Recent Application in Biometrics

Edited by Jucheng Yang and Norman Poh

RECENT APPLICATION IN BIOMETRICS

Edited by **Jucheng Yang** and **Norman Poh**

Contributors

Yongwha Chung, Woo Yong Choi, Jin-Won Park, Jing Liu, Shuo Wang, Jianchu Yao, Steve Warren, Yongbo Wan, Chin-Tsan Wang, Larissa Panina, Aythami Morales Moreno, Miguel A. Ferrer, Ping Wang, Chih-Chiang Ku, Tzu Chia Wang, Alberto de Santos, Carmen Sánchez-Avila, Javier Guerra-Casanova, Gonzalo Bailador-del Pozo, Tzovaras, Anastasios Drosou, Dimosthenis Ioannidis, Georgios Stavropoulos, Konstantinos Moustakas, Nobuyuki Nishiuchi, Hiroka Soya, Stefano Tennina, Luigi Pomante, Francesco Tarquini, Fabio Graziosi, Fortunato Santucci, Christophe Rosenberger, Rima Belguechi, Vincent Alimi, Patrick Lacharme, Estelle Cherrier, Mariano Fons, Francisco Fons, Robert Brumnik, Vladislav Kovtun, Sergii Kavun, Iztok Podbregar, Sam Milshtein, Anup Pillai, Vishal Oliyil Kunnil, Michael Baier, Paula Bustos

Notice

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

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

4,000+
Open access books available

116,000+
International authors and editors

120M+
Downloads

151
Countries delivered to

Our authors are among the
Top 1%
most cited scientists

12.2%
Contributors from top 500 universities

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

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

Meet the editors

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

Dr. Norman Poh received the Ph.D. degree in computer science in 2006 from the Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland. He is a Research Fellow in the Centre for Vision, Speech, and Signal Processing (CVSSP) at the University of Surrey, Guildford, Surrey, U.K. Since Sept 2006. He is one of the Work-Package Leaders in the Mobile Biometry (MOBIO) project (2008-2010), responsible for designing adaptive multimodal biometric systems. Prior to joining CVSSP, he was a research assistant at IDIAP research institute. His areas of interest are: pattern recognition, video processing, biometric authentication, and information fusion. He has authored more than 50 peer-reviewed publications, such as IEEE Trans. on Pattern Analysis and Machine Intelligence, IEEE Trans. on Information Forensics and Security, IEEE Trans. Audio, Speech and Language Processing and so on. He received five best paper awards (AVBPA'05, ICB'09, HSI 2010, ICPR 2010, Pattern Recognition Journal 2006) and two personal research grants from the Swiss National Science Foundation.

Contents

Preface

In the recent years, a number of recognition and authentication systems based on biometric measurements have been proposed. Algorithms and sensors have been developed to acquire and process many different biometric traits. Moreover, the biometric technology is being used in novel ways, with potential commercial and practical implications to our daily activities.

The key objective of the book is to provide a collection of comprehensive references on some recent theoretical development as well as novel applications in biometrics. The topics covered in this book reflect well both aspects of development. They include biometric sample quality, privacy preserving and cancellable biometrics, contactless biometrics, novel and unconventional biometrics, and the technical challenges in implementing the technology in portable devices.

The book consists of 15 chapters. It is divided into four sections, namely, biometric applications on mobile platforms, cancelable biometrics, biometric encryption, and other applications. Chapter 1 gives an overarching survey of existing implementation of biometric systems on mobile devices. Apart from the conventional biometrics, biomedical data such as blood pressure, ECG and heart beat signal are considered. The authors highlight the technical challenges that need to be overcome. Chapter 2 presents a biometric system based on hand geometry oriented to mobile devices. Chapter 3 exploits the recent advances in the biometric and heterogeneous wireless networks fields to provide an authentication platform that supports both physical and logical access management.

Section 2 is a collection of four chapters on cancelable biometrics. Chapter 4 provides a comprehensive overview on privacy preserving biometrics, also the state of the art in this field and presents the main trends to be solved. In Chapter 5 the author proposes a new attack algorithm for fingerprint template protection which applies a fast polynomial reconstruction algorithm based on the consistency theorem. Also, the proposed attack method is evaluated, and compared with the known attack methods. Chapter 6 introduces the technology of contactless fingerprinting and explores its application. In chapter 7 the author proposes a novel method of cancelable biometric identification that combines biological data with the use of artifacts and is resistant to spoofing.

