represent a linearized image as a covariance matrix with its corresponding eigenvectors as a set of linearly independent basis for capturing the variation of the image.

Covariance matrix of an image representation is symmetric. In fact, any linearized data vector D will yield a symmetric matrix out of $D \cdot D^T$. Let A1 and A2 be the proper transformation of some linearized data D1 and D2 through $D1 \cdot D1^T$ and $D2 \cdot D2^T$ respectively. Let $\{(\lambda_i, V_i) \mid i = 1, 2, \dots\}$ and $\{(\lambda_i, V_i) \mid i = 1, 2, \dots\}$ be the corresponding sets of eigenvalues and unity normalized eigenvectors for A1 and A2 respectively. For the sake of discussion and without the loss of generality, we will focus on only the largest Eigen components (λ_1, V_1) and (λ_2, V_2) . Let x be a vector such that $(A1+A2)x = \lambda_1 \cdot V_1 + \lambda_2 \cdot V_2$. By definition, $A1V_1 = \lambda_1 \cdot V_1$ and $A2V_2 = \lambda_2 \cdot V_2$. These relationships manifest an endomorphism mapping with the following three observations:

Observation 1: $\lambda_1 \cdot V_1$ and $\lambda_2 \cdot V_2$ are the transformation of the eigenvectors V_1 and V_2 through A1 and A2 respectively.

Observation 2: The resultant sum of the vectors $\lambda_1 \cdot V_1$ and $\lambda_2 \cdot V_2$ are the transformation of the vector x through $(A1+A2)$.

Observation 3: Vector x can be decomposed into components with V_i ($i=1,2$) as basis; i.e., $x = V_1 + \epsilon_1$ and $x = V_2 - \epsilon_2$; whereas ϵ_i ($i=1,2$) can be considered as an error/offset term accounting for the deviation of x from V_i ($i=1,2$).

The graphical interpretation of the observations just mentioned above is illustrated in figure 1 below:

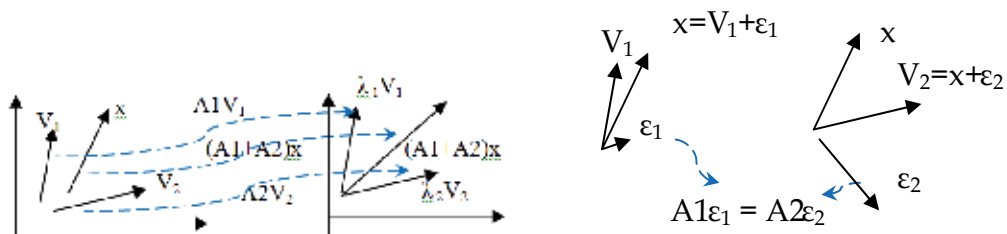


Fig. 1. Eigen-based Approach for SIPPA

As one could notice, x converges to the eigenvector as (λ_1, V_1) and (λ_2, V_2) converge to each other. This, together with observation 3, leads to the following trivial, yet important property:

Property 1: As (λ_1, V_1) and (λ_2, V_2) converge to each other, x converges to the eigenvector V_1 ($i=1,2$) and ε_i ($i=1,2$) converge to zero.

The key mathematical structure of SIPPA is the algebraic system of linear equations defined by $(A1+A2)x = \lambda_1 \cdot V_1 + \lambda_2 \cdot V_2$. These linear equations define the constraint relationship between a separator boundary and the eigenvectors of the two parties; whereas the solution to the algebraic system reveals the information needed to derive the lower bound distance of the norm deviation of the eigenvectors. If the lower bound is deemed acceptable, the eigenvalue and some scalar value would be sent by the server (P2) to the client (P1) as the helper data. Based on the helper data and D1, P1 can derive the norm deviation between the eigenvectors of P1 and P2 under the assumption of equidistance. Subsequently, P1 can derive a sufficiently good approximation of D2.

As one could note in property 1, it essentially states that one could infer the closeness between V_1 and V_2 through V_1 and x without knowing V_1 – the basis of SIPPA; where $(i j) = (1 2)$ or $(2 1)$. Furthermore, it can easily be shown that $A1\varepsilon_1 = A2\varepsilon_2$, or $D1 \cdot D1^T \cdot \varepsilon_1 = D2 \cdot D2^T \cdot \varepsilon_2$, which provides a convenient way in the SIPPA scheme to derive D_j when the scalar $D_j^T \cdot \varepsilon_j$ is known; where $x = V_1 + \varepsilon_1$ and $V_2 = x + \varepsilon_2$. In other words, SIPPA can easily facilitate “separation of duty” in the sense that data exchange/processing is only possible when both parties collaborate. Similarly, SIPPA separates the step for similarity comparison and that for retrieval. Therefore, the security principles “need-to-know” and “least privilege” can be implemented in the SIPPA environment.

3.1 SIPPA algorithm details

There are three major aspects of SIPPA:

1. Derivation of the eigenvalues and the corresponding unity normalized eigenvectors of the symmetric matrix representation of the data.
2. Derivation of a boundary vector separating the eigenvectors of the two parties, which is formulated as a two-party PPCSC secure computation.
3. Reconstruction of the source data based on the helper data composed of the eigenvalue and a scalar derived from the vector product between the transpose of the linearized source data vector and the boundary vector.

We will first outline the key steps of SIPPA, and then the secure computation protocol for PPCSC.

Let D_v and D_e be the sample and source linearized data respectively. The key steps of SIPPA are summarized below:

- Step 1.** Derive symmetric matrix representation of the data in the form of $D_v \cdot D_v^T$ ($= A1$) and $D_e \cdot D_e^T$ ($= A2$).
- Step 2.** Derive the eigenvalues and the corresponding unity normalized eigenvectors $\{(\lambda_i, V_i) \mid i=1,2,\dots\}$ of $D_v \cdot D_v^T$ and $\{(\lambda_i, V_i) \mid i=1,2,\dots\}$ of $D_e \cdot D_e^T$.
- Step 3.** Compute x such that $(A1+A2)x = \lambda_1 \cdot V_1 + \lambda_2 \cdot V_2$ based on our PPCSC (Sy 2009b) summarized in Fig. 2 below
- Step 4.** Derive the closeness between D_e and D_v via the minimum distance between V_1 and V_2 . The minimum distance between V_1 and V_2 is estimated via $|V_2 - x|$.
- Step 5.** If D_e and D_v are sufficiently close as measured by $|V_2 - x| <$ some pre-defined threshold, proceed to send the following helper data: λ_2 and $D_e^T \cdot x$.

Step 6. Derive $\underline{V}_2 = V_1 + k(x - V_1)$, and then the estimated source data $De' = \lambda^2_1 \cdot \underline{V}_2 / De^T \cdot x$; where $k = 2$ under the assumption of equidistance.

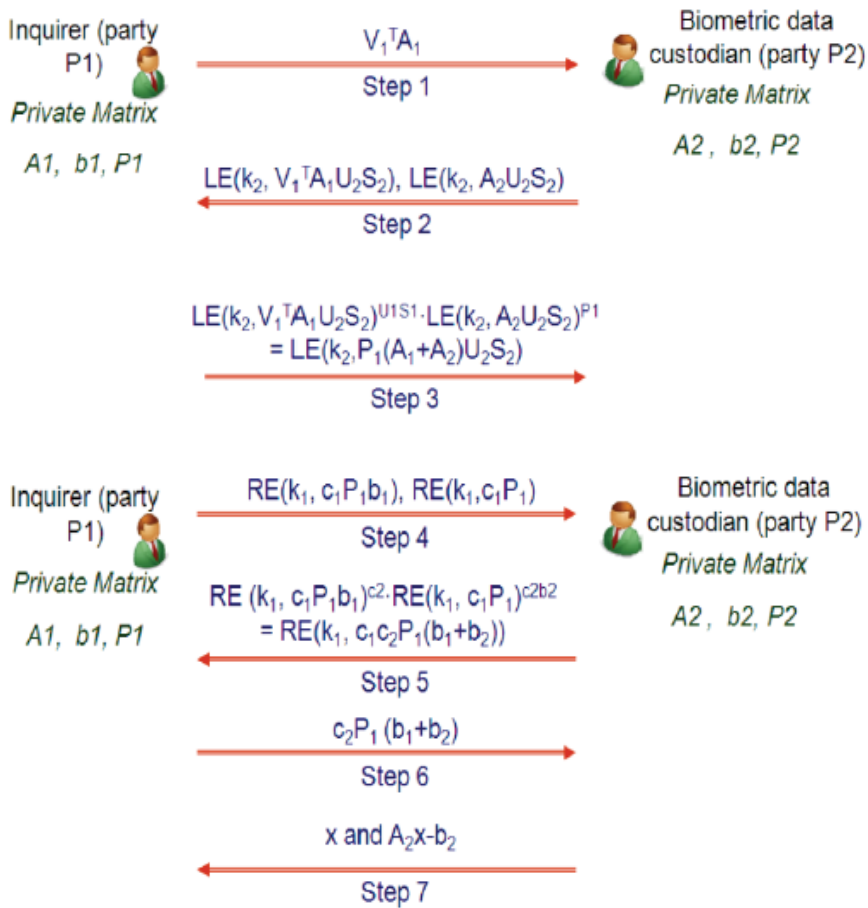


Fig. 2. PPCSC Illustration

The boundary vector x in step 3 will be derived under PPCSC. The basic idea behind the secure computation protocol for PPCSC required in step 3 for solving $(A_1 + A_2)x = b_1 + b_2$ is to solve $P_1(A_1 + A_2)P_2 y = P_1(b_1 + b_2)$; where (P_1, A_1, b_1) are private to P1, and (P_2, A_2, b_2) are private to P2. Note that even if P2 knows $P_1(A_1 + A_2)P_2$ and $P_1(b_1 + b_2)$, P2 can only derive y but could not know $A_1, b_1,$ and P_1 ; thus the privacy of $A_1, b_1,$ and P_1 for P1 is preserved. Once y is solved, P2 can derive $x = P_2 y$ and sends it to P1. Note that any adversary intercepting or sniffing from the network the value of x cannot derive De or Dv unless the adversary also has either (A_1, b_1, V_2) or (A_2, b_2, V_1) .

During the process of secure computation, A_1, b_1, A_2, b_2 are never exposed individually except $P_1(A_1 + A_2)P_2$ and $P_1(b_1 + b_2)$ for P2. Therefore, it is information-theoretic secure; i.e., the privacy of (A_1, b_1) for P1, and the privacy of (A_2, b_2) for P2, is guaranteed. Readers interested in the step-by-step details on the secure computation mechanism for solving x in $(A_1 + A_2)x = b_1 + b_2$ are referred to our articles elsewhere (Sy 2009b) (Sy 2009c).

4. Experimental studies on SIPPA parameters & Usability for biometric security

For proof-of-concept, we conducted experimental studies to better understand the potential of SIPPA for biometric applications. This experimental study consists of three parts. The first part is a simulation study targeting the specific parameters to investigate their inter-relationship. This experimental study aimed at tackling the following two questions:

1. How is the quality of the estimate on the closeness between D_e and D_v affected by the dimension of x ?
2. How is the closeness between D_e and D_e' (estimated D_e) affected by (a) $|V_{2_1-x}|$, and (b) dimension of x ?

The second part is an application of SIPPA to the private reconstruction of digital images, and the third part is the private reconstruction of the voiceprint.

4.1 Experimental study part 1: SIPPA parameters

In the simulation study, we generated 5 test data sets categorized by different dimensions. The vector dimensions in these five data sets are 5, 10, 20, 40, and 60 respectively. In each data set, 10 pairs of client (D_v)/server (D_e) vectors are generated, thus totaling 50 pairs for all dimensions. Figure 3 shows the normalized data difference $|D_e - D_v| / \text{Dim}$ in y-axis and the normalized eigenvector difference $|V_{2_1} - V_{1_1}| / \text{Dim}$ in x-axis. Figure 4 shows graphically the relationship between the normalized deviation measure $|V_{2_1} - V_{1_1}| / \text{Dim}$ and $|V_{2_1-x}| / \text{Dim}$.

Figure 3 shows an approximately linear relationship between the deviation of the client (D_v)/server (D_e) and that of the corresponding eigenvectors. The degree of closeness between two sources of data is reflected by the closeness between the corresponding eigenvectors. In other words, the client with D_v and the corresponding eigenvector V_{1_1} will be able to deduce the degree of closeness between D_v and D_e if the similarity distance as measured by 2-norm $|V_{2_1} - V_{1_1}|$ is known. And if the degree of closeness between D_v and D_e is known, it provides a basis for reconstructing D_e .

Figure 4 also shows an approximately linear relationship in relatively lower vector dimension (≤ 20) between the degree of closeness of the eigenvectors (of the client and server), and that of the server (thus the client in an inversely proportional sense) and the boundary vector x . Consequently, if the dimension of the data vector is relatively low (i.e., ≤ 20), then the degree of closeness as measured by the 2-norm $|V_{1_1-x}|$ or $|V_{2_1-x}|$ can act as a predictor for $|V_{2_1} - V_{1_1}|$.

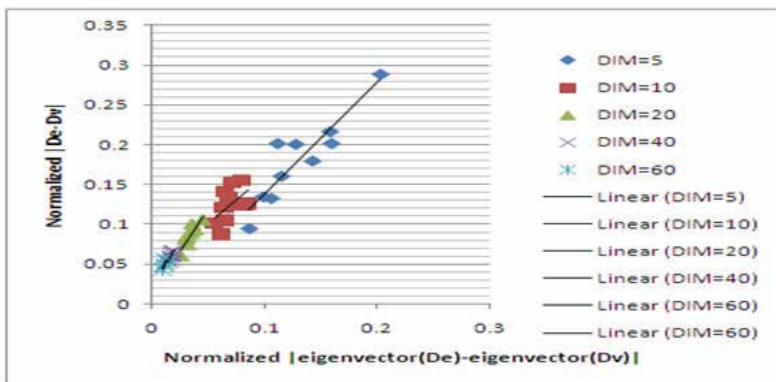


Fig. 3. Data deviation vs. eigenvector deviation

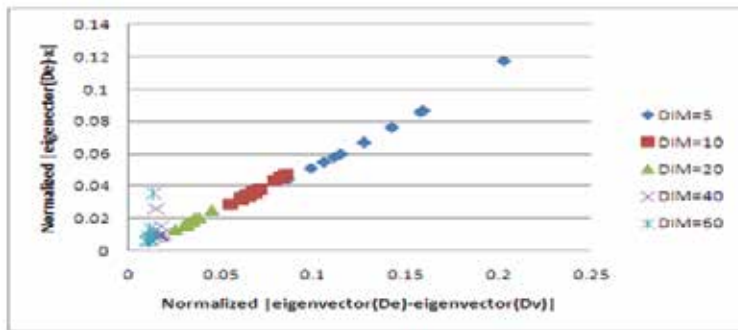


Fig. 4. Goodness of fit of the difference estimator

Based on the result of the simulation study, we chose the data dimension to be 10 in the second part of our experimental study.

4.2 Experimental study part 2: Biometric face image

In the SIPPA experimentation with digital images, three doll faces and one real person image containing biometric face information were used. All images are stored in PGM (Portable Gray Map) format. In the experimentation using doll faces, Figure 5 shows the result of the experimentation. The first row shows two black-and-white original images of two Barbies. The first column shows the reference sample images consisting of two other Barbies, and a generic blonde hair doll. The resolution of each image is 64x85 with 255 gray-level.

Altogether six doll faces are reconstructed from every combination pair of a reference image and a sample image. In the magnified version of Figure 5, the best reconstruction is from Barbie-to-Barbie, while the worst reconstructions are the two on the last rows.

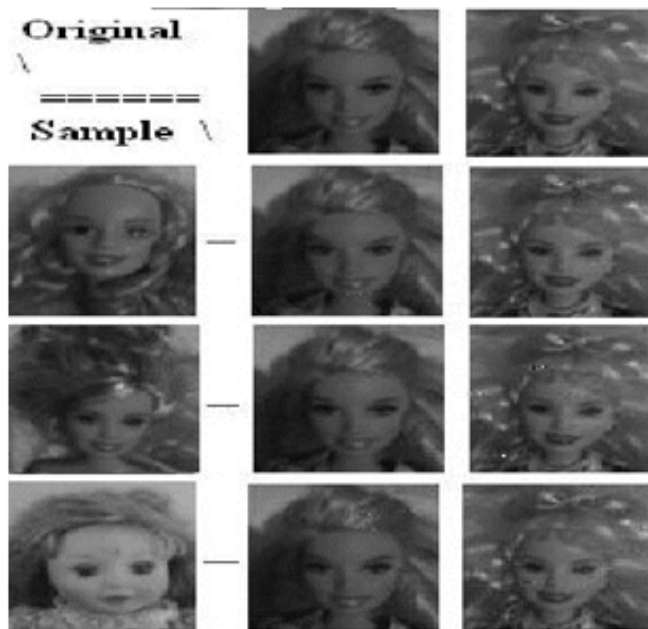


Fig. 5. SIPPA application to doll's face

In the experimentation using real human images, the server side utilizes a black-and-white image with 255 gray levels (Figure 6). The resolution of the image is 256x192. The linearization of the image results in a 49152 (=256x192) x 1 vector. This is about 9 times larger in comparing to the linearization of the doll images, which have a vector size of 5440 (=64x85) x 1. Obviously the dimensions of the linearized vector in both cases are beyond the optimal performance range of SIPPA. As such, the 49152x1 (or 5440x1) linearized image vector is split into multiple 10x1 vectors. SIPPA is then applied iteratively to reconstruct the original image.

During the reconstruction, two parties – referred to as client and server – will participate in a SIPPA session. The client has an image, referred to as a sample image. The server has an image, referred to as a source image. In the SIPPA session, each party takes a portion of the linearized image sequentially in the form of a 10x1 vector to construct a symmetric matrix, and derives the eigenvalue/eigenvector of the matrix. Then both parties participate in a secure computation protocol for PPCSC as described in step 3 of SIPPA in the previous section.

The outcome of PPCSC is the boundary vector x . The client party then uses the boundary vector x and the client side 10x1 linearized vector of the sample image to reconstruct the server side 10x1 linearized vector of the source image. The only information that the server provides is the helper data composed of the eigenvalue of the server side 10x1 linearized source image and a scalar. The original 10x1 linearized source image is never shared with the client party. Once the reconstruction of the 10x1 server side linearized source image is completed, the SIPPA process repeats for the next block of the 10x1 linearized vector until all 10x1 image blocks are processed. The entire source image is then reconstructed by the client party using the estimated 10x1 source image blocks.



Fig. 6. Source image



Fig. 7. Similar sample image

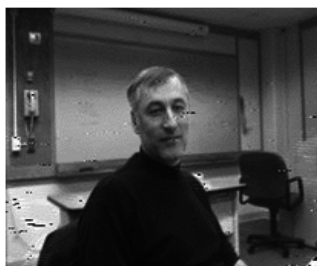


Fig. 8. Reconstruction using similar image



Fig. 9. Reconstruction using dissimilar image

The SIPPA reconstruction of the entire human source image is conducted twice in this experiment. In the first trial, the client side provides a sample image shown in Figure 7 that is sufficiently similar to the server side source image as shown in Figure 6. All images have a gray scale of 255 and have the same resolution 256x192.

The outcome of the reconstruction by applying SIPPA based on the server side source image (Figure 6) and the client side sample image (Figure 7) is shown in Figure 8. During the reconstruction, the source image is never shared with any party except the corresponding information used in the SIPPA processing.

During the second trial, a sample image of a different human subject (not Figure 6 or 7) is used on the client side for SIPPA. The image of the other human subject is not shown due to the restriction on the human subject clearance. The outcome of the reconstruction is shown in Figure 9, which shows a poorer quality in comparison to Figure 8.

By visual inspection and using Figure 6 as a reference, the face biometrics in Figure 8 is preserved better than that in Figure 9 - when the client side sample image is closer to that of the server side source image (Figure 6).

4.3 Experimental study part 3: Utilizing SIPPA in voice biometrics

In this part, a study of SIPPA utilizing voiceprints is conducted using the speaker verification system reported in our paper elsewhere (Sy 2009b). The objective of the experimental study is to study SIPPA in terms of its ability to reconstruct voiceprints that can cause the speaker verification system to behave in the same way as if the original voiceprints of its users are applied to the system. Twenty four speakers of different native languages participated in the experimental study. Altogether 118 different viable True User attempts and 101 different viable Impostor attempts were used to evaluate the system as described in the following experiments.

After filtering the instances of FTE (Failure to Enroll) and FTA (Failure to Acquire) due to noise introduced by phone devices and background environment, each voiceprint for verification and the enrolled voice template are used by SIPPA to reconstruct the voiceprint of a speaker for the speaker verification system; i.e., the speaker does not present his/her voiceprint to the speaker verification system. The distance between the enrolled voice template and the voiceprint reconstructed by SIPPA - abbreviated by $KL\text{-dist}(\text{enroll}, \text{sippa}(\text{VectorDim}))$, is computed. VectorDim specifies the SIPPA vector dimension. For the control experiment, the distance between the enrolled voice template and the verification voiceprint - abbreviated by $KL\text{-dist}(\text{enroll}, \text{verify})$ is also computed. Kullback-Leibler divergence is the distance function used to calculate a similarity score in this case, as described in our paper elsewhere (Sy 2009b).

The system behavior characterized by ROC using original speaker voiceprints and SIPPA reconstructed voiceprints are shown in Figure 10 for a comparison purpose. In this experimentation we deliberately set the *pre-defined threshold* as stated in step 5 of the SIPPA algorithm in the previous section to be infinity. In doing so, the "usability" of SIPPA is the highest, while the performance is expected to be the lowest when compared with the cases where the pre-defined threshold is used to filter the cases where the source and sample data are significantly different. In other words, all the voiceprints reconstructed by SIPPA were used in the derivation of the ROC irrespective to the closeness of a reconstructed voiceprint and its reference voiceprint.

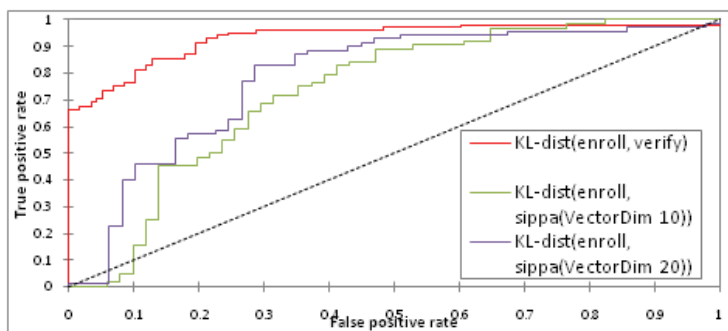


Fig. 10. ROC with threshold as infinity

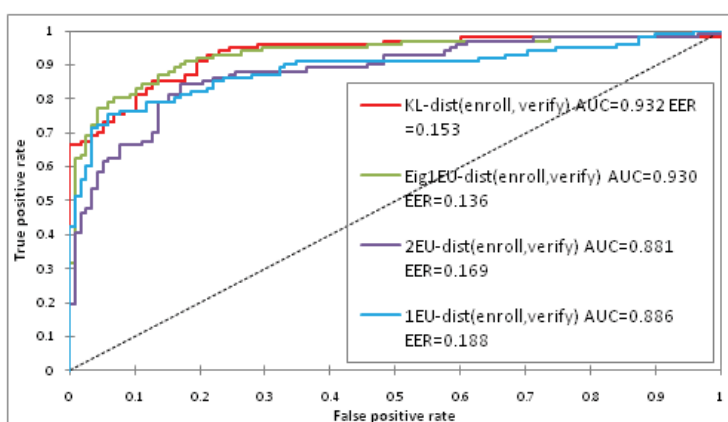


Fig. 11. ROC using original enroll voiceprint

4.3.1 Other Distance Functions more suitable for SIPPA reconstructions

Since SIPPA is Eigen-Based and utilizes Euclidean Distance of Eigenvectors to estimate the similarity between datasets, we tried to determine if an Eigen, Euclidean based distance function may be used to yield similar performance to the KL-Distance mentioned above. If such a distance function can be found, SIPPA reconstructions may be used more successfully in voice biometrics. The essential problem here is obtaining a distance function that performs as-good-as/better than the KL-Distance function used in our voice system described elsewhere (Sy 2009b), while also not being adversely affected by “noise” that may be generated by SIPPA reconstructions of the Verification voiceprint.

As described elsewhere (Sy 2009b), the voiceprints used by our system are created by extracting Mel-Spectrum features from a wav file (using 8KHz sampling rate) to derive a 20X1 mean vector and a 20X20 covariance matrix of the corresponding multivariate Gaussian model. The 20X1 mean vector and 20X20 Covariance matrix are the basis used in all the varied Eigen, Euclidean based distance functions that we evaluated. Their performance relative to the KL-Distance described above is characterized by ROC curves shown in Figure 11. Their performance with SIPPA reconstructions, where SIPPA threshold is set to infinity is plotted in Figure 12. Table-2 describes these distance functions. All ROC curves were generated by utilizing the same dataset.

As observed in Figure 11, the voice system using Eig1EU-dist distance function performs as good as that using the KL-dist function on the same dataset. The Eig1EU-dist function leads to an EER of 0.136 with an AUC (Area Under Curve) of 0.930, while the KL-dist function leads to an EER of 0.153 with an AUC of 0.932. What is interesting about the Eig1EU-dist function is the fact that even when utilizing SIPPA reconstructions (SIPPA threshold at infinity) of Verification Voiceprints, the performance of the voice biometric system is not affected in too adverse a manner as seen in Figure 12. The system configured with Eig1EU-dist distance function performs as good as one configured with the KL-Distance function without SIPPA i.e. KL-distance comparing Original Enroll and Sample Voiceprints.

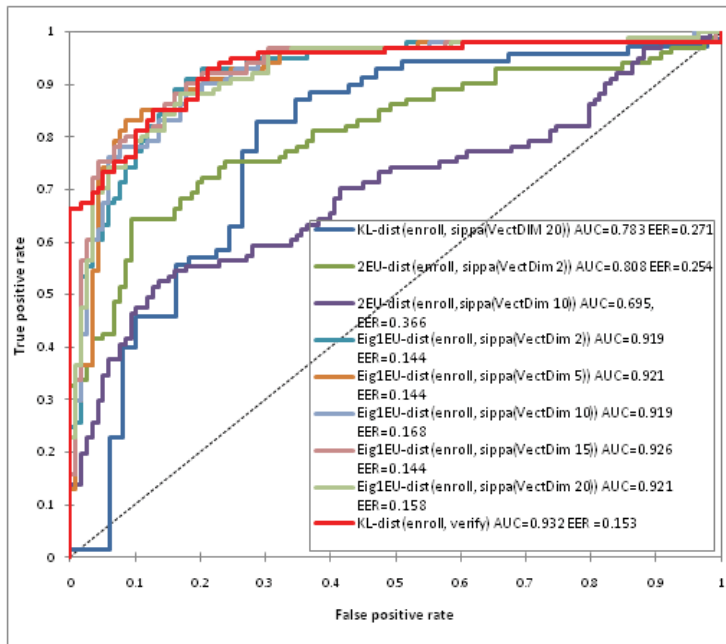


Fig. 12. ROC using original enroll and SIPPA

4.3.2 SIPPA privacy protection & thresholds

When a person interacts with a voice verification system, the privacy of the person and his/her voiceprint can be protected through SIPPA. Specifically, the user may engage in a SIPPA session as a SIPPA server, and chooses to send a close approximation of a sample verification voiceprint only if the verification system has a reference voiceprint of the person. The person can choose to set her/his SIPPA threshold to infinity. However, this might not provide effective privacy protection across different biometric template standards and modalities. Setting a SIPPA threshold provides for finer control. The basic idea is to utilize the one piece of information that the SIPPA server knows about the SIPPA client - Vector X. If one can devise a distance function that provides an ROC similar to the ROC using a KL-distance function for comparing the *Original Enroll* and *Original Verify* voiceprints, the SIPPA server can set an effective threshold which can be customized intelligently for varied levels of privacy protection. The essential question here is, by knowing the vector X, how effectively (based on some distance function) can the SIPPA

Distance Function	Description
KL-dist	Distance function described elsewhere (Sy 2009b)
Eig1EU-dist	<p>Similarity score calculated by:</p> <p>Obtaining a symmetric matrix by multiplying the 20X1 multivariate mean vector (stored in each of the voiceprints being compared, as detailed earlier) with its transpose. Symmetric Matrix (Me) is obtained from the enrolled voiceprint and Symmetric Matrix (Mv) is obtained from the verify voiceprint. Obtaining 20 symmetric matrices by multiplying each column of the 20X20 covariance matrix (stored in each of the voiceprints being compared, as detailed earlier) with its transpose. 20 Symmetric Matrices (Ce1 ... Ce20) are obtained from the enrolled voiceprint and a further 20 Symmetric Matrices (Cv1 ... Cv20) are obtained from the verify voiceprint.</p> <p>Obtaining the largest Eigenvalue and its corresponding unity-normal Eigenvector for the Symmetric Matrices Me,Mv,Ce1...Ce20,Cv1...Cv20 – producing Eigenvectors MEe,MEv,CEe1...CEe20,CEv1...CEv20.</p> <p>Obtaining 2-Norm Euclidean distances for $MEe-MEv$, $CEe1-CEv1 \dots CEe1-CEv1$</p> <p>The Similarity score (i.e. the output of this distance function) is the Geometric Mean of the 21 2-Norm Euclidean distance values calculated in the previous step.</p>
2EU-dist	<p>Similarity score calculated by:</p> <p>Obtaining the 2-Norm Euclidean distance(2EM) between the 20X1 Enroll-Voiceprint mean vector and the 20X1 Verify-Voiceprint mean vector.</p> <p>Obtaining two 400X1 vectors by converting the 20X20 covariance matrices (in each of the Voiceprints being compared) into a vector by essentially copying values from the first column of the matrix into the first 20 values of the vector and then obtaining values from the second column of the covariance matrix into the next 20 values of the vector and so on. One 400X1 vector (Ve) is obtained from the Enroll Voiceprint and one 400X1 vector (Vv) is obtained from the Verification Voiceprint.</p> <p>Obtaining the 2-Norm Euclidean distance(2EC) for $Ve - Vv$.</p> <p>The Similarity score i.e. the output of this distance function is the Geometric Mean of 2EM and 2EC</p>
1EU-dist	<p>Similarity score calculated by:</p> <p>Obtaining the 1-Norm Euclidean distance(1EM) between the 20X1 Enroll-Voiceprint mean vector and the 20X1 Verify-Voiceprint mean vector.</p> <p>Obtaining two 400X1 vectors by converting the 20X20 covariance matrices (in each of the Voiceprints being compared) into a vector by essentially copying values from the first column of the matrix into the first 20 values of the vector and then obtaining values from the second column of the covariance matrix into the next 20 values of the vector and so on. One 400X1 vector (Ve) is obtained from the Enroll Voiceprint and one 400X1 vector (Vv) is obtained from the Verification Voiceprint.</p> <p>Obtaining the 1-Norm Euclidean distance (1EC) for $Ve - Vv$.</p> <p>The Similarity score (i.e. the output of this distance function) is the Geometric Mean of 1EM and 1EC</p>

Table 2. Tested distance functions described.

server/client determine if both voiceprints being compared are from the same person, or if they are from two different persons.

To illustrate, let's consider that a user would like to verify his voiceprint with a verification system. The user assumes the role of a SIPPA server while the verification system assumes the role of a SIPPA client. Using Mel spectrum as discussed (Sy 2009b), each voiceprint is a 420×1 vector. 21 SIPPA sessions are required with each session assuming the responsibility for reconstructing one 20×1 vector. Each of the SIPPA server sessions obtains the vector X , and computes the 2-norm Euclidean distance between their own eigenvector and vector X . The user at this point can utilize these 21 2-norm Euclidean distance scores to determine if her/his voiceprint is sufficiently similar to the reference voiceprint of the verification system (SIPPA client). If they are sufficiently similar, indicating the voiceprints may be from the same person, the user may opt to send helper data to the SIPPA client; i.e. the verification system.

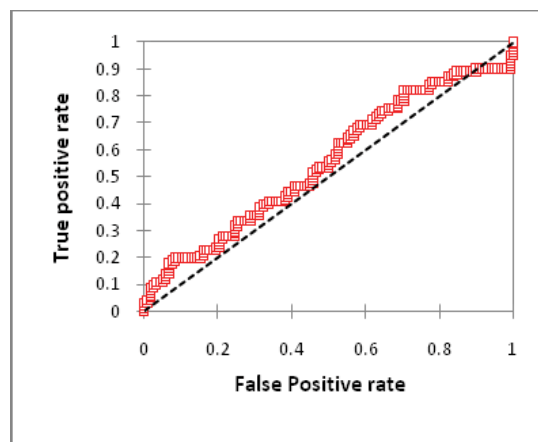


Fig. 13. Max Function for Calculating Distance

An ROC curve can be used to examine the quality of a distance function which produces a Similarity Score based on these 21 2-norm Euclidean distance scores, or in general the 21 SIPPA Server Eigenvectors and the 21 X vectors. We can generate the ROC by varying the SIPPA threshold. Specifically, the SIPPA threshold, in conjunction with Similarity Score produced by a distance function, can be used to decide whether the two voiceprints may be from the same person, or two different persons; thus an instance of true/false positive. If SIPPA server relies on a distance function that merely uses the maximum of the set of 21 2-norm-Euclidean distance scores as a distance value, and, say, the distance function results in a ROC as in Fig. 13, then the voice system will have behaved no better than flipping a coin to determine whether the voiceprints are from the same person.

The performance of various distance functions producing a similarity score for the two voice prints being compared utilizing only the SIPPA server Eigenvectors and X vectors are shown in Figure 14. Table 3 describes these distance functions. To reiterate, all ROC curves in this chapter were generated by utilizing the same dataset.

4.3.3 Dual SIPPA, verification-engine threshold

The ROC curves at Figure 15 show the behavior of the Voice-Biometric system described elsewhere (Sy 2009b) when two different distance functions were used; namely, a two-

threshold SIPPA based distance function and KL-distance function. In this case, the Voice-Biometric System is the SIPPA client. The ROC of the system employing the two- threshold SIPPA distance function is calculated in the following manner: Let's consider that \mathbf{n} denotes the set of possible SIPPA thresholds we would like to evaluate the system on and \mathbf{m} denotes the set of Voice-Biometric system thresholds we would like to evaluate the system on. The total possible combinations of thresholds are $\mathbf{m} \times \mathbf{n}$. Hence the ROC will contain $\mathbf{m} \times \mathbf{n}$ data points. For each possible combination of thresholds, either a True-Negative, or a False-Negative, will be counted if the SIPPA threshold leads to not sending helper data. If SIPPA server (i.e., the user) does send helper data according to the (SIPPA) threshold, then the user voiceprint will be reconstructed by the SIPPA server (i.e., the system) and compared against the template. The comparison will then yield a count of either True-Negative, True-Positive, False-Positive or False-Negative count.