Section 3 groups three biometric encryption applications. Chapter 8 proposes a user-specific biometric key with multimodal biometric for encryption. In Chapter 9, the authors apply a cryptography scheme known as the "Co-Z approach" to biometric systems. Chapter 10 presents a novel remote authentication scheme based on the secret-splitting concept for cloud computing applications.

Finally, Section 4 groups a number of novel biometric applications. Chapter 11 provides a comprehensive review of existing research work that exploits electrocardiograms (ECGs) for human identification as well as addresses several important technical challenges arise from this application. Chapter 12 investigates new magnetic sensing technologies for use in biometrics. In Chapter 13, the authors study run-time reconfigurable hardware platforms and hardware-software co-design techniques for biometric systems. Embedded systems based on programmable logic devices such as field programmable gate arrays (FPGA) are presented as case studies. Chapter 14 proposes a contactless biometric system based on the combination of hand geometry and palmprint using only low cost devices for medium security environments. The device uses infrared illumination and infrared camera in order to handle changing lighting conditions as well as complex background that contains surfaces and objects with skin-like colors. In Chapter 15 the biometric concept is applied to the fuel cells, microbial fuel cells and micromixer. The findings suggest that a novel flow slab design would be useful to improve Proton Exchange Membrane Fuel Cells (PEMFC) and can even be expanded to other types of cell, and the prototype will be useful in the design of a optimal biophysical passive micromixer and even show the feasibility and potential of biometric concept widely applied in biochemical, biological, chemical analysis, fuel cell and bioenergy.

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

Jucheng Yang
School of Information Technology
Jiangxi University of Finance and Economics , Nanchang, Jiangxi province,
China

Norman Poh
Centre for Vision, Speech and Signal Processing (CVSSP)
Faculty of Engineering and Physical Sciences , University of Surrey, Guildford, Surrey
U.K

Part 1

Application of Mobile Phone

Biometrics on Mobile Phone

Shuo Wang and Jing Liu
Department of Biomedical Engineering,
School of Medicine, Tsinghua University
P. R. China

1. Introduction

In an era of information technology, mobile phones are more and more widely used worldwide, not only for basic communications, but also as a tool to deal with personal affairs and process information acquired anywhere at any time. It is reported that there are more than 4 billion cell phone users over the world and this number still continues to grow as predicted that by 2015 more than 86% of the world population will own at least one cell phone (Tseng et al., 2010).

The massive volume of wireless phone communication greatly reduces the cost of cell phones despite their increasingly sophisticated capabilities. The wireless communication capability of a cell phone has been increasingly exploited for access to remote services such as e-commerce and online bank transaction. Smart phones are providing powerful functionality, working as a miniaturized desktop computer or Personal Digital Assistant (PDA). More excitingly, most of the state-of-the-art mobile phones are now being incorporated with advanced digital imaging and sensing platforms including various sensors such as GPS sensors, voice sensors (microphones), optical/electrical/magnetic sensors, temperature sensors and acceleration sensors, which could be utilized towards medical diagnostics such as heart monitoring, temperature measurement, EEG/ECG detection, hearing and vision tests to improve health care (Wang & Liu, 2009) especially in developing countries with limited medical facilities.

These scenarios, however, require extremely high security level for personal information and privacy protection through individual identification against un-authorized use in case of theft or fraudulent use in a networked society. Currently, the most adopted method is the verification of Personal Identification Number (PIN), which is problematic and might not be secured enough to meet this requirement. As is illustrated in a survey (Clarke & Furnell, 2005), many mobile phone users consider the PIN to be inconvenient as a password that is complicated enough and easily forgotten and very few users change their PIN regularly for higher security as can been seen from Fig. 1. As a result, it is preferred to apply biometrics for the security of mobile phones and improve reliability of wireless services.

As biometrics aims to recognize a person using unique features of human physiological or behavioral characteristics such as fingerprints, voice, face, iris, gait and signature, this authentication method naturally provides a very high level of security. Conventionally, biometrics works with specialized devices, for example, infrared camera for acquisition of

iris images, acceleration sensors for gait acquisition and relies on large-scale computer servers to perform identification algorithms, which suffers from several problems including bulky size, operational complexity and extremely high cost.
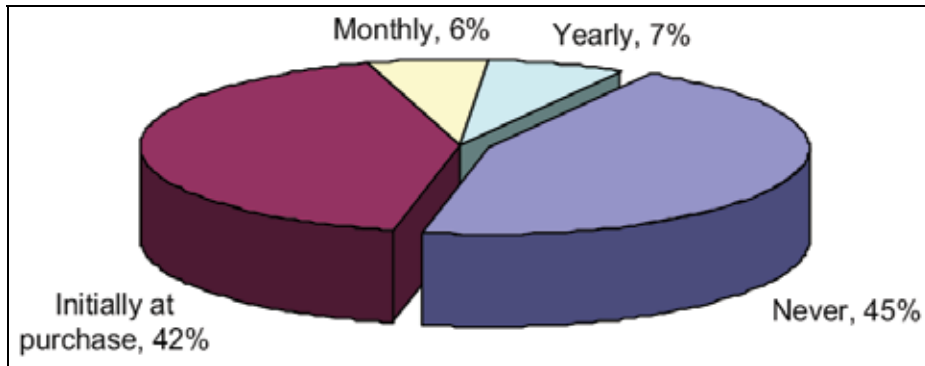
Fig. 1. Frequency of the change of PIN code. Reprinted from Computers & Security, Vol. 24, Clarke & Furnell, 2005, Authentication of Users on Mobile Telephones - A Survey of Attitudes and Practices, pp. 519-527, with permission from Elsevier

Mobile phone, with its unique features as small size, low cost, functional sensing platforms, computing power in addition to its wireless communication capability, is opening up new areas in biometrics that hold potentials for security of mobile phones, remote wireless services and also health care technology. By adding strong security to mobile phones using unique individual features, biometrics on mobile phones will facilitate trustworthy electronic methods for commerce, financial transactions and medical services. The increasing demand for pervasive biomedical measurement would further stimulate the innovations in extending the capabilities of a mobile phone as a basic tool in biometric area. This chapter is dedicated to drafting an emerging biomedical engineering frontier-- Biometrics on Mobile Phone. To push forward the investigation and application in this area, a comprehensive evaluation will be performed on the challenging fundamental as well as very practical issues raised by the biometrics on mobile phone. Particularly, mobile phone enabled pervasive measurement of several most important physiological and behavioural signals such as fingerprint, voice, iris, gait and ECG etc. will be illustrated. Some important technical issues worth of pursuing in the near future will be suggested. From the technical routes as clarified and outlined in the end of this chapter, it can be found that there is plenty of space in the coming era of mobile phone based biometric technology.

2. Feasible scenarios of biometrics on mobile phone

Incorporated with advanced sensing platforms which could detect physiological and behavioural signals of various kinds, many types of biometric methods could be implemented on cell phones. This offers a wide range of possible applications such as personal privacy protection, mobile bank transaction service security, and telemedicine monitoring. The use of sensor data collected by mobile phones for biometric identification and authentication is an emerging frontier and has been increasingly explored in the recent decade. A typical architecture of this technology can be seen in Fig. 2.
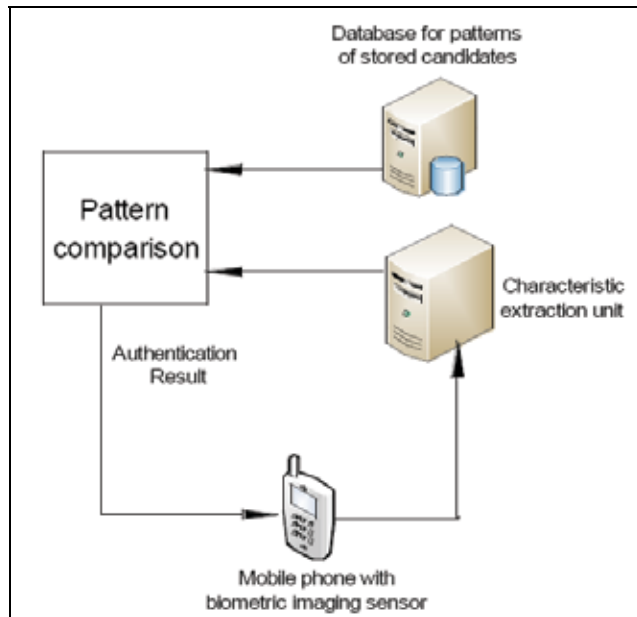
Fig. 2. Mobile biometric authentication system (Xie & Liu, 2010)

Several typical examples of recent advances which successfully implemented biometrics on mobile phones are described below.