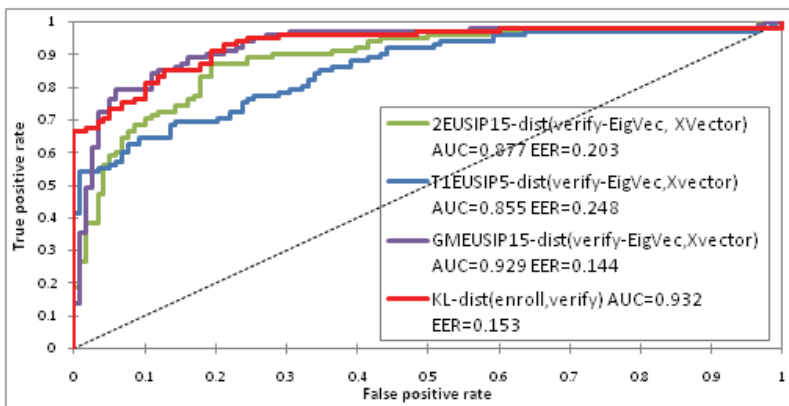


Fig. 14. ROC of different distance functions.

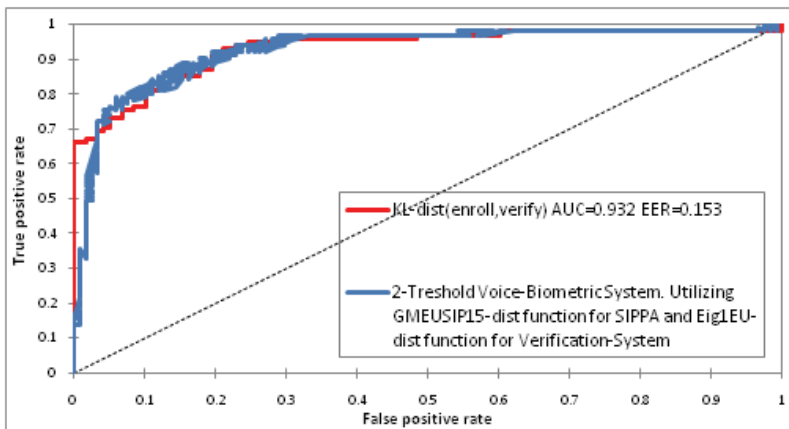


Fig. 15. ROC of 2- threshold with SIPPA

Deducing from the plots in Figure 15, adopting the two-threshold SIPPA based distance function allows the system to perform as good as one based on KL-distance. But the two-threshold SIPPA based distance function allows privacy preserving authentication without

Distance Function	Description
KL-dist	Distance function described elsewhere (Sy 2009b)
2EUSIP15-dist	<p>Similarity score calculated by:</p> <ol style="list-style-type: none"> 1. SIPPA vector dimension is 15X1; therefore 28 SIPPA Sessions are required to process the two Voiceprints. Generating 28 15X1 SIPPA Server Eigenvectors(SE1...SE28) and 28 15x1 X vectors (X1...X28) 2. 2-NormEuclideanDistance is obtained for $SE1-X1 \dots SE28-X28$ producing 28 Euclidean distance values. 3. Similarity score is obtained by calculating the mean of the 28 Euclidean distance values calculated above.
T1EUSIP5-dist	<p>Similarity score calculated by:</p> <ol style="list-style-type: none"> 1. SIPPA vector dimension is 5X1; therefore 84 SIPPA Sessions are required to process the two Voiceprints. Generating 84 5X1 SIPPA Server Eigenvectors(SE1...SE84) and 84 5x1 X vectors (X1...X84) 2. 2-NormEuclideanDistance is obtained for $SE1-X1 \dots SE84-X84$ producing 84 Euclidean distance values. 3. Similarity score is obtained by calculating the t-value utilizing the one sample t-test with a population mean of 0.
GMEUSIP15-dist	<p>Similarity score calculated by:</p> <ol style="list-style-type: none"> 1. SIPPA vector dimension is 15X1; therefore 28 SIPPA Sessions are required to process the two Voiceprints. Generating 28 15X1 SIPPA Server Eigenvectors(SE1...SE28) and 28 15x1 X vectors (X1...X28) 2. 2-NormEuclideanDistance is obtained for $SE1-X1 \dots SE28-X28$ producing 28 Euclidean distance values. 3. Similarity score is obtained by calculating the geometric-mean of the 28 Euclidean distance values calculated above.

Table 3. Details of SIPPA distance functions plotted in Figure 14

compromising accuracy! Given the promise of this result, it will be worthwhile to conduct a large scale study to quantify the characteristics of various distance functions described in this section, and to observe how SIPPA behaves with respect to each of them.

5. Standard-based parallel SIPPA design and implementation

The experimental study in the previous section is focused on the accuracy performance. Currently SIPPA would need parallel processing, referred to as Parallel SIPPA, in order to deliver real time computation performance on image processing. The essential idea behind Parallel SIPPA is the creation of multiple SIPPA-Client and SIPPA-Server instances. The parallel SIPPA architecture is summarized in Figure 16. SIPPA realized under parallel computing can produce faster results. In parallel SIPPA, a Server Application (SA) initializes multiple instances of SIPPA-Server in different physical/virtual machines. As SIPPA-Server instances are created, their contact information is added to a queue maintained by SA. In dealing with digital images in PGM/PPM file format, SA linearizes the file into multiple

10X1 vectors where each vector is used to produce a symmetric matrix described in step 1 of the SIPPA algorithm reported in the previous section.

All SIPPA-Server instances have access to any of these symmetric matrices based on a matrix's unique id. When the Client Application (CA) starts, it takes possession of a sample client PGM/PPM image and linearizes the file into multiple 10X1 vectors; where each vector is used to produce a symmetric matrix in a similar fashion as the server side.

After CA contacts SA, SA de-queues an element and sends the contact information of SIPPA-Server instance to CA. CA continuously creates a new SIPPA-Client instance on available SIPPA-Server instances until all 10X1 vectors have been assigned resulting in parallel processing. CA provides every SIPPA-Client instance with an array of symmetric matrices and a chosen SIPPA-Server instance's contact information. This SIPPA-Client instance and its assigned SIPPA-server instance now engage in a sequential SIPPA session, processing each symmetric matrix at a time. As the SIPPA-Server/Client pairs finish processing their allotted set of matrices, the SIPPA-Server instances are added back to SA's queue and the SIPPA-Client instance is destroyed.

The reconstructed 10X1 vector produced by each SIPPA-client/SIPPA-Server pair is sent back to CA. When the entire set of 10X1 vectors is completed, all these vectors are combined to reconstruct an approximation of the server side image by CA.

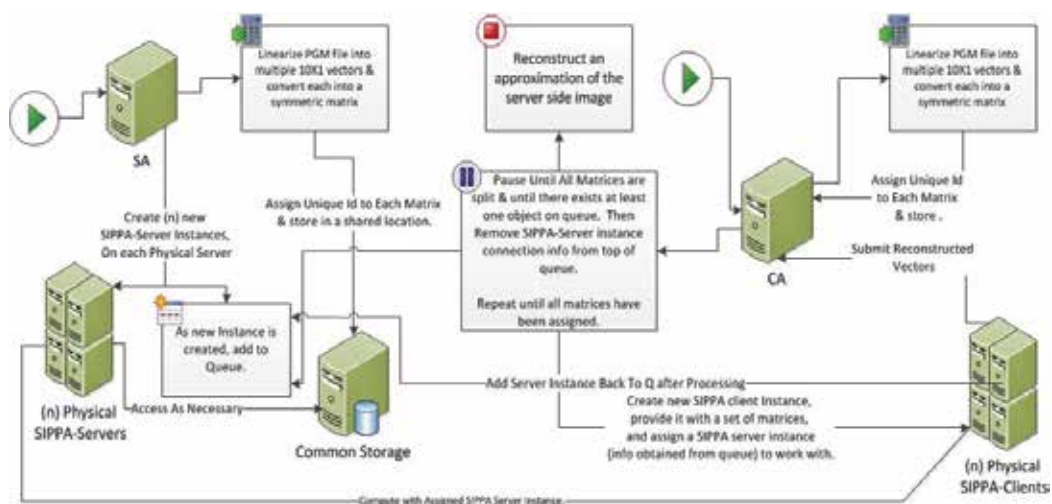


Fig. 16. Parallel SIPPA system delineation

5.1 Parallel SIPPA experimental results

In this experimental study we try to determine the optimal configuration for Parallel SIPPA i.e. the number of physical/virtual machines that need to be assigned for SIPPA-Servers and SIPPA-Clients respectively, and the number of SIPPA-Server instances that need to be created initially by SA (Server Application).

Performance of the SIPPA system is highly dependent on the degree of parallelism that can be obtained with the available hardware under which SIPPA operates. In our tests, we experiment with various platforms, both sequential and parallel; to pinpoint increases in performance as well as where our system is bottle-necked. By sequential we mean one instance of SIPPA-Server and SIPPA-Client runs at any given time in the entire system, and

the 10X1 vectors are processed one at a time, whereas by parallel we mean multiple instances of SIPPA-Server and SIPPA-Client run at any given time processing multiple 10X1 vectors at any given time. Our testing configurations included: Single Core virtual machines (VM), Dual Core physical machines (DC), and a Quad Core machine (QC).

To obtain 90 - 100% CPU utilization on all physical/virtual machines (both server and client), a proper client to server ratio must be applied. This is because SIPPA-client instances have an additional workload when compared with SIPPA-Server instances; i.e. they obtain their matrices dynamically from CA and then store them locally for processing. This additional network I/O coupled with the fact that they also have to reconstruct the server side matrix makes the Client Machines consume 6 times more resources than a Server machine. With a naïve parallel approach of using one physical/virtual machine as a server for every physical/virtual client machine, performance is improved in comparison to a sequential SIPPA approach. As we increase the client to server ratio by decreasing the amount of servers, the servers are more efficiently used which results in a dramatic decrease in completion time and full utilization of the CPU.

The data shows where our performance increase lies and where possible bottlenecks are. As seen in Figure 17, even when parallel SIPPA is run on a single quad core machine, a performance increase of 1600% is obtained when compared with using Two Dual Core machines engaging in SIPPA sequentially. Also notable is the performance increase by switching from the naïve parallel approach to an optimized parallel approach. With the VMs we had an increase of 75% in performance when comparing a naïve parallel approach to an optimized parallel approach. With the large data set experiment on the physical machines, we noticed an increase of 350%.

During Parallel SIPPA startup, there is an initialization and job distribution phase; this seems to take anywhere between 30 and 60 seconds depending on the size of image. Due to this, processing small images with a resolution 256x192 by parallel SIPPA do not seem to benefit much from increasing the number of physical/virtual machines within a certain setup.

In parallel SIPPA when an optimized setup is doubled and presented with an image of high resolution, we notice a reduction of processing time by about 40%; i.e., when the total number of machines is doubled from 8 (# 7) to 16 (# 6), the processing time for a 640X480 image decreases from 26 min to 15 min. We observe a similar improvement in the case with virtual machines (# 10 and # 9). Doubling of resources reduces processing time by approximately 40%. This indicates that given a reasonable size cluster of professional level servers, most of the typical biometric images could be processed in a Parallel SIPPA framework within a reasonable amount of time.

5.2 BioAPI Standard based SIPPA Implementation

When job requests are presented to Parallel SIPPA (PSIPPA), PSIPPA can achieve near real time performance as shown in the experiment results in the previous section. While the near real time results are noteworthy, our aim is to expand PSIPPA towards a true interoperable system. To achieve this we integrate PSIPPA into a BioAPI system which results in *BioSIPPA*. By integrating PSIPPA into a BioAPI 2.0 framework (May 2006), PSIPPA is made available in an interoperable environment allowing multiple users on various platforms to access PSIPPA through standard based service component. In order for *BioSIPPA* to handle simultaneous job requests from different users, we also investigate a Slice Based Architecture (SBA) design to optimize the resource utilization of *BioSIPPA*.

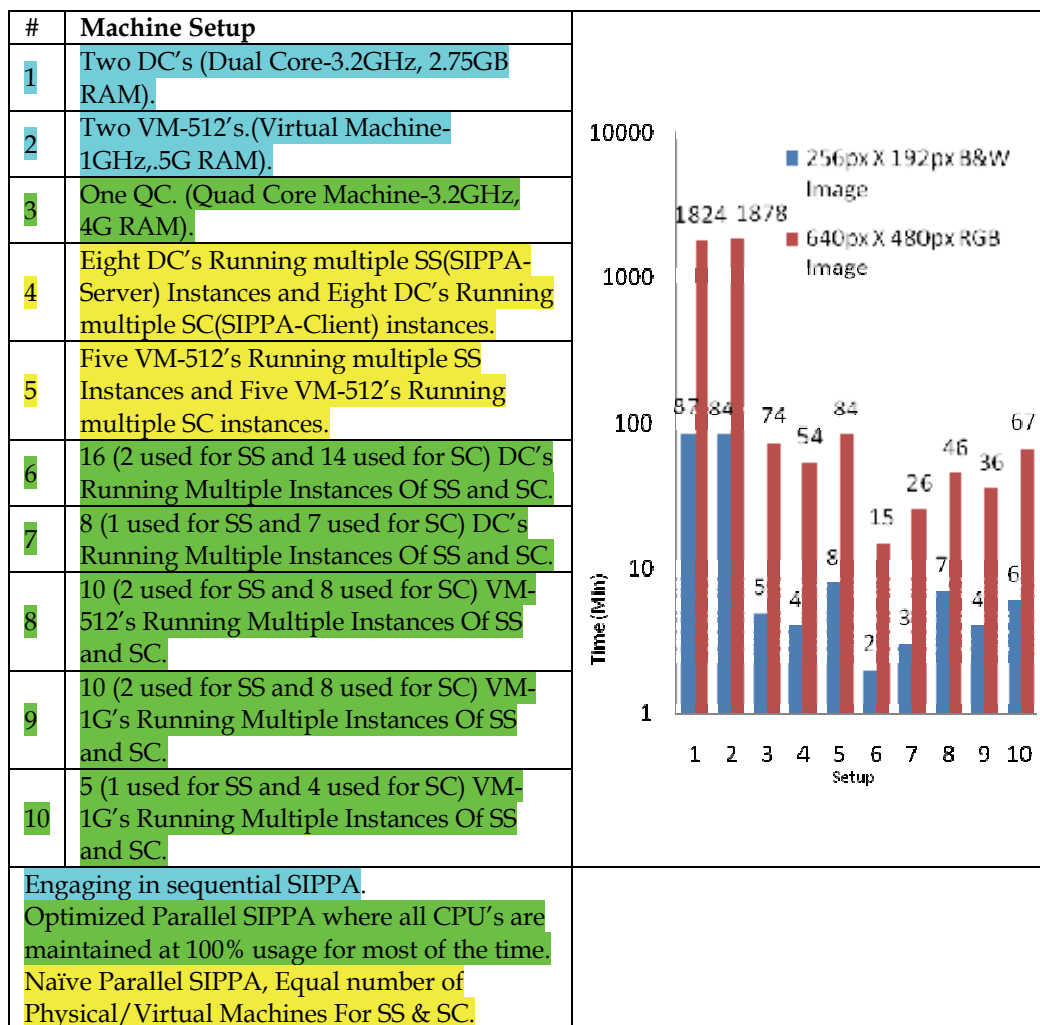


Fig. 17. Parallel SIPPA processing time performance

At the BioAPI level, realizing PSIPPA as a BioAPI service component involves a design and structural definition for various BioAPI framework components. At the minimum, we need to determine the appropriate level of PSIPPA instantiation in the BioAPI framework, and to define the corresponding framework components including Biometric Service Providers (BSP), Biometric Function Providers (BFP), and various BioAPI Units managed by the BSP or BFP such as Processing Algorithm Units (PAU) or Archive Units (AU). As defined in the BioAPI 2.0 standard, the PSIPPA schema and access interfaces are registered with the BioAPI component registry. This allows end users to search for our service.

Our targeted SBA-BioSIPPA system is composed of various elements that are the basis for its interoperable environment. Specific to biometric system development, the most atomic element in the SBA proposed in this project is *resource*. A resource could be a Java functional method exposed for an external call via Java RMI, a biometric software/hardware interface or a software/hardware processing unit for vector processing. A collection of resources for

the purpose of completing a specific task (e.g., enrollment) is referred to as an *aggregate*. An aggregate only defines the set of resources; it does not define how the resources of an aggregate should be used. The concept of a service can be thought of as one or more aggregates with additional information on how the resources may be used to accomplish a specific task. Lastly, a service *slice* is essentially one or more aggregates provisioned by a slice manager. A slice manager coordinates how these aggregates may be utilized; e.g., by whom, for how long, under what pre-conditions or constraints on bandwidth, memory, CPU, or storage size. In the application of SBA-BioSIPPA, the slice manager drives the provisioning of slices which creates the necessary element for a true parallel and interoperable environment.

An objective of SBA-BioSIPPA is to expose PSIPPA as an interoperable service component for a BioAPI 2.0 framework (May 2006). As a long term goal, we also aim for an easy adaptation of the service component for the standards developed for other application domains; e.g., ISO 19092:2008. Our strategy towards staging the PSIPPA service component for flexibility and adaptation is to employ RMI technology, which provides a lightweight, multithreaded, reliable communication bridge between the BioAPI and PSIPPA systems.