2.1 Fingerprint identification on mobile phone

Fingerprint biometric has been adopted widely for access control in places requiring high level of security such as laboratories and military bases. By attaching a fingerprint scanner to the mobile phone, this biometric could also be utilized for phone related security in a similar manner.

A typical example can be seen from a research that utilizes a fingerprint sensor for acquisition of fingerprint images and implements an algorithm on internal hardware to perform verification of users (Chen et al., 2005). Experiment results show that this implementation has a relatively good performance. The prototype of this mobile phone based fingerprint system could be seen in Fig. 3.

Fig. 3. A schematic for fingerprint mobile phone (Redrawn from Chen et al., 2005)

Fig. 4. Snapshots of fingerprint security - Pro (retrieved from company release news
http://itunes. apple.com/us/app/fingerprint-security-pro/id312912865?mt=8)

One major inconvenience with mobile phone based fingerprint biometric is that it requires
an external attachment as a scanner of fingerprint images. Recently, iPhone launched an
application named Fingerprint Security by using its touch screen which does not require
external scanner (shown in Fig. 4).

2.2 Speaker recognition on mobile phone

A voice signal conveys a person's physiological characteristics such as the vocal chords,
glottis, and vocal tract dimensions. Automatic speaker recognition (ASR) is a biometric
method that encompasses verification and identification through voice signal processing.
The speech features encompass high-level and low level parts. While the high-level features
are related to dialect, speaker style and emotion state that are not always adopted due to
difficulty of extraction, the low-level features are related to spectrum, which are easy to be
extracted and are always applied to ASR (Chen & Huang, 2009).

One major challenge of ASR is its very high computational cost. Therefore research has been
focusing on decreasing the computational load of identification while attempting to keep the
recognition accuracy reasonably high. In a research concentrating on optimizing vector
quantization (VQ) based speaker identification, the number of test vectors are reduced by
pre-quantizing the test sequence prior to matching, and the number of speakers are reduced