As in most biometric systems, SFA-BioSIPPA utilizes various resources similar to a face recognition system that has a Sensor Unit, a Matching Unit, and an Archive Unit. In SFA-BioSIPPA each necessary resource may be supplied by a different vendor and we provide a means for an end user to choose among the various resources.

To illustrate the utilization of SFA-BioSIPPA, consider a user A requests for privacy preserving verification using SIPPA. User A will interact with the Slice Manager, which queries the Component Manager. The slice manager then displays what resources are available to execute the job request. User A, with the assistance of the Slice Manager, chooses the appropriate resources for the job and sends a request for a slice allocation. The request is recorded in an internal registry within the Slice Manager and then sent to the Component Manager. The Component Manager accesses the Resource Management Database (RMDB) and acquires the details of resources specific to the slice request. The resources are accessed through their BSP, at which point the slice is in its initiation stage.

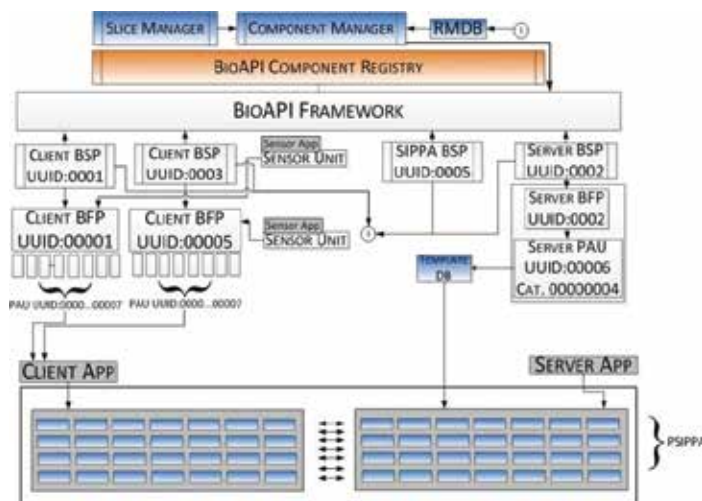


Fig. 18. SIPPA Deployment for BioAPI Framework

When a resource is accessed through its BSP, the BSP will initiate execution and immediately update a record in the RMDB incrementing the number of jobs executing on the particular resource. On the termination of a job on each resource, the resource's BSP will again update the database decrementing the number of jobs currently running on the resource. Any changes in the RMDB will allow future queries from the Slice Manager to reflect a near real time status of the system. Note that SFA-BioSIPPA aims at a standard based framework to facilitate not only interoperability, but also a scalable computation environment for PSIPPA where additional computing resources for SIPPA could be easily added to the framework via the registration of a RMI client with remote RMI server acting as a SIPPA client or SIPPA server. This is essentially a kind of cloud computing model to harness the computing power from multiple sources for PSIPPA. In this cloud computing model one must be careful in the design of SFA-BioSIPPA in regard to the risk in information leak and collusion when multiple parties are involved in PSIPPA. An alternative to this without the same degree of the risk just mentioned is MPSIPPA (Massive Parallel SIPPA) that relies on the GPU technology.

5.3 MPSIPPA

SIPPA, a client server mechanism for protecting privacy, is not merely a scientific project intended to run on clusters/supercomputers. Its true wide-ranging purpose is to protect the privacy of individuals as they interact with other individuals and large central databases. Ideally SIPPA would most effectively protect privacy if SIPPA clients/servers were deployed among personal computers without the need for third parties. However, sequential SIPPA, when used for comparing two 640X480 RGB images shown previously, takes a prohibitive amount of time (well over 1800 minutes). PSIPPA, as shown above, with a cluster of 16 Dual-Core machines reduces this time to around 15 minutes.

PSIPPA, as mentioned before, requires a cluster. PSIPPA could potentially be bottlenecked by network data transfer issues. Since GPU's are designed for around 500 cores running at the same time, trying to access memory; they have an asynchronous memory read/write bandwidth of around 180GB/s. This helps to reduce the data transfer workload to only the necessary cross network data transfer required for 2-party secure computation. Given the widespread availability of GPU's, which might facilitate wider adoption of SIPPA as a privacy protection mechanism in households and to examine its usability/performance over PSIPPA, we proceeded to implement SIPPA in CUDA-C (Nvidia 2010).

As shown in section 4, the optimal vector dimension for SIPPA is less than 20, which leads to a typical SIPPA session on a 640X480 RGB PPM image consisting of 92160, 10X10 individual matrices. Since these individual matrices are small for utilizing GPU LAPACKS, we proceeded to convert required matrix operations like SVD, Eigen Decomposition, Matrix Multiplication etc. from standard C LAPACKS to CUDA-C. These converted C functions now can be called by each GPU thread to allow for the simultaneous execution of operations like SVD on many different matrices at the same time. Each thread on the GPU which is sequential, takes ownership of one of those 92160 matrices and performs the required matrix operations of SIPPA iteratively. Thousands of these threads are scheduled in parallel and execute in parallel on the many cores available on a GPU.

5.3.1 NVIDIA CUDA-C enabled GPU thread execution overview

Unlike CPU threads, there is negligible overhead in creating thousands of CUDA threads. In fact, for the GPU to function at anywhere close to 100% occupancy, tens of thousands of

threads must be scheduled. This is due to the fundamental problem of parallel computing; i.e. memory access. Although CUDA architecture allows for 500 cores executing threads in parallel, allowing random access to GPU memory for each of these parallel threads creates bottlenecks where a thread executing in one of the 500 cores needs to wait on the memory access queue. By scheduling tens of thousands of threads, CUDA swaps threads waiting for memory access with other threads which might perhaps use the core more effectively.

A thread in CUDA executes a specified function iteratively. In CUDA when a function is launched by the Host (CPU) on the Device (GPU), the Host specifies the number of threads that need to be scheduled to launch the same function in parallel. Every thread is assigned a unique ID by CUDA; these ID's can be of one, two or three dimensions. Threads are further organized into blocks, each of these blocks can have up to a three dimensional ID. In modern GPU's the Host can schedule to execute a CUDA-C function on around 65000 blocks, with each containing up to 1024 threads. Threads in a block have access to certain CUDA privileges among threads in the same block; e.g., threads in the same block can cooperate to share limited low latency memory on each of the streaming multiprocessors. This is particularly useful when the output of threads in a block is input to other threads in the same block, or if a certain set of data is shared among threads, threads in a block can also be synchronized and be forced to execute certain parts of code before proceeding. The key here is, by utilizing these ID's assigned by the CUDA scheduler, one can precisely direct each of these thousands of threads to access, process and deposit to unique locations in memory, or perhaps to precisely choreograph a certain pattern of interaction with certain other specific threads.

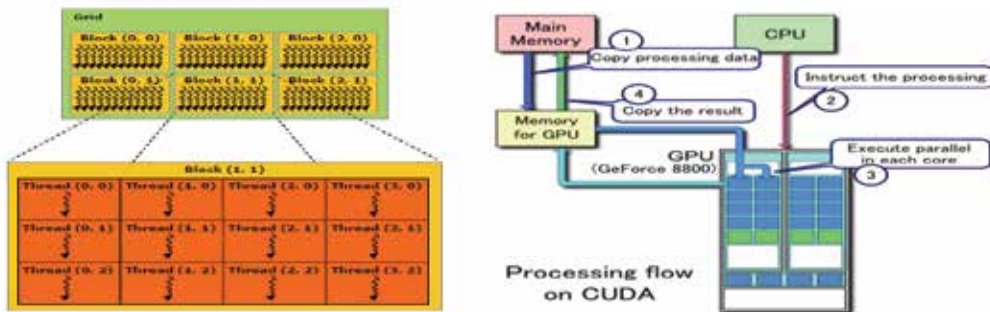


Fig. 19. Block, Thread Illustration (Kirk 2009)(Sander 2010)

5.3.2 Some notes on how we achieved MPSIPPA

To illustrate our process of achieving Massively Parallel SIPPA (MPSIPPA), perhaps it will be fruitful to describe how we implemented the simplest of our CUDA-C functions i.e. Vector Multiplication. As stated before, libraries for utilizing GPU's to perform matrix operations on multitudes of small independent matrices don't really exist. Our idea was, since usable SIPPA essentially leads to thousands of independent SIPPA sessions with each pair of sessions working on matrices ranging from 5X5 to 20X20, why not use the ability of modern GPU's to create and run thousands of parallel threads, and assign each of these threads ownership to one particular SIPPA session.

For the purposes of this example, let's assume that the Client would like to reconstruct a ppm image on the Server, and they both are using a candidate 640X480 RGB PPM image. Let's also assume that they settled on a vector size of 10X1(V). Processing this 640X480 RGB

PPM image would require 92160 independent client server SIPPA sessions. As the initial step, both the Client and Server machines read a particular candidate file in their disk drives into Host Memory (CPU Memory). All 92160 10X1 vectors are stored in a long one dimensional array, both on the client and server. On both client and server, The Host (CPU) then allocates a similar sized one dimensional array on Device (GPU) memory. Following allocation, the entire array of 921600 elements are copied from Host memory to Device memory. The Host also allocates on both Host and Device memory another long one dimensional array of size 92160*100 to store the results of $V.V^T$ (i.e. a 10X10 matrix), 100 values for each of the 92160 threads. Certainly a two dimensional or even a three dimensional array would be more intuitive. However, we noticed that a single malloc on the GPU, no matter the size of memory allocated, takes a constant amount of time. Allocating multiple smaller regions of memory and then storing their pointers in other arrays was too time prohibitive. Since threads can be arranged in two or ever more dimensions, their multi-dimensioned ID's can be used to specify what Vector they should work on.

```
VecMul<<<MAXTHR/THDSIZE, THDSIZE >>>(InputV, InputMat);
```

Once all required memory is allocated on the Host and Device, and Vector values copied from Host to Device, the Vector Multiplication function which multiplies the vector with its transpose to produce a 10X10 matrix for each of the required 92160 sessions is called by the host. Just shown above is the function call. The values inside "<<<>>>" specify the number of threads to create on the GPU, to execute the function in parallel. We surely require 92160 sessions, and therefore 92160 threads in total on the client and a corresponding number of 92160 threads on the server side; i.e. each thread assumes all computation iteratively for one SIPPA session. The first value in the angle bracket specifies the number of blocks to create, while the second value specifies the number of threads per block to create. Certainly there is a limit of 1024 threads per block; therefore multiple blocks need to be created. The optimal setting for the number of blocks to create and the number of threads to assign to each block to achieve optimal performance is not a precise endeavor, and requires experimentation. Factors that influence this are the number of registers used by a function, amount of shared memory used per block etc. For this function, our experiments show that optimal performance is reached (i.e. most cores are occupied for the majority of time) when the number of threads per block is 256 and the number of blocks are (No of threads required, 92160 in this case)/256. In this straightforward implementation, one entire SIPPA session is executed in a single thread. There are conceivably other implementations. For example, there are roughly six steps that SIPPA performs according to its algorithm, and one of these steps involving PPCSC is an aggregate of seven other steps for message exchange under a 2-party secure computation scenario. Each of these (12) steps could be implemented to run as a thread, thus the corresponding (12) threads in the same block realizes the implementation of one SIPPA session, and the parallel computing process for multiple SIPPA sessions is realized by scheduling multiples of these blocks to run in parallel. There are also other implementations that could be studied in future research to determine the optimal realization of the SIPPA implementation in a GPU environment.

The CUDA API takes care of all scheduling, and will only run threads as resources become available. When the Host calls Vector Multiplication, 92160 threads are scheduled on the GPU to execute the above function. Many of these threads execute in parallel, i.e. many of these SIPPA sessions (with each thread representing one particular SIPPA session) run in

parallel on the GPU. But from the programmer's perspective they all execute in parallel, unless any block specific synchronization is requested. In order to determine the place in memory to obtain the specific 10X1 vector a particular thread owns, the Unique Thread id (Utid in the above function) is needed. Utid is calculated utilizing the block id of the specific thread multiplied by the thread id of the particular thread. Using this Utid offsets one can calculate where to read the vector from, and where to store the resulting matrix (as shown in Fig. 20). Each thread block has the privilege of utilizing a 16 Kilobyte area of on-chip low latency memory, which a thread can access in around one to two clock cycles. In contrast, on-board device memory takes about 20-30 cycles to access. Given the fact that during Vector multiplication the same 10 numbers will be accessed repeatedly, creating a shared memory array of size 256*10 (when vector size is 10) will lead to a noticeable performance increase. It is because each of the threads in the block can copy over the 10X1 vector and perform multiplications utilizing the shared memory. Sure enough doing so leads to a performance increase of over 50% in vector multiplication. Since CUDA-C is essentially a subset of C, i.e. the GPU instruction set can essentially perform most functions a general purpose i386 processor can perform, we implemented SIPPA required functions like SVD, Calculating Eigenvectors, Pseudo Inverses using SVD etc. in CUDA-C by essentially translating widely available Open Source C LAPACK code. Translations were essentially trivial - mostly replacing standard C math functions with CUDA-C provided math functions. These functions were called in a similar manner to Vector Multiplication with each thread being assigned a particular matrix and that thread will call these iterative functions as necessary. Eigenvalues were calculated using the Power method since at this time SIPPA only uses the largest Eigenvalue and its corresponding Eigenvector. Inverses were Moore-Penrose pseudo-inverses calculated for precision utilizing the SVD method. Once required calculations are complete, and results are copied back to Host memory, on both the Client and Server Side. The client initiates a communication request with the server, and sends over $V_1^T A_1$. It sends the entire packet of 92160 10X10 matrices, in one stream of bytes. This is because sending individual matrices does not make sense when in GPU one has to provide all the matrices at the same time to utilize the power of parallel computing. Moreover, all results are also obtained at the same time when a GPU function is called. After a series of back and forth communication between the Server and Client as detailed before in the section discussing SIPPA algorithm, the server proceeds to send the entire block of X vectors (i.e. an array of size 92160*10 elements) to the client when the threshold condition is met. In other words, the server determines whether the threshold condition is met by comparing the predefined threshold to the distance between its Eigenvector and its estimate of the client's Eigenvector. The client also receives two other arrays from the client, a 92160 element array containing all server Eigenvalues and a 92160 element array containing the helper data $De^T \cdot x$ for all 92160 Vectors. Once the Client receives these three sets of data from the Server, it can proceed to reconstruct the server side image. One interesting issue encountered during the display of reconstructed images was that if the value of the reconstructed pixel was even slightly above the max intensity of 255 (in the case of ppm) the pixel would be displayed as black. This was remedied by re-writing all values above 255 as 255. This slight change has led to improvements in reconstruction.

5.3.3 MPSIPPA performance

We were surprised to note the immense increases in performances when using GPU's for SIPPA. The same task i.e. running SIPPA on a 640X480 RGB PPM image that took 1800



Image 1



Image 2

Reconstruction of server
image 2 with client using
image 1Reconstruction of
server image 1 with
client using image 1Reconstruction of
server image 1 with
client using image 3

Image 3

per second which vary from 5-60 with each frame representing one image. With each of these images a SIPPA session can be performed to reconstruct an estimate of an *image* located on the server side. After each image is reconstructed it may be sent to a matching engine to compare the template image with the reconstructed image. If this process happens 100 times, the matching engine may provide the system with 100 different scores each representing the similarity of the reconstructed image versus the template image. Doing this process multiple times may provide the system capabilities for better true acceptance rates and lower false rejection rates. The larger the sample of video results in a larger test bed of reconstructed images. As this test bed grows larger, the performance of the system may improve.

In testing MPSIPPA's performance with image reconstruction from surveillance footage we noticed a slight increase in performance in comparison to the performance for image reconstruction from static images. The performance time for a surveillance footage providing 50 images (5fps at 10seconds), resulted in approximately 5,000ms per image with a total run time of 246,380ms. Per image, this is approximately 2,000-3,000ms faster than static image comparison (Both Client & Server were on the same machine i.e localhost). With the initial experimental results, using MPSIPPA as a component in a real-time video surveillance face recognition and matching system seems promising. A short video of 5 seconds may be captured in 640x480 quality with 5fps and SIPPA would be capable of reconstructing estimates of the server side data within 5000ms per image. In approximately 120 seconds MPSIPPA would have generated 25 images for use in an efficient matching engine. These results do not consider time to write to file, however when using MPSIPPA as a component in a biometric system, it may not be necessary to write the file to disk, but merely keep it in memory and give the next component a stream of data directly from

memory. Eliminating the write to file seems a more practical and likely setup and with this setup MPSIPPA will provide close to 1400 images for verification in the same time it took MPSIPPA to write 25 images. If 25 images are all that are required to increase the performance of a biometric system, they may be reconstructed in approximately 2 seconds.

6. Conclusion

In summary, we present a secure computation technique that guarantees security and privacy for the reconstruction of voiceprints and images containing biometric information. Our secure computation technique, referred to as Secure Information Processing with Privacy Assurance (SIPPA), aims for sufficient privacy homomorphism and computational efficiency. Although it is not complete privacy homomorphism, SIPPA can be applied to the class of secure computation problems that is linearly decomposable. For proof of concept, we conducted both a simulation study and an application of SIPPA to reconstruct images with face biometrics as well as voice biometrics. Our experimental study showed that the results were consistent with the theory behind SIPPA. To achieve near real time performance, PSIPPA is studied and integrated in a slice-based architecture for realizing the full potential of SIPPA to achieve real time performance.

As a closing remark, we would also re-iterate many interesting research problems arose from this project; for example, what is the impact on the security and privacy when the original SIPPA protocol for 2-party secure computation is expanded to multi-party scenario in SFB-BioSIPPA? How do we handle failover in PSIPPA and SFB-BioSIPPA environment? What real time computation performance impact should be anticipated when MPSIPPA is scaled to operate in a wide-LAN environment?

Notwithstanding these astronomical increases in performance, NVIDIA CUDA provides mechanisms to further optimize performance. One might look into utilizing the ability of the newer CUDA enabled GPUs, which allow for asynchronous write to Host memory, or perhaps the ability to both load data into GPU memory while the GPU cores are simultaneously busy executing a function; or perhaps the ability of GPU's to directly display content on screen at over 300 Frames per second eliminating any file write (which seems to have taken over 8 seconds in our particular MSIPPA case). GPU programming surely holds promise in various areas including the field of biometrics, privacy protection including privacy preserving surveillance and computing: Real time multiple Face tracking/recognition, Object tracking, Biometric Intelligence involving identification of attributes such as gender, ethnicity, age etc. The possibilities are truly endless.

7. Acknowledgment

This work is supported in part by a grant from PSC-CUNY Research Award, PSC-CUNY CIRG17, and NSF DUE 0837535.

8. References

BioAPI 2.0 (May 2006), ISO/IEC 19784-1, Information Technology - BioAPI - Biometric Application Programming Interface - Part 1: BioAPI Specification (<http://www.itl.nist.gov/div893/biometrics/standards.html>)

- Du, W.; Atallah M.J. (2001a). Secure Multi-Party Computation Problems and Their Applications: A Review and Open Problems. In *Proceeding of New Security Paradigms Workshop*, pp. 11-20.
- Du, W.; Atallah, M.J. (2001b). Privacy-Preserving Cooperative Scientific Computations. In *Proceeding of 14th IEEE Computer Security Foundations Workshop*, pp. 273-282.
- Gentry. C. (2009). Fully Homomorphic Encryption Using Ideal Lattices, In the *Proceeding of 41st ACM Symposium on Theory of Computing (STOC)*.
- Goldreich, O.; Micali, S.; Wigderson A. (January 1987). How to play ANY mental game, *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, p.218-229, New York, New York, United States
- Goldwasser, S. (1987). Multi party computations: past and present, *Proceedings of the sixteenth annual ACM symposium on Principles of distributed computing*, p.1-6, August 21-24, 1997, Santa Barbara, California, United States
- Gross, R; Airoldi, E; Malin, B.; Sweeney, L. (2005) Integrating Utility into Face De-Identification.
- Kirk, D.B.; Hwu, W.W; (2009). Programming Massively Parallel Processors: A Hands-on Approach, ISBN 0123814723, Morgan Kaufmann.
- Nvidia (2010) http://www.nvidia.com/object/cuda_showcase_html.html as seen Oct 2010.
- Newton, E.; Sweeney, L.; Mali, B. (February 2005). Preserving Privacy by De-identifying Facial Images. *IEEE Transactions on Knowledge and Data Engineering*, 17 (2), pp. 232-243.
- Sanders, J.; Kandrot, E; (2010). CUDA by Example: An Introduction to General-Purpose GPU Programming, ISBN 0131387685, Addison-Wesley Professional.
- Solove, D.J.; Rotenberg, M.; Schwartz P.M. (2006). Privacy, Information, and Technology, ISBN 0-7355-6245-8, Aspen Publishers, Inc.
- Sy, B.K. (2009a). Slice-based Architecture for Biometrics: Prototype Illustration on Privacy Preserving Voice Verification. *Proceeding of Biometrics: Theory, applications and systems Conference*, Washington D.C.
- Sy, B.K. (2009b). Secure Computation for Biometric Data Security Application to Speaker Verification. *IEEE Systems Journal*, Vol. 3, Issue 4.
- Sy, B.K. (2009c) http://www.qcwireless.net/biometric_ppr/he_primer.pdf
- Sy, B.K.; Ramirez, A.; Arun P. Kumara Krishnan (2010). Secure Information Processing with Privacy Assurance Standard based Design and Development for Biometric Applications, *Proceeding of 8th International Conference on Privacy, Security and Trust*.
- Turk and A. Pentland (1991). Face recognition using eigenfaces. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. pp. 586-591. Available online at: <http://www.cs.ucsb.edu/~mturk/Papers/mturk-CVPR91.pdf>.
- Yao. A.C. (1982). Protocols for secure computations. In *Proceeding of 23rd IEEE Sym. On Foundations of Computer Science*.

Implementing Multimodal Biometric Solutions in Embedded Systems

Jingyan Wang, Yongping Li, Ying Zhang and Yuefeng Huang
*Shanghai Institute of Applied Physics, Chinese Academy of Science
P.R. China*

1. Introduction

Embedded systems are widely used in the areas of PIM (Personal Information Management) and safety-critical mechanical manipulation. With the increasing demands of privacy protection and safety reliability, these systems confront with all kinds of security concerns. To start with, they are possibly operated in physically insecure environment. The small-size feature of the devices such as cellphones and PDAs lends them easy to be lost and stolen. Furthermore, increasing programmability and networking function of these devices make them feeble to secure against various hacker assaults. While recent advances in embedded system security have addressed issues like secure communication, secure information storage, and tamper resistance (protection from physical and software attacks), objectives such as user-device authentication have often been overlooked, placing a hidden danger on the overall security of the system (Yoo Jang-Hee; Ko Jong-Gook; Chung Yun-Su; Jung Sung-Uk; Kim Ki-Hyun; Moon Ki-Young; Chung Kyoil, 2008).

Traditional methods for personal identification depend on third-party objects such as keys, passwords, certifications, etc. However, these media could be lost or forgotten. Another possible way to solve these problems is through biometrics, for each person has his own special biometric features definitely. Biometric features that can be used for identification include fingerprints, palm prints, handwriting, vein pattern, facial characteristics, iris, and some others like voice pattern and gait. Biometrics-based authentication system is emerging as the most reliable solution (Zuniga AEF; Win KT; Susilo W, 2010). However, personal identity recognition based on any unimodal biometric may not be sufficiently robust or may not be feasible to a particular user group or under a particular situation. Unimodal biometric systems are usually affected by problems including noisy sensor data, inconformity and lack of individuality of the chosen biometric trait, absence of an invariant representation for the biometric trait and susceptibility to circumvention. Some of these problems can be relieved by using multimodal biometric systems, which consolidate evidence from multiple biometric sources (Fan Yang; Baofeng Ma, 2007). Multimodal biometric technology has been developed to an important approach to alleviate the problems intrinsic to unimodal biometric systems and getting more concerns in biometric area (Xiuqin Pan; Yongcun Cao; Xiaona Xu et al., 2008). The most recent commercial and research multi-biometric systems adopt software implementation on PC computer and require a dedicated computer for the image or digital

signal-processing task—a large, expensive, and complicated-to-use solution, which is not practical for embedded devices like mobile phones. In order to make biometric recognition ubiquitous, the system's complexity, size, and price must be substantially reduced. This chapter investigates the problem of supporting efficient multimodal biometrics-based user authentications on embedded devices fusing two or more biometrics. In these devices, most traditional ways of interaction (e.g. keyboard and display) are limited by small size, power source and cost. The embedded system based on biometric authentication is applied as the platform of personal identification.

On the one hand, compared with traditional biometric identification systems, the embedded devices of biometric recognition have plenty of advantages. It is low-cost, simple-to-use, no dedicated image sensor; On the other hand, compared with the unimodal biometric systems in embedded device, the embedded multimodal biometrics need more capture devices and should run more than one algorithms. Additionally, it also needs a fusion method to improve the accuracy performance. In this chapter, we will introduce how to design an embedded multimodal biometric system, and describe several embedded multimodal biometric solutions, including the algorithms and the designs of the software and hardware.

The purpose of section 2 is to provide a general guidance for the readers to design a high performance embedded multimodal biometric system. In this section, we discuss two main problems which should be considered in the design of an embedded multi-biometric system: the selection of embedded platform and the biometric algorithms. In the first place, we investigate several embedded platforms suited for biometrics systems, including ARM based MPU processor, Multi-Core Processor combining ARM and DSP cores and so on. Afterwards, we introduce several biometrics algorithms designed for the implementation on embedded devices and the rules to select and optimize them. Following the guidance in section 2, we present three examples for the design of embedded multi-biometrics system in the following sections.

In section 3, we present a multi-biometric verification solution aiming at implementing on embedded systems within a wide range of applications. The system combines the voiceprint with fingerprint and makes the decision at score level. The fusion strategy is based on score normalization and support vector machine (SVM) classifier. This embedded platform adopts an ARM9-Core based S3C2440A microprocessor and the Microsoft Windows CE operation system. An external module PS1802 produced by Synochip Corporation is employed as fingerprint sub-system whilst the voiceprint sub-system uses the microphone of the developing board to capture vocal biometric samples.

In section 4, a new multi-biometrics system is designed for multi-core OMAP3 processor combing GPP and DSP cores, fusing iris and palmprint at sensor level (image level). The algorithm is based on phase-based image matching, which is effective for both iris and palm recognition tasks. Hence, we can expect that the approach can be useful for multimodal biometrics system with palmprint and iris recognition capabilities. The system accomplishes the fusion of palmprint and iris biometric at image level. A new image fusion algorithm, Baud limited image product (BLIP), designed especially for phase-based image matching is proposed. The algorithm is particularly useful for implementing compact iris recognition devices using the state-of-the-art DSP technology. OMAP3 process is utilized to realize this algorithm and then the new effective multi-biometrics system is proposed. Experiment results prove that the new scheme can not only improve the system accuracy performance, but also reduce the memory size used to store the templates and the time consumed for the matching.

In section 5, we introduce a DaVinci based multi-biometrics verification system mainly from the prospective of system design and fusion strategy. Verification systems require flexibility to solve different sorts of situations, so we adopt component-based architecture combined with simultaneous hardware and software considerations to address the problem. In addition, because methods to fuse multiple biometrics have also determined the improvement of the systems' performance, we raise the FAR-score strategy, which normalizes the scores into false acceptance rate. Once scores from all classifiers are normalized into FARs, common fusion rules could be utilized to calculate a singular scalar to make the final decision. The proposed system could fulfill the goals of flexibility and the enhancement of verification accuracy. The paper would be concluded in section 6.

2. General guidance: How to select multi-biometrics algorithms and embedded platforms

In this section, we discuss the general principles for designing an embedded multiple biometrics authentication system. The discussion is two-fold: how to choose the embedded platform and design the multiple biometrics algorithms. We should notice that, though we mention the above two sections of embedded multiple biometrics system separately, they must be considered jointly when you are going to complete the system. Moreover, the essential rule of design is not to choose the most powerful embedded platforms or the most effective algorithms, but to satisfy the requirements of the user. We recommend that the readers keep this in mind, so that you can understand the followings are just options, not the necessarily optimal choice for the designer.

To sum up, the embedded multi-biometrics system, such as a hand-held personal authentication system owns the following two characters:

1. Unlike the traditional uni-biometrics system, it combines two or more biometric modalities for more secure authentication. The advantage of the fusion multimodal biometrics lies on the improvement of the accuracy of the system by fusing more information. Of cause, the improvement depends on the deft fusion strategy. However, it requires the embedded processor to run more than one biometrics algorithms, which might be fatal for resource-limited embedded system on both processor and memory. Thus, when the design chooses the modalities, the complexity of the algorithm should be considered with the fusion algorithms. At the same time, the way to capture the biometric data is also an important factor.
2. The system differs with other traditional CP based systems, for the embedded system is limited on resources for complex biometrics verification algorithms. Considerable though the advantages of the prospective embedded biometrics solutions enjoy, they can not diminish the realistic difficulties current systems suffer from, such as the limited resources of the embedded device, high computational expenses of the biometrics algorithm and so on.

2.1 Algorithms

The biometric fusion procedure usually involves two steps. The first is to choose appropriate biometrics, which could provide essential information for recognition. The second is to design an effective method for fusing the biometrics. First of all, we introduce some commonly used

biometric modalities; then we will discuss the fusion methods. The biometrics can be an option for multi-biometrics including the following examples as displayed in (Fig 1).

- **Iris** is a kind of biometrics with high security. However, it needs a special capture device, which limits its applications. We design a high-security authentication system in case II, using iris as one modality, providing for the readers as reference.
- **Fingerprint** is another high performance biometrics. Similar to iris, it also needs a special fingerprint sensor. However, among the most frequently used biometric solution, there are many commercial devices which could be integrated into self-designed systems directly, making it an excellent choice for identity authentication.
- **Face** can be captured easily with a general camera integrated in a cellphone or a PDA, but it is easily influenced by the clients' posture, the environment's illumination and so on. Nonetheless, this disadvantage can be compensated by fusing with other biometrics insensitive to the above factors.
- **Palmprint** has several advantages compared with other biometrics (Ito K.; Aoki T.; Nakajima H. et al., 2006): palmprint capture devices are cheaper than iris devices; and palmprints contain additional distinctive features which can be extracted from low-resolution images. However, the accuracies of these approaches are not so satisfied for the requirement of some high secure applications.
- **Voiceprint** is the most natural modality for PDA or cellphone based embedded system, for most mobile devices can capture voice signals using a microphone. This feature is used in case I.

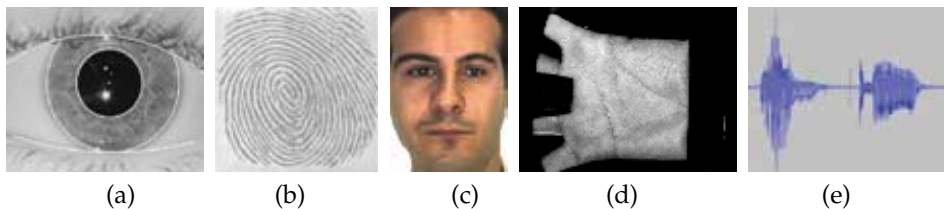


Fig. 1. Biometric modalities used in embedded multi-biometrics system: (a)Iris; (b)Fingerprint; (c)Face; (d)Palmprint; (e)Voiceprint.

Multi-biometric systems fuse information from multiple biometric sources in order to achieve better recognition performance and overcome other limitations of unibiometric systems (Nandakumar K; Chen Y; Dass SC et al., 2008). Fusion can be performed at four different levels of information, namely, sensor, feature, match score, and decision levels (Fig 2).

1. Fusion at the sensor level means the biometric data is fused directly before the features are extracted. This kind of fusion can preserve most parts of the information, for it combines the biometric modalities before they are processed further. The case II in this chapter utilizes this fusion strategy, integrating palmprint and iris image at the pixel level. This fusion needs flexible algorithms and is not general for all the other biometric modality.
2. In fusion at the feature-extraction level, the features extracted using two or more sensors are concatenated. The fusion is established by joining two or more features into a long vector. This category of modes is not practical because the features of the various

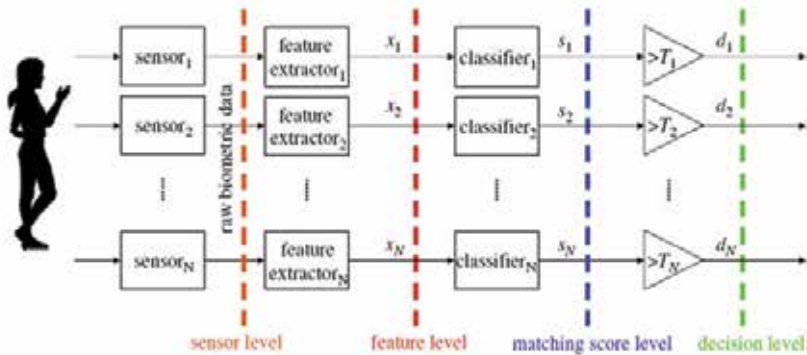


Fig. 2. Four levels at which multimodal biometrics can be fused.

modalities could be incompatible. For example, face images normally have larger sizes than those of finger images. Moreover, in this type of fusion modes, the recognition system does not work if one or more modalities of testing samples are not available.

- In fusion at the matching-score level, the matching scores obtained from multiple matchers are combined. Furthermore, score fusion techniques can be again divided into the following three categories:
 - Transformation-based score fusion. The match scores are first normalized (transformed) to a common domain and then combined with each other. The case III is based on this fusion strategy.
 - Classifier-based score fusion. Scores from multiple matchers are treated as a feature vector; then, a classifier is constructed to discriminate genuine and impostor scores. We adopt this strategy in case II.
 - Density-based score fusion. The approach is based on the likelihood ratio test. It requires explicit estimation of genuine and impostor match score densities (Jingyan Wang; Yongping Li; Xinyu Ao et al., 2009).
- In fusion at the decision level, the accept/reject decisions of multiple systems are jointly considered.

In practice, fusion at match score level and decision level are usually employed since they are much easier to accomplish, but in these modes, the useful information has never been exploited for fusion before the match and decision.

Here, we recommend the readers to choose the fusion strategy jointly with the biometric modalities. Two rules can be referenced as follows:

- Because fusion at the lower level can preserve more useful information than higher level, we should first consider lower fusion (sensor or feature level). However, lower level algorithms are hard to design, for the original data is quite distinctive for different modalities. The best chance to use low-level fusion is when different modalities can be matched in the same way. An example is given in case II in section 4 of this chapter, as iris and palmprint can both be matched by POC function. Another advantage of low-level fusion is that it only matches once for verification, which reduces the time and space complexity of the algorithm.

2. When the modalities cannot be fused at low-level, the reader can consider the matching score level. Fusion at this level has been studied a lot by researchers and plenty of effective algorithms have been developed. An example is given in section 3.