by pruning out unlikely speakers during the identification process (Kinnunen et al., 2006). The best variants are then generalized to Gaussian Mixture Model (GMM) based modeling. The results of this method show a speed-up factor of 16:1 in the case of VQ-based modeling with minor degradation in the identification accuracy, and 34:1 in the case of GMM-based modeling.
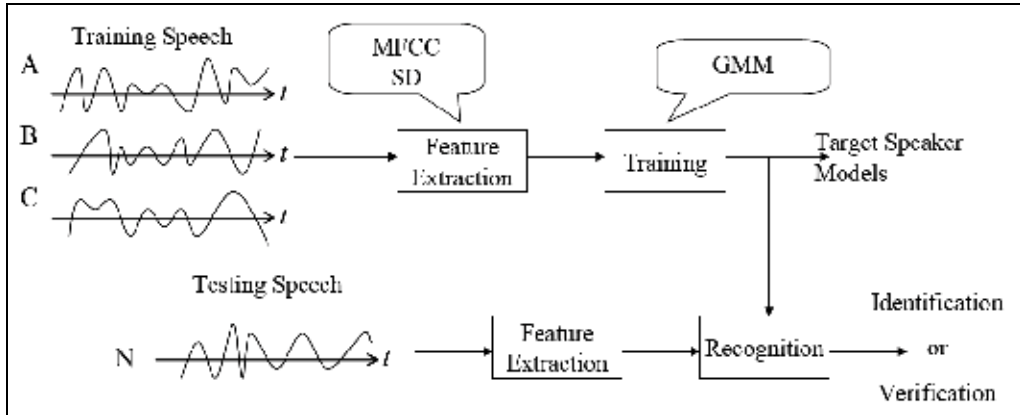
Fig. 5. Structure of a proposed ASR system. Reprinted from Proceedings of the 2009 Fourth International Multi-Conference on Computing in the Global Information Technology, Chen & Huang, 2009, Speaker Recognition using Spectral Dimension Features, pp. 132-137, with permission from IEEE
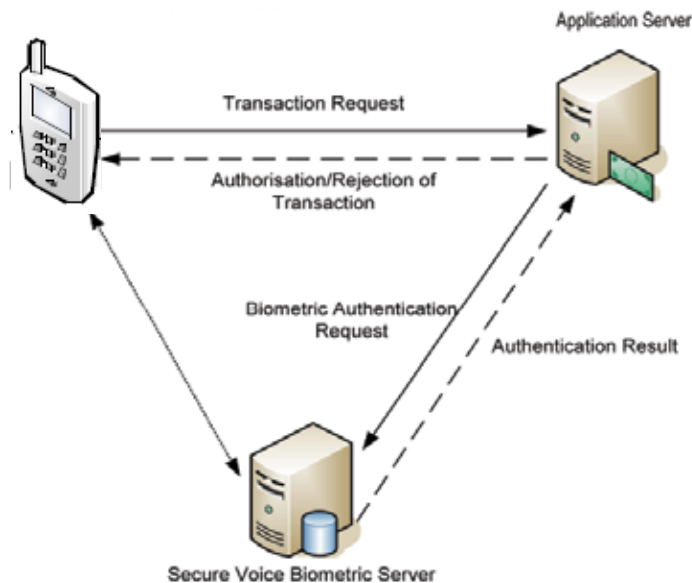
Fig. 6. Voice biometric authentication for e-commerce transactions via mobile phone. Reprinted from Proceedings of 2006 2nd International Conference on Telecommunication Technology and Applications, Kounoudes et al., 2006, Voice Biometric Authentication for Enhancing Internet Service Security, pp. 1020-1025, with permission from IEEE

By far, Mel Frequency Cepstral Coefficients (MFCC) and GMM are the most prevalent techniques used to represent a voice signal for feature extraction and feature representation in state-of-the-art speaker recognition systems (Motwani et al., 2010). A recent research presents a speaker recognition that combines a non-linear feature, named spectral dimension (SD), with MFCC. In order to improve the performance of the proposed scheme as shown in Fig. 5, the Mel-scale method is adopted for allocating sub-bands and the pattern matching is trained by GMM (Chen & Huang, 2009).

Applications of this speaker verification biometric can be found in person authentication such as security access control for cell phones to eliminate cell phone fraud, an identity check during credit card payments over the Internet or for ATM manufacturers to eliminate PIN number fraud. The speaker's voice sample is identified against the existing templates in the database. If the claimed speaker is authenticated, the transaction is accepted or otherwise rejected as shown in Fig. 6 (Kounoudes et al., 2006).

Although the research of speech processing has been developed for many years, voice recognition still suffers from problems brought by many human and environmental factors, which relatively limits ASR performance. Nevertheless, ASR is still a very natural and economical method for biometric authentication, which is very promising and worth more efforts to be improved and developed.

2.3 Iris recognition system on mobile phone

With the integration of digital cameras that could acquire images at increasingly high resolution and the increase of cell phone computing power, mobile phones have evolved into networked personal image capture devices, which can perform image processing tasks on the phone itself and use the result as an additional means of user input and a source of context data (Rohs, 2005). This image acquisition and processing capability of mobile phones could be ideally utilized for mobile iris biometric.

Iris biometric identifies a person using unique iris patterns that contain many distinctive features such as arching ligaments, furrows, ridges, crypts, rings, corona, freckles, and a zigzag collarette, some of which may be seen in Fig. 7 (Daugman, 2004). It is reported that the original iris patterns are randomly generated after almost three months of birth and are not changed all life (Daugman, 2003).

Recently, iris recognition technology has been utilized for the security of mobile phones. As a biometric of high reliability and accuracy, iris recognition provides high level of security for cellular phone based services for example bank transaction service via mobile phone.