2.2 Embedded platform

For the biometric solutions, there are many embedded platforms to be considered:

GPP A common ARM-Core processor, such as LPC2106 (Martin T., 2004) by NXP Semiconductors, or S3C2440A (SAMSUNG Electronic Company, 2004) by SAMSUNG Electronics Company, can be a good choice for multi-biometrics system. For example, an advanced smart card chip ($5mm \times 5mm$) can employ 32-bit ARM7 or ARM9 CPU, 256KBytes of ROM program memory, 72KBytes of EEPROM data memory, and 8KBytes of RAM at most. Since the smart card chip has very limited memory, typical biometric verification algorithms cannot run on the card successfully. However, we can use a commercial device for a uni-biometric verification, just like what we do in section 3, so that the complex verification algorithm is finished outside the ARM processor with only the fusion being done by ARM. The advantage of using ARM is that it can usually run an Embedded WinCE or Linux operating system. This is very useful for real-world applications, because it can provide a good user interface.

DSP DSP might be the most suitable processor for biometric algorithms. Some powerful DSP solutions provided by Texas Instruments have already been used in biometric systems, which could refer to Wencang Zhao; Zhen Yang; Haiqing Cao (2010); Xin Zhao; Mei Xie (2009); Shah D.; Han K.J.; Narayanan S.S. (2009); Yanushkevich (S.N.; Shmerko) for more information. We must notice that the only use of DSP is not enough for the real-world applications, for it usually cannot provide a friendly user interface. This is mainly because it cannot run an OS like ARM which is designed to work with an embedded system. This shortage can be overcome by the so-called multi-core processor.

Multi-core processor Recently, some powerful embedded multimedia platforms have been proposed by TI. Two typical examples are the DaVinci and OMAP. These platforms are often combined with an ARM based GPP processor and a DSP processor. At the same time, some software and hardware components have also been provided to the developers to establish the communication between them. For multi-biometrics systems, the complex verification algorithms can be implemented on the DSP core, and the user interface can be implemented in the OS on ARM core. We will give two examples in this chapter, in section 4 and section 3 separately.

FPGA FPGA or CPLD is another choice for multi-biometrics solution. However, due to the complexity of the design, we usually won't consider it as a practical option.

At last, we should note that the choice of embedded platform should consider the two following factors:

1. Are the algorithms complex? If yes, we recommend you to implement in Multi-core system like DaVinci; else, a sample MCU or ARM processor will be enough.
2. Is this system stand-alone or integrated to existing embedded system? If it is a stand-alone system, you will have more freedom to choose a platform; if it is integrated, apparently, it should match the existing processor, and what you need to do is just to develop a new software system.

3. Case details I: ARM based multi-biometrics fusing fingerprint and voiceprint

In this case, we propose a new multi-biometric verification solution aiming at implementing on an embedded system within a wide range of applications. The system combines the voiceprint with fingerprint and makes the decision at score level. Fusion strategy is based on score normalization and support vector machine (SVM) classifier. We test the performance of SVM using three kernel functions for system adaptation. Experimental results demonstrate that proposed multi-biometric verification approach achieve. 1.0067% in equal error rate (EER), which means it can be deployed in the majority of embedded devices such as PDA and smart cellphone for user identity verification. We first introduce the single biometric verifiers, including the fingerprint and voiceprint. Then the proposed score level fusion method is given. Afterwards, we describe the design and implementation of multi-biometric system. Finally, we show the performance testing results.

3.1 Fingerprint and voiceprint verifiers

3.1.1 Fingerprint verifier

For fingerprint verification implementation, the full-functioned fingerprint identification system-on-chip (SOC) PS1802 produced by Synochip Corporation (Synochip Corporation, 2006) is employed. PS1802 fingerprint module uses a commercial minutia extraction algorithm, including image preprocessing, binarization, thinning and minutia finding. The output image of each process is given, as we can see from Fig. 3. With these minutia features, the alignment-based elastic matching algorithm is used.



Fig. 3. Output images of PS1802 modala's minutia extraction algorithm.

3.1.2 Voiceprint verifier

The voiceprint recognition system is content-dependent; it accepts voice samples for up to 10 seconds and enrolls the user in less than 4 seconds. The speech recordings used for feature extraction are utterances of a 4-digit PIN in English. The recording speech is divided into several small segments with a fixed length. Then a 34-dimensional feature vector is calculated using 20ms Hamming windows with 10ms shift. Each feature vector consists of (concatenated):

- the Mel Frequency Cepstral Coefficients (size 16);
- the energy coefficient (size 1),
- the first order derivatives of the MFCC (size 16)
- the delta energy (size 1).

The number of feature vectors between users and presentations may differ. With these feature vectors, we train the code book for each speaker with VQ (Vector Quantization) (Cai Geng-ping; Huang Shun-zhen; Xu Zhi-hong, et al.).

3.2 Multiple biometrics fusion method

As to fuse fingerprint and voiceprint verification systems, a score vector $X = (x_1, x_2)$ representing the score output of multiple verification systems is constructed, where x_1 and x_2 correspond to the scores obtained from the fingerprint and voiceprint verification system respectively. Then the identity verification turns to be the problem of separating the 2-dimension score vector $X = (x_1, x_2)$ into two classes, genuine or impostor. In other words, identity verification typically equals to binary classification problem, i.e. accept (genuine) or reject (impostor). We adopt SVM as the fusion strategy of the fingerprint and voiceprint identity verification system.

3.2.1 Score normalization

In our approach, the raw scores from fingerprint and voiceprint match system are normalized before they can be inputted into SVM, following (Jain, A; Nandakumar). These score can be normalized by max-min method as follows:

$$x = \frac{x - \min}{\max - \min} \quad (1)$$

where \min and \max are the minimum and the maximum values of these scores x .

3.2.2 Support vector machine fusion

Support vector machine (SVM) is based on the principle of structural risk minimization (Suykens JAK; Vandewalle J., 1999). In Yuan Wang; Yunhong Wang; Tieniu Tan. (2004), SVM is compared with other fusion methods of fingerprint and voiceprint, and its performance is the best. In this paper, we pay attention to the performance of SVM with different kernel functions. The detailed principle of SVM has not been shown in this paper, but it can be seen in reference Suykens JAK; Vandewalle J. (1999). Three kernel functions of SVM used in our study are:

Polynomials $K(x, z) = (x^\top z + 1)^d, d > 0$

Radial Basis Functions $K(x, z) = \exp(-g||x - z||^2)$

Hyperbolic Tangent $K(x, z) = \tanh(\beta x^\top z + \gamma)$

In order to choose the best kernel function for our system, we test the performances of SVMs based on three kernel functions mentioned above, and the results can be seen at the Fig. 4.

Fig. 4 shows different SVMs with different kernel functions classifying genuine and impostor of fingerprint and voiceprint after normalization. We can see that three SVMs can all separate the two classes correctly. Their performances are similar; however, the number of support vectors and the difficulty to adjust parameters of kernel function are different. In our experiment, the SVM-poly is easier to be trained than SVM-RBF and SVM-sigmoid; the latter two need more patience during training period. Moreover, the classification error of the SVM-poly is the lower (0.3%) than the other two (0.4% and 0.5%), which makes polynomials kernel function our final choice.

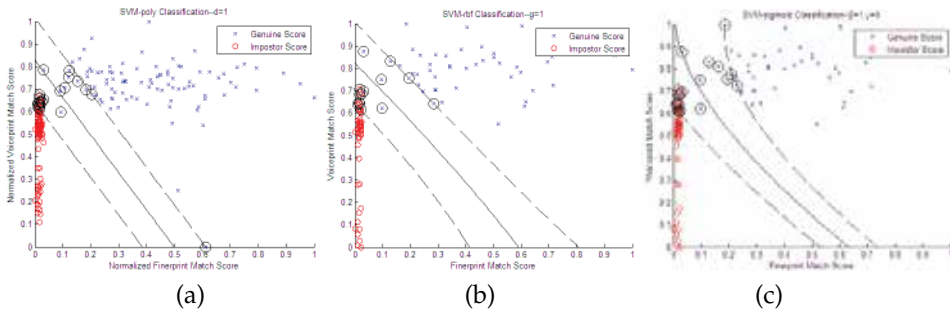


Fig. 4. SVM Classification results with different kernel functions: (a) SVM-Poly; (b) SVM-RBF; (c) SVM-Tanh.

3.3 Design and implementation of multi-biometric system

3.3.1 System frame and design scheme

The multi-biometric verification system is composed of three sub-systems: fingerprint sub-system, voiceprint sub-system and score level fusion sub-system. The embedded platform adopts an ARM9-Core based S3C2440A microprocessor and the Microsoft Windows CE operation system. An external module PS1802 produced by Synochip Corporation is employed as fingerprint sub-system whilst the voiceprint sub-system uses the microphone of the developing board to capture voice biometric samples. System software is developed by using Microsoft Embedded Visual C++. The system frame is shown in Fig. 5.

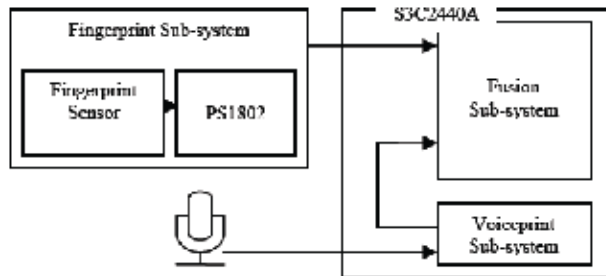


Fig. 5. The frame of multi-biometric verification system.

3.3.2 Hardware architecture

We utilize the S3C2440A, a 32-bit RISC microprocessor made by Samsung Company (SAMSUNG Electronic Company, 2004). To meet the demand of the audio capturing of voiceprint, the IIS-bus interface model with a UDA1341 audio CODEC is adopted. The Universal Asynchronous Receiver and Transmitter (UART) model is used as interface to fingerprint model PS1802. The techniques concerning the voiceprint recognition and multi-biometric fusion algorithm have been used in the system. Fig. 6 shows the hardware structure of the multi-biometric embedded system.

3.3.3 System software implementation

A multi-biometric verification system works in two models: enrollment model and verification model. In the off-line enrollment model, an enrolled fingerprint image and

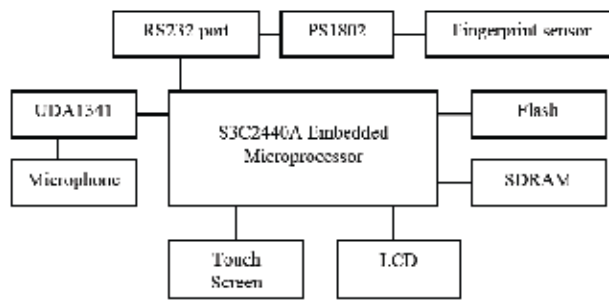


Fig. 6. The hardware structure of ARM-based multi-biometric verification system.

voice signal is preprocessed, and the features are extracted and stored into the on-board memory or the external SD card. In the on-line verification model, the similarity between the enrolled features and the features of real-time captured fingerprint image and voice signal are examined, giving two match scores. After fusing the two scores using SVM, decision can be determined by comparing the fusion score with the threshold. Fig. 7 shows the working models and data flow of a multi-biometric verification system.

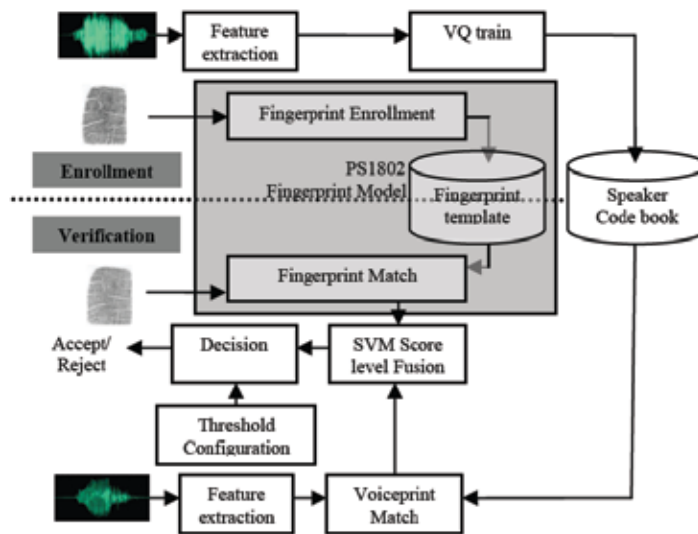


Fig. 7. The working models and data flow of a multi-biometric verification system.

4. Case details II: OMAP3 based multi-biometrics fusing iris and palm at image level

To improve the performance of the multimodal biometrics system, we proposed an effective image fusion method—band limited image product (BLIP) (Jingwang Liu; Yan Hou; Jingyan Wang; Yongping Li; Ping Liang, 2007), especially for phase based image matching. Using this method, we fuse iris and palmprint images to construct a multimodal framework, which can not only improve the security, but also can reduce the time and space complexity. Based on

OMAP3530's 'Dual Processor' character, we implement the algorithm and optimize it in terms of algorithm and programming, improving the execution efficiency further.

4.1 Multi-biometrics verification algorithm fusing iris and palmprint at image level

Figure 8 shows the overview of the proposed algorithm, which fuses the iris and palmprint image to one single image and uses it for verification. In this section, we describe the detailed process of the proposed algorithm, which consists of effective region extraction (to be explained in Section 4.1.1), image fusion (to be explained in Section 4.1.2), and matching score calculation (to be explained in Section 4.1.3).

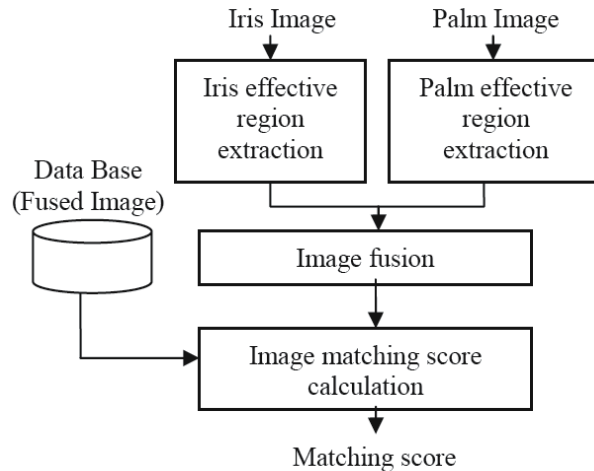


Fig. 8. Flow diagram of the proposed algorithm.

4.1.1 Iris and palm effective region extraction

To extract region from the iris, two circular boundaries of iris are searched by the integro-differential operators. Then, the disk-like iris area is unwrapped to a rectangular region by using doubly dimensionless projection, as shown in Figure 9.

In order to detect the effective palmprint areas in the palm image, we examine the n_1 -axis projection and the n_2 -axis projection of pixel values. Only the common effective image areas with the same size are extracted for the succeeding image matching step, as is shown in Fig. 10.

4.1.2 Baud limited image product fusion

In previous work (Miyazawa, Kazuyuki; Ito; Ito K.; Aoki T.; Nakajima H. et al., 2006; Miyazawa K; Ito K; Aoki T et al., 2006; Miyazawa, K.; Ito), the idea of the Phase-Only Correlation (POC) function for matching of iris and palm is proposed. Inspired by the POC function, we can fuse the image of iris and palm into one single image which containing all the phase information of iris and palm used for POC match. Since the product of two complex number's phase is the sum of the two phases, we can use the product of the iris and palm image, and the product image will contain the sum of the phase of the two. To solve the mismatch of the iris and palm' size and to improve the matching performance, Baud

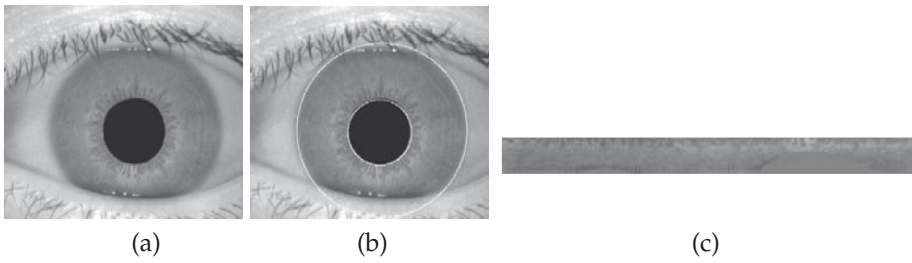


Fig. 9. Iris effective region extraction: (a) Iris image; (b) Two circular boundaries of iris; (c) rectangular region.

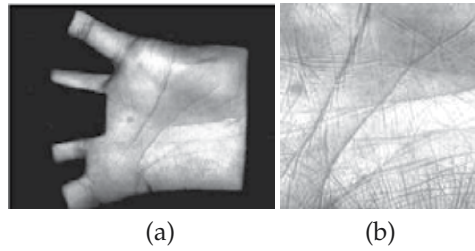


Fig. 10. Palm effective region extraction: (a) The palm image; (b) extracted common regions.

Limited 2D-IDFT (BL 2D-IDFT) considering the inherent frequency components of images is raised. Our observation shows that the 2D DFT of a normalized iris image and the extracted common palm regions sometimes includes meaningless phase components in high-frequency domains as illustrated in Fig. 11. The idea to improve the matching performance is to eliminate meaningless high frequency components in the calculation of 2D IDFT depending on the inherent frequency components of palmprint images.