One major challenge of the implementation of iris biometric on mobile phone is the iris image quality, since bad image quality will affect the entire iris recognition process. Previously, the high quality of iris images was achieved through special hardware design. For example, the Iris Recognition Technology for Mobile Terminals software once used existing cameras and target handheld devices with dedicated infrared cameras (Kang, 2010). To provide more convenient mobile iris recognition, an iris recognition system in cellular phone only by using built-in mega-pixel camera and software without additional hardware component was developed (Cho et al., 2005). Considering the relatively small CPU processing power of cellular phone, in this system, a new pupil and iris localization algorithm apt for cellular phone platform was proposed based on detecting dark pupil and corneal specular reflection by changing brightness & contrast value. Results show that this algorithm can be used for real-time iris localization for iris recognition in cellular phone. In 2006, OKI Electric Industry Co., Ltd. announced its new Iris Recognition Technology for

Mobile Terminals using a standard camera that is embedded in a mobile phone based on the original algorithm OKI developed, a snapshot of which can be seen in Fig. 8.
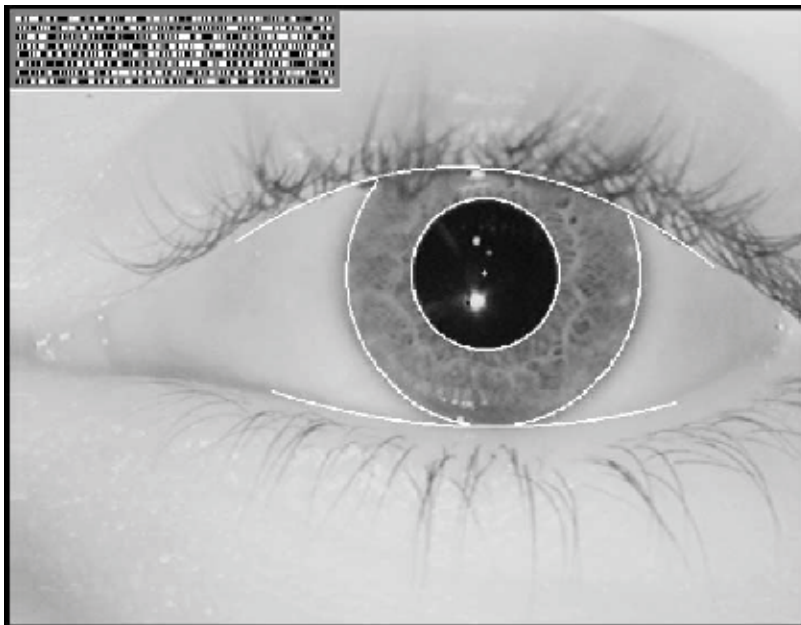
Fig. 7. Example of an iris pattern image showing results of the iris and pupil localization and eyelid detection steps. Reprinted from Pattern Recognition, Vol. 36, Daugman, 2003, The Importance of Being Random: Statistical Principles of Iris Recognition, pp. 279-291, with permission from Elsevier

Fig. 8. Iris recognition technology for mobile terminals (OKI introduces Japan's first iris recognition for camera-equipped mobile phones and PDAs, In: *OKI Press Releases*, 27.11.2006, Available from http://www.oki.com/en/press/2006/z06114e.html)