For mathematical simplicity, consider $(2M_1 + 1) \times (2M_2 + 1)$ iris images $f_{iris}(n_1, n_2)$ and $(2N_1 + 1) \times (2N_2 + 1)$ iris image $f_{palm}(n_1, n_2)$, where we assume that the index ranges are $n_1 = -M_1, \dots, M_1$, $n_2 = -M_2, \dots, M_2$ for iris image and $n_1 = -N_1, \dots, N_1$, $n_2 = -N_2, \dots, N_2$ for palm image. Moreover, we assume that the ranges of the inherent frequency band are given by $k_1 = -K_1, \dots, K_1$ and $k_2 = -K_2, \dots, K_2$, where $K_1 = \text{Min}(M_1, N_1)$ and $K_2 = \text{Min}(M_2, N_2)$. Thus, the effective size of frequency spectrum is given by $L_1 = 2K_1 + 1$ and $L_2 = 2K_2 + 1$. The Baud Limited 2D-IDFT (BL 2D-IDFT) function is given by

$$f'(n_1, n_2) = \frac{1}{L_1 L_2} \sum_{k_1=-K_1}^{K_1} \sum_{k_2=-K_2}^{K_2} F(k_1, k_2) W_{K_1}^{-n_1 k_1} W_{K_2}^{-n_2 k_2} \quad (2)$$

Where the $F(k_1, k_2)$ is the 2D-DFT of $f(n_1, n_2)$, and $f(n_1, n_2)$ can be $f_{iris}(n_1, n_2)$ or $f_{palm}(n_1, n_2)$, as illustrated in Figure 5. Then the BLIP algorithm can be described in Algorithm 1.

The flow diagram of Baud Limited Image Product Fusion is shown in Figure 11.

4.1.3 Image matching using POC

We calculate the POC function $r_{fg}(n_1, n_2)$ between the two fusion images $f'_{fusion}(n_1, n_2)$ and $g'_{fusion}(n_1, n_2)$, and evaluate the matching score. Let $F(k_1, k_2)$ and $G(k_1, k_2)$ denote the 2D-DFTs of the two images $f'_{fusion}(n_1, n_2)$ and $g'_{fusion}(n_1, n_2)$. The cross-phase spectrum

Algorithm 1 Baud Limited Image Product Algorithm.

Require: The iris image $f_{iris}(n_1, n_2)$ and the palm image $f_{palm}(n_1, n_2)$;

Require: The fused image $f'_{fusion}(n_1, n_2)$;

Calculate 2D-DFTs of $f_{iris}(n_1, n_2)$ and $f_{palm}(n_1, n_2)$ to obtain $F_{iris}(k_1, k_2)$ and $F_{palm}(k_1, k_2)$;

Calculate BL 2D-IDFTs of $F_{iris}(k_1, k_2)$ and $F_{palm}(k_1, k_2)$ to obtain $f'_{iris}(n_1, n_2)$ and $f'_{palm}(n_1, n_2)$;

Calculate the pixel product of $f'_{iris}(n_1, n_2)$ and $f'_{palm}(n_1, n_2)$ to get the final fused image $f'_{fusion}(n_1, n_2)$, as follows,

$$f'_{fusion}(n_1, n_2) = f'_{iris}(n_1, n_2) \times f'_{palm}(n_1, n_2) \quad (3)$$

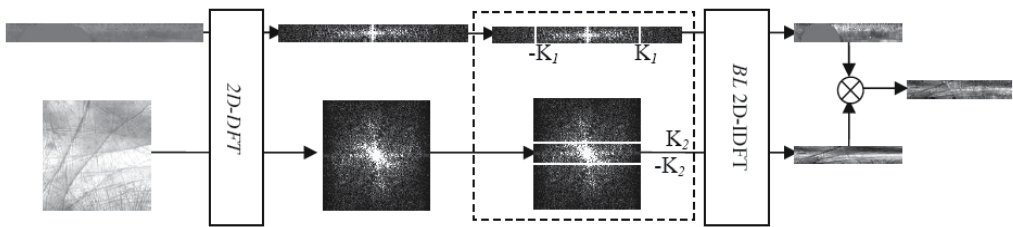


Fig. 11. Flow diagram of Baud Limited Image Product Fusion

$R_{FG}(k_1, k_2)$ is given by

$$R_{FG}(k_1, k_2) = \frac{F(k_1, k_2) \times \overline{G(k_1, k_2)}}{|F(k_1, k_2) \times \overline{G(k_1, k_2)}|} \quad (4)$$

where $\overline{G(k_1, k_2)}$ is the complex conjugate of $G(k_1, k_2)$. The POC function $r_{fg}(n_1, n_2)$ is the 2D Inverse DFT (2D-IDFT) of $R_{FG}(k_1, k_2)$ and is given by

$$r_{fg}(n_1, n_2) = \frac{1}{L_1 L_2} \sum_{k_1=-K_1}^{K_1} \sum_{k_2=-K_2}^{K_2} R_{FG}(k_1, k_2) W_{K_1}^{-n_1 k_1} W_{K_2}^{-n_2 k_2} \quad (5)$$

Figure 12 shows examples of genuine and impostor matching respectively. When two images are similar, their POC function $r_{fg}(n_1, n_2)$ gives a distinct sharp peak, or the peak value drops significantly, so the matching score is the highest peak value.

4.2 Design of OMAP3 based multi-biometrics system

The personal authentication system's framework is given in Fig. 13. According to the requirements of personal authentication, the system can be divided to two parts: the user enrollment module and the verification module.

- **Enrollment Module.** This module should capture the templates of users and store them into the database. First, the iris and palmprint images are captured; then the effective region of both iris and palmprint are extracted; finally they are fused using the BLIP

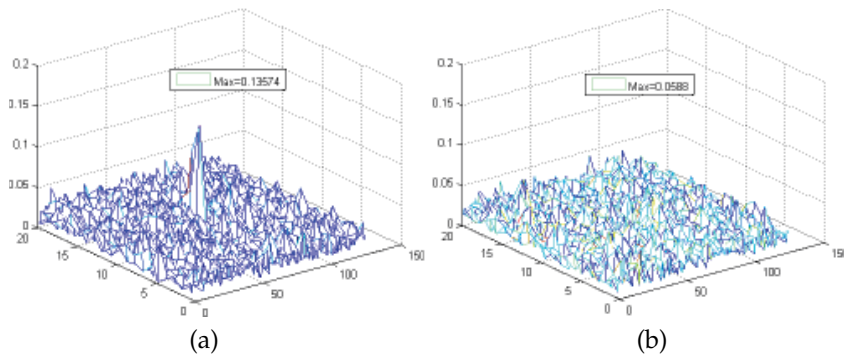


Fig. 12. Example of image matching using the POC function: (a) Genuine matching; (b) Impostor matching.

algorithm and stored in the database as templates. This procedure is illuminated in the pink dashed block in Fig. 13.

- **Verification Module.** In this module, the client first input his iris and palmprint image, and similarly, the effective region are extracted and fused to a single image, which contains the phase information of both iris and palm. Finally, this image with a small size is matched with the template using the POC function and the matching score is compared with the threshold to make a decision.

4.2.1 OMAP3530 applications processor

To design a multi-biometrics personal authentication system, a great variety of embedded system platforms and strategies are available for choice, but the following advantages have made TI's DSP based open multimedia application platform (OMAP) an excellent one. Firstly, it can accomplish complex DSP algorithm in real-time with very low power consumption and very small package size. Secondly, it integrates widely used and supported processors which represent the leading technical level. At last, The OMAP platform is based on a highly extensible architecture that can be expanded with application-specific processing capabilities and additional I/O so that even the most complicated multimedia applications will execute smoothly and seamlessly. In this case, a mobile multi-biometrics authentication terminal based on OMAP MPSoC device would be presented.

The device integrates multiple processors. The main parts consist of MPU and DSP (Texas Instruments, 2009):

MPU The MPU is a 600-MHz ARM Cortex-A8 processor core with a high-effectively, lower power consumption, and the DSP core is a 430-MHz TMS320DMC64x+ which is designed for digital signal processing. The MPU controls all resources of the device though running generic operating system such as Linux or Windows CE.

DSP The DSP acts as a coprocessor of the MPU. In addition, the device includes other processors or subsystems, such as 2D/3D hardware accelerator to deal with vector graphics processing.

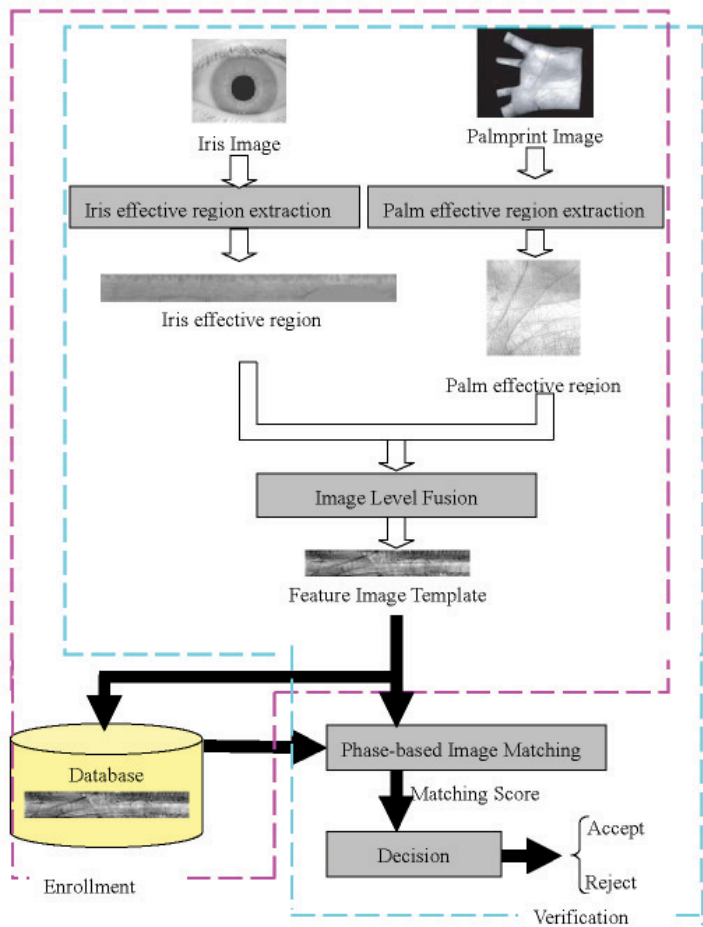


Fig. 13. Framework of the multi-biometrics software system.

4.2.2 System software implementation and optimization

The software design starts from embedded operating system running on ARM Cortex-A8 core. MontaVista Linux and Google Android have been successfully ported to the terminal respectively.

Another sort of important software for the terminal is DSP/BIOS Bridge (DSP Bridge). It is a software package designed by TI Instruments for OMAP platform. It enables asymmetric multiprocessing on target platforms which contain a general purpose processor (GPP) and one or more attached DSPs. It is a combination of software for both the GPP Operating System (OS) and DSP OS that links the two operating systems together. The linkage enables applications on the GPP and the DSP to easily communicate messages and data in a device-independent, efficient fashion (Jinhe Zhou; Tonghai Wu; Rongfu Wu, 2009). The software of OMAP platform includes the application layer and signal processing layer. The application layer runs in the Linux on OMAP's GPP core, while signal processing layer runs in OMAP's DSP core, and their interaction is through the Codec Engine.

Although the time complexity of the algorithms is not high, the preliminary implementation on the Linux in MPU core is not satisfying in real-time performance. ARM Cortex-A8 Core shows poor real-time performance in reading the image, extracting the effective region, 2D-FFT and 2D-IFFT transformation, image production and image matching. To finish this procedure, it takes ARM Cortex-A8 as long as 1 minute. To satisfy the real-time requirement of real-world applications, we need to optimize the algorithm. The strategy is to employ the DSP core in OMAP3530 processor to do the operation like FFT, so that the system can improve its performance. At the same time, we will also optimize the programs running in DSP core to save the operating time.

4.2.3 Algorithm transplantation on dual-core processor

According to the characteristics of the algorithms, we transplant parts of the program to the DSP core. The main program will call these functions and finish the enrollment and verification. The strategy of transplantation is shown in Table 1.

Program	GPP	DSP
Flow Control	all	-
User Input and Output	all	-
Iris Effective Region Extraction	Locating Iris Center	Inner and outer boundary detection; Iris Normalization
Palm Effective Region Extraction	Constructing axes; Extracting the center subimage	Palmprint binaryzation; Boundary detection
Image Fusion	-	all
Image Matching	-	all

Table 1. Strategy of Algorithm'S Transplantation

On the one hand, in the whole Flow Control procedure, there are many *if-else* and *switch* statements, and GPP is better than DSP on running these statements; on the other hand, the iris and palmprint images are both stored in the Linux file system on GPP, which cannot be accessed by DSP directly, so we decide to implement the main flow control and the input output procedure in GPP. The algorithms in the system, including the iris and palmprint effective region extraction, fusion, and matching, utilize many 2D-FFT and 2d-IFFT operations; at the same time, they are suitable for parallel computing, which can be executed by DSP's pipelines efficiently, so we choose to transplant them to DSP core. All the algorithms are implemented according to xDAIS standard, and are called by application program in Linux on GPP based on the Codec Engine (CE) module. The corresponding enrollment and verification flow-process diagram is given in Fig 14.

4.2.4 Program optimization on DSP

Because the TMS320C64x+ DSP core in OMAP3530 is a fixed-point processor and most of our algorithms are floating-point algorithms, we carry out fixed-point programming to the programs to improve the efficiency. According to our experiments, we scale the data using Q11, which can both keep the precision of the program and improve the algorithm's efficiency. Other technologies are also taken to improve the efficiency:

- Optimization of compiling options;

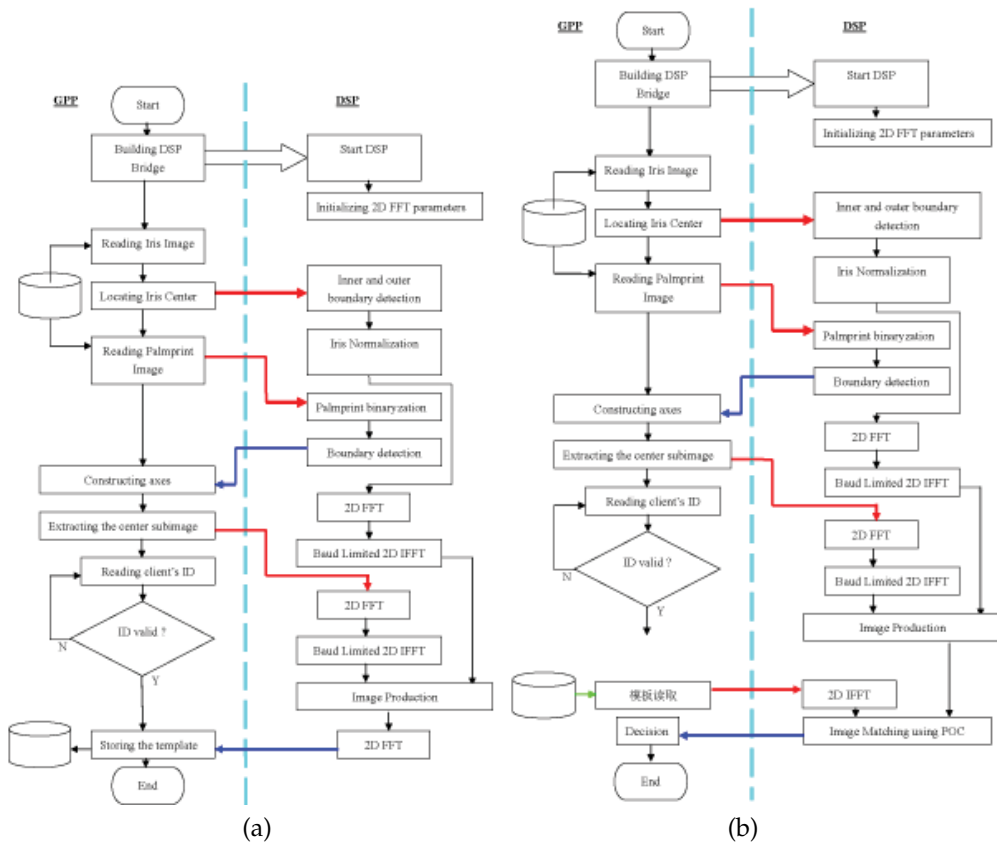


Fig. 14. Enrollment and Verification Flow-Process Diagram on Dual-Core Processor OMAP3530.

- Loop Unrolling;
- Block Data Movement Using DMA Module.

5. Case Details III: DaVinci based multi-biometrics verification

In this case, a multi-biometric verification system based on TI’s DaVinci DSP platform is presented. It aims to achieve the following goals: (1) deployed friendly with environment; (2) flexible to networking circumstances; and (3) configurable with changing on scheme of biometrics matching. Accordingly, we extend the component-based architecture to the embedded computing environment and this will be introduced in the first part. Based on the systematic design, a face recognition subsystem and fingerprint recognition subsystem are constructed to test the solution and explore the capability of multiple biometrics subsequently. In the end, conclusions would be drawn on the DaVinci based verification system.