Since iris image quality is less controllable with images taken by common users than those taken in the laboratory environment, the iris image pre-processing step is also very important for mobile applications. In recent research, a new pupil & iris segmentation method was proposed for iris localization in iris images taken by cell phone (Cho et al., 2006; Kang, 2010), the architecture and service scenarios of which is shown in Fig. 9. This method finds the pupil and iris at the same time, using both information of the pupil and iris together with characteristic of the eye image. It is shown by experimental results that this method has good performance in various images, even when they include motion or optical blurring, ghost, specular refection, etc.
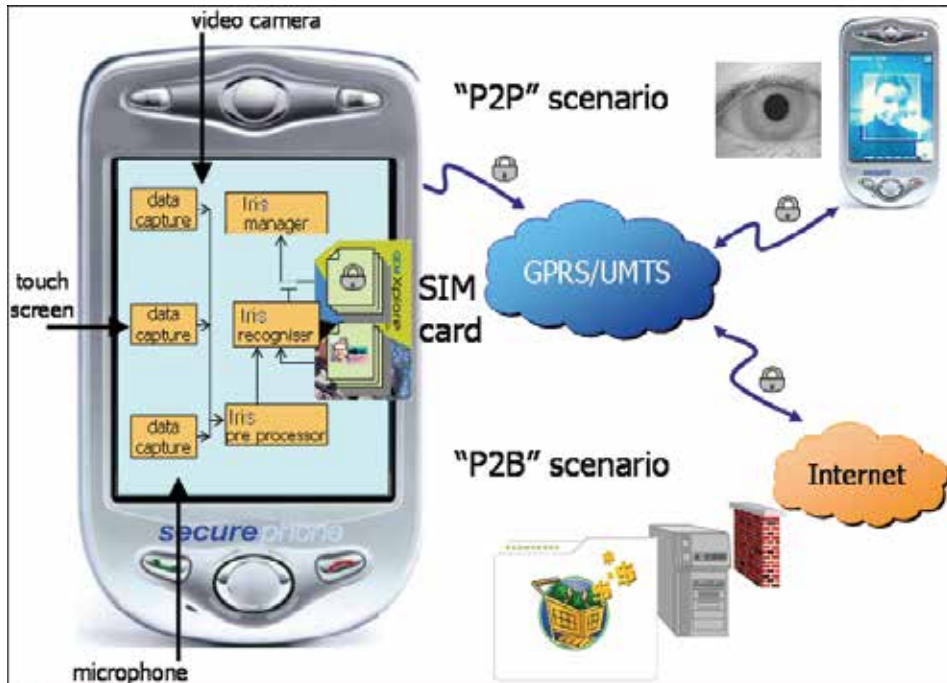
Fig. 9. Architecture and service models of mobile iris system. Reprinted from Procedia Computer Science, Vol. 1, Kang, 2010, Mobile Iris Recognition Systems: An Emerging Biometric Technology, pp. 475-484, with permission from Elsevier

2.4 Unobtrusive user-authentication by mobile phone based gait biometrics

Mobile phones nowadays contain increasing amount of valuable personal information such as wallet and e-commerce applications. Therefore, the risk associated with losing mobile phones is also increasing. The conventional method to protect user sensitive data in mobile phones is by using PIN codes, which is usually not secured enough. Thus, there is a need for improving the security level in protection of data in mobile phones.

Gait, i.e., walking manner, is a distinctive characteristic for individuals (Woodward et al., 2003). Gait recognition has been studied as a behavioral biometric for more than a decade, utilized either in an identification setting or in an authentication setting. Currently 3 major approaches have been developed for gait recognition referred to as the Machine Vision (MV) based gait recognition, in which case the walking behavior is captured on video and

video processing techniques are used for analysis, the Floor Sensor (FS) based gait recognition by placing sensors in the floor that can measure force and using this information for analysis and Wearable Sensor (WS) based gait recognition, in which scenario the user wears a device that measures the way of walking and recognize the pattern recognition for recognition purposes (Bours & Shrestha, 2010). Smart phone, such as an iPhone, is now incorporated with accelerometers working along three primary axes (as shown in Fig. 10), which could be utilized for gait recognition to identify the user of a mobile phone (Tanviruzzaman et al., 2009).