5.1 System design on embedded system

The unpredictability of complex scene requires flexibility of the verification system. To address this issue, sensors capturing biometrics features should be plugged at any time, which

should be coined as "Plug and Play" standard updating its state dynamically. The system architecture as shown in Fig. 15 is given in the following. However, there are several other

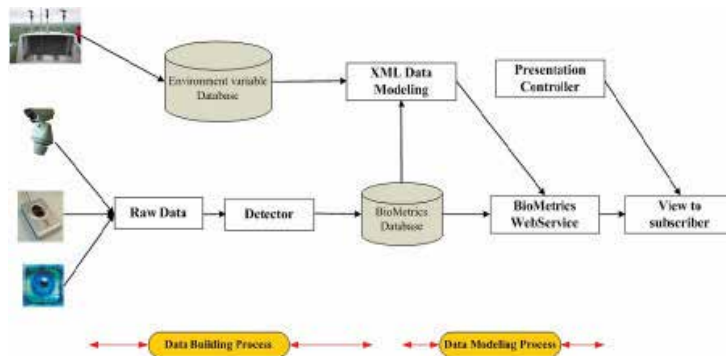


Fig. 15. Architecture of biometrics system.

problems existing in the biometrics verification system. First, verification system usually is based on simple assumptions. Biometrics is idealized to be captured in the best condition, where faces are deemed to be acquired with the frontal view. Secondly, the traditional dumb terminal model fails to adapt in more volatile circumstances, such as at the POS site or in the subway check-in station. Thirdly, occlusions and camouflages decrease the verification performance greatly in reality, which become the main obstacle for companies to adopt biometrics verification solutions in their business. Furthermore, limited biometrics excludes particular groups of people from the verification process, such as disabled person without fingerprint or palm print.

To eliminate problems above, the system should add precaution mechanisms to facilitate the usage. In our system, knowledge of the angle and distance to the reference point is added to the sensors. We try to find location mismatch itself and collect the input again from a sensor which can minimize such disparity between different locations. Resultantly, the system is equipped with the capability to find the change of its attached sensors and its related configuration. Currently, the change of sensors and its configuration depend on manual work, so does the trigger of reloading the new configuration.

As shown in Fig. 16, the red line between data layer and service layer is triggered when the system tries to send feedback to data layer for new data from different conditions. The hardware to support this kind of behavior is smart sensors based on DSP technology. The sensor is connected to the system using network cables. HTTP server is installed on the DSP server to listen to the port for possible commands action specified by the system. Command and control input is fed from specific port after the DSP chip receives the signal and will be dispatched to the GPIO interface of the DSP to control movement of the sensor. In the experiment, DaVinci Multimedia platform is only used for video processing. Hopefully, it could be extended to other sensors which could be sensitive to the position when capturing biometrics data in the future.

5.1.1 Hardware platform

In the DaVinci system, the TMS320Dm6446 DSP chip is used here. Within the DSP chip, there are ARM926EJ-S kernel, TMS32064x+DSP, video/image compressors (VICP), and Video

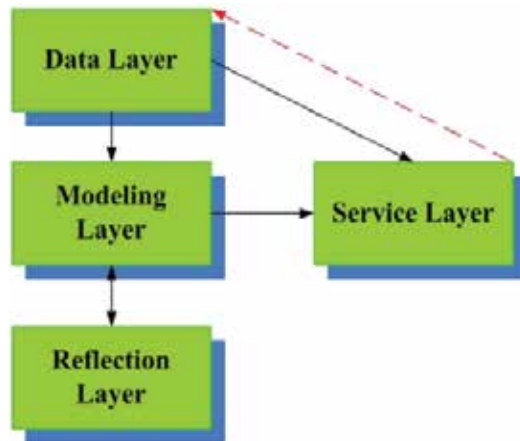


Fig. 16. Architecture of biometrics system.

Processing Sub System (VPSS). The DaVinci system is placed in the smart camera. The verification system is running on an IBM eServer with four Itanium CPUs, which is running Linux Server.

5.1.2 Software subsystem

The software module running on the DSP system includes a real-time Linux, which will communicate with the DSP hardware through DSP link as illustrated in Fig. 17. The software running on the DSP-enabled sensors abides the standard of TI as shown in Fig. 18. That is to say software module should be conformed to xDAIS standard. The numbers in the figure correspond to the following actions.

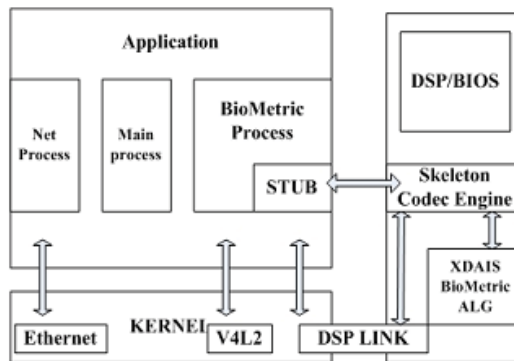


Fig. 17. Architecture of biometrics system.

- The GPP (General Processor Platform) side application makes an algorithm call;
- Codec Engine forwards this calling conventions to the GPP side algorithm stub;
- The stub places the argument in a compact inter-CPU message and replaces all GPP-side (virtual address) pointer values with DSP-side (physical address) values, which is called "marshalling" the argument;

- CE delivers the message to the DSP-side algorithm skeleton;
- The skeleton unmarshals the argument and calls the actual xDAIS algorithm's process function;
- On the way back, the skeleton marshals any return arguments, places them in a message and the stub unmarshals them for the application.

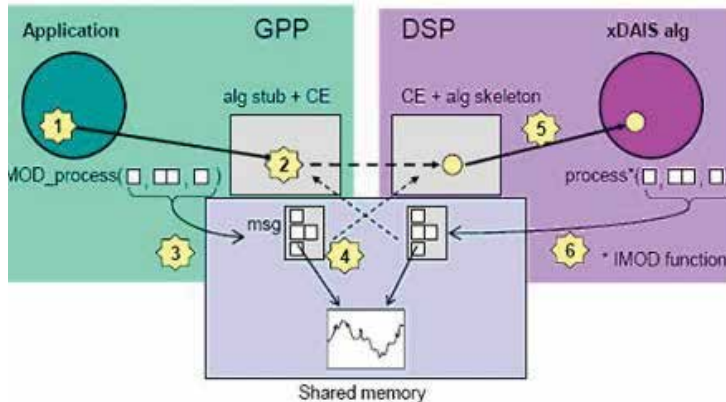


Fig. 18. Architecture of biometrics system.

These steps are excerpted from reference (Texas Instrument, 2007).

5.2 Fusion on multiple biometrics

There are three kind of multi-modal biometrics verification system: the first is the multi-algorithm system which employs different algorithms to verify a single biometric trait; the second is the multi-biometric system that involves two or more distinct modules of biometric traits; the third is the hybrid system wherein the multi-algorithm and multi-biometric systems are integrated together. The paradigm employed here takes the second way to fusion on multiple biometrics to reach a single conclusion. The fusion method employed here is to convert the matching score into the false acceptance rate. Score normalization for multi-classifier fusion refers to transform the various scores obtained by different classifiers into a common domain. Distinct matcher produces score diversely in numerical range and meaning, so the evaluation standards vary accordingly. It is necessary to normalize the scores into homogeneous domain before combination. When normalizing scores of different classifiers, two factors should be considered.

In practice, classifier outputs a matching score s to reflect the similarity between the testing sample Z and the claimed class. In general, s can be modeled as shown in Eq. 6.

$$s = f[P(\text{genuine}|Z)] + \eta(Z) \quad (6)$$

f is a monotonic function and η is the bias of the classifier and often supposed to be zero.

Jain et. al recommends normalizing scores with a certain functions such as z-score and their functions as below.

$$n = \frac{s - \text{mean}(S)}{\text{std}(S)} \quad (7)$$

$$n = \frac{1}{2} \left[\tanh\left(0.01 \frac{s - \text{mean}(S)}{\text{std}(S)}\right) + 1 \right] \quad (8)$$

n is the normalized score, mean and std denote the arithmetic mean and standard deviation operators respectively. S is the set composing of scores from the classifier. Although these functions use the statistical characters of scores such as means and variances, they do not follow the distributions of the scores from different classifiers.

We introduce a novel normalization method here, which converts scores into false acceptance rate. In the typical Receiver Operating Characteristic (ROC) curve of a classifier, two sorts of probabilities are relevant to the scores: the false acceptance rate and false rejection rate. They are functions of threshold (denoted as h) and can be written as following.

$$\begin{aligned} f_{far}(h) &= P(\text{genuine} | \text{imposter}, s < h) \\ &= \frac{\text{false positive}}{\text{positive instances}} \end{aligned} \quad (9)$$

$$\begin{aligned} f_{frr}(h) &= P(\text{imposter} | \text{genuine}, s > h) \\ &= \frac{\text{false negative}}{\text{negative instances}} \end{aligned} \quad (10)$$

To learn the FAR-score curve, a series of thresholds h should be calculated beforehand. For a training set of K classes, each class has m samples, so there are $mK(K-1)$ imposter samples in sum. For the j classifier, at the i th threshold of h_i^j , the false acceptance rate is $f(h_i^j)$. Using a set of thresholds ($h_{i-1}^j < h_i^j < h_{i+1}^j$), the FAR-score curve can be calculated. When a testing sample Z comes with a claim, the score S^j from the j th matcher can be normalized by the curve. If FAR monotonically increase with h^j , s^j is normalized by the following equations.

$$n^j = far_i^j(h_i^j) + \frac{d far_i^j}{dh^j} \Big|_{h=h_i^j} (s_i^j - h_i^j); \quad h_{i-1}^j \leq h_i^j \leq h_{i+1}^j \quad (11)$$

Otherwise, Eq. 12 is used to normalize S .

$$n^j = far_i^j(h_{i+1}^j) + \frac{d far_i^j}{dh^j} \Big|_{h=h_i^j} (s_i^j - h_{i+1}^j); \quad h_{i-1}^j \geq s_i^j \geq h_{i+1}^j \quad (12)$$

When scores from all classifiers are normalized into FARs, the common fusion rules such as sum, min, med and max can be adopted to compute a single scalar to make a final decision. In the experiment, the face module and fingerprint module will compute their score independently first, then it will be combined in the way in different algorithm to get the single value.

5.3 Conclusions on DaVinci based verification system

The DaVinci based system performs well and reaches the active responsiveness standard with the aid from outside. It could launch the newly developed modules at run time with the only need being that you specify the change in related configuration document. Currently, face and fingerprint modules are tested. The FAR-score curve of each classifier is computed without assumptions of observing any distributions, and scores from all classifiers can be normalized

by its own FAR-curve. Therefore, the method can be adapted to scores from any classifiers. However, the responsiveness needs extra improvement.

6. Conclusion

In this chapter, we discuss the design of multi-biometrics authentication system for embedded devices like high-end cellphones or PDAs. The aim here is to provide the readers on ideals about how to design a multi-biometrics satisfying the requirement of embedded devices. The general guidance is first given on how to select proper algorithms and embedded platforms. Then we introduce three useful examples in the following sections to show how to fuse them together.

With the fast development of mobile communication, embedded devices advance with each passing day. Because it is power efficient, fast authentication and compact in size, embedded device-based multi-biometrics verification system would be in widespread use in the near future.

7. References

- Yoo, Jang-Hee; Ko, Jong-Gook; Chung, Yun-Su; Jung, Sung-Uk; Kim, Ki-Hyun; Moon, Ki-Young; Chung, Kyoil. Design of embedded multimodal biometric systems. Proceedings - International Conference on Signal Image Technologies and Internet Based Systems, SITIS 2007, p 1058-1062, 2007.
- Zuniga, AEF; Win, KT; Susilo, W. Biometrics for Electronic Health Records. JOURNAL OF MEDICAL SYSTEMS Volume: 34 Issue: 5 Pages: 975-983 Published: 2010
- Fan Yang; Baofeng Ma A new mixed-mode biometrics information fusion based-on fingerprint, hand-geometry and palm-print. 2007 4th International Conference on Image and Graphics Pages: 689-93 Published: 2007 .
- Xiuqin Pan; Yongcun Cao; Xiaona Xu, et al. Ear and face based multimodal recognition based on KFDA. 2008 International Conference on Audio, Language and Image Processing Pages: 965-9 Published: 2008.
- Yuefeng Huang; Xinyu Ao; Yongping Li; Chengbo Wang (2008), Multiple biometrics system based on DavinCi platform, 2008 International Symposium on Information Science and Engineering (ISISE) Pages: (vol.2) 88-92 Published: 2008.
- Jingyan Wang; Yongping Li; Ping Liang, et al.(2009), An effective multi-biometrics solution for embedded device Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics. SMC 2009 Pages: 917-22 Published: 2009.
- Jingyan Wang; Yongping Li; Xinyu Ao, et al. Multi-modal biometric authentication fusing iris and palmprint based on GMM. 2009 IEEE/SP 15th Workshop on Statistical Signal Processing (SSP) Pages: 349-52 Published: 2009.
- Nandakumar, K; Chen, Y; Dass, SC, et al. Likelihood ratio-based biometric score fusion. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE Volume: 30 Issue: 2 Pages: 342-347 Published: 2008.
- Martin, T. Introduction to the LPC 2100 architecture. Embedded System Engineering Volume: vol.12, no.1 Pages: 30-2 Published: Jan.-Feb. 2004.

- Wencang Zhao; Zhen Yang; Haiqing Cao. The System of Face Detection Based on DSP. 2010 8th World Congress on Intelligent Control and Automation (WCICA 2010) Pages: 2834-7 Published: 2010.
- Xin Zhao; Mei Xie. A practical design of iris recognition system based on DSP 2009 International Conference on Intelligent Human-Machine Systems and Cybernetics. IHMSC 2009 Pages: 66-70 Published: 2009.
- Shah, D.; Han, K.J.; Narayanan, S.S. A low-complexity dynamic face-voice feature fusion approach to multimodal person recognition. Proceedings of the 2009 11th IEEE International Symposium on Multimedia (ISM 2009) Pages: 24-31 Published: 2009.
- Yanushkevich, S.N.; Shmerko, A.V. Fundamentals of biometric system design: new course for electrical, computer, and software engineering students. Proceedings of the 2009 Symposium on Bio-inspired Learning and Intelligent Systems for Security (BLISS 2009) Pages: 3-8 Published: 2009.
- Jwu-Sheng Hu; Ming-Tang Lee; Chia-Hsing Yang An embedded audio-visual tracking and speech purification system on a dual-core processor platform. Microprocessors and Microsystems Pages: 274-84 Published: 11 2010 Nov. 2010.
- Synochip Corporation: PS1802 DSP Brief Introduction. Technical report, Synochip Corporation (2006)
- Cai Geng-ping; Huang Shun-zhen; Xu Zhi-hong, et al. Voiceprint recognition system. Journal of Shenzhen University Science & Engineering Volume: vol.19, no.2 Pages: 78-80 Published: June 2002.
- Jain, A; Nandakumar, K; Ross, A. Score normalization in multimodal biometric systems. PATTERN RECOGNITION Volume: 38 Issue: 12 Pages: 2270-2285 Published: 2005.
- Suykens, JAK; Vandewalle, J. Least squares support vector machine classifiers. NEURAL PROCESSING LETTERS Volume: 9 Issue: 3 Pages: 293-300 Published: JUN 1999.
- Yuan Wang; Yunhong Wang; Tieniu Tan. Combining fingerprint and voiceprint biometrics for identity verification: an experimental comparison. Biometric Authentication. First International Conference, ICBA 2004. Proceedings (Lecture Notes in Comput. Sci. Vol.3072) Pages: 663-70 | xvii+800.
- SAMSUNG Electronic Company: S3C2440A 32-BIT RISC MICROPROCESSOR USER'S MANUAL.
- Miyazawa, Kazuyuki; Ito, Koichi; Aoki, Takafumi; Kobayashi, Koji; Nakajima, "An effective approach for Iris recognition using phase-based image matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, , v 30, n 10, p 1741-1756, 2008.
- Ito, K.; Aoki, T.; Nakajima, H., et al. A phase-based palmprint recognition algorithm and its experimental evaluation. 2006 International Symposium on Intelligent Signal Processing and Communications (IEEE Cat. No.06EX1444) Pages: 215-18 | CD-ROM Published: 2006
- Miyazawa, K; Ito, K; Aoki, T, et al., An iris recognition system using phase-based image matching, 2006 IEEE International Conference on Image Processing, ICIP 2006, Proceedings Pages: 325-328 Published: 2006.
- Miyazawa, K.; Ito, K.; Aoki, T., et al., An efficient iris recognition algorithm using phase-based image matching. 2005 International Conference on Image Processing Pages: II-49-52 | CD-ROM Published: 2006
- Texas Instruments,OMAP3530/25 Applications Processor,
<http://focus.ti.com.cn/cn/docs/prod/folders/print/omap3530.html>.

Jinhe Zhou; Tonghai Wu; Rongfu Wu, A mobile multimedia network terminal based on MPSoC 2009 First International Conference on Future Information Networks. ICFIN 2009 Pages: 121-5 Published: 2009.

Texas Instrument. Codec Engine Algorithm Creator User's Guide, Published: 2007.

Jingwang Liu, Yan Hou, Jingyan Wang, Yongping Li, Ping Liang, Fusing Iris and Palmprint at Image Level for Multi-Biometrics Verification, 2010 The 3rd International Conference on Machine Vision, Accepted.

Edited by Midori Albert

Biometrics-Unique and Diverse Applications in Nature, Science, and Technology provides a unique sampling of the diverse ways in which biometrics is integrated into our lives and our technology. From time immemorial, we as humans have been intrigued by, perplexed by, and entertained by observing and analyzing ourselves and the natural world around us. Science and technology have evolved to a point where we can empirically record a measure of a biological or behavioral feature and use it for recognizing patterns, trends, and or discrete phenomena, such as individuals' and this is what biometrics is all about. Understanding some of the ways in which we use biometrics and for what specific purposes is what this book is all about.

Photo by NicoElNino / iStock

IntechOpen