Fig. 10. Three axes of accelerometers on an iPhone (Redrawn from Tanviruzzaman et al., 2009)
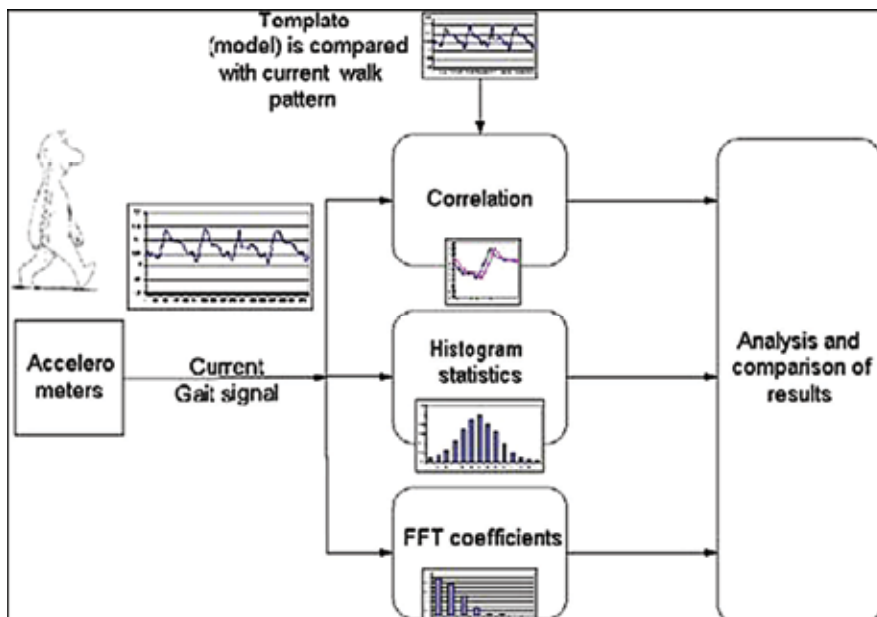
Fig. 11. Block diagram of a gait based identification method. Reprinted from Proceedings of 2005 30th IEEE International Conference on Acoustics, Speech and Signal Processing, Mäntyjärvi et al., 2005, Identifying Users of Portable Devices from Gait Pattern with Accelerometers, pp. 973-976, with permission from IEEE

Mobile phone based biometrics uses the acceleration signal characteristics produced by walking for verifying the identity of the users of a mobile phone while they walk with it. This identification method is by nature unobtrusive, privacy preserving and controlled by the user, who would not at all be disturbed or burdened while using this technology. The principle of identifying users of mobile phones from gait pattern with accelerometers is presented in Fig. 11. In this scenario, the three-dimensional movement produced by walking is recorded with the accelerometers within a mobile phone worn by the user. The collected data is then processed using correlation, frequency domain methods and data distribution statistics. Experiments show that all these methods provide good results (Mäntyjärvi et al., 2005).

The challenges of the method come from effect of changes in shoes, ground and the speed of walking. Drunkenness and injuries also affect performance of gait recognition. The effect of positioning the mobile phone holding the accelerometers in different places and positions also remains to be studied in future.

2.5 ECG biometrics for mobile phone based telecardiology

Cardiovascular disease (CVD) is the number one killer in many nations of the world. Therefore, prevention and treatment of cardiovascular disorders remains its significance in global health issues.

With the development of telemedicine, mobile phone based telecardiology has been technologically available for real-time patient monitoring (Louis et al., 2003; Sufi et al., 2006; Lee et al., 2007; Lazarus, 2007; Chaudhry et at., 2007; Plesnik et al., 2010), which is becoming increasingly popular among CVD patients and cardiologists. In a telecardiology application, the patient's Electrocardiographic (ECG) signal is collected from the patient's body which can be immediately transmitted to the mobile phone (shown in Fig. 12) using wireless communication and then sent through mobile networks to the monitoring station for the medical server to perform detection of abnormality present within the ECG signal. If serious abnormality is detected, the medical server informs the emergency department for rescuing the patient. Prior to accessing heart monitoring facilities, the patient first needs to log into the system to initiate the dedicated services. This authentication process is necessary in order to protect the patient's private health information. However, the conventional user name and password based patient authentication mechanism (as shown in Fig. 13) might not be ideal for patients experiencing a heart attack, which might prevent them from typing their user name and password correctly (Blount et al., 2007). More efficient and secured authentication mechanisms are highly desired to assure higher survival rate of CVD patients.

Recent research proposed an automated patient authentication system using ECG biometric in remote telecardiology via mobile phone (Sufi & Khalil, 2008). The ECG biometrics, basically achieved by comparing the enrollment ECG feature template with an existing patient ECG feature template database, was made possible just ten years ago (Biel et al., 2001) and has been investigated and developed by a number of researchers (Shen et al., 2002; Israel et al., 2005; Plataniotis et al., 2006; Yao & Wan, 2008; Chan et al., 2008; Fatemian & Hatzinakos, 2009; Nasri et al., 2009; Singh and Gupta, 2009; Ghofrani & Bostani, 2010; Sufi et al., 2010b). The common features extracted from ECG signals contain three major feature waves (P wave, T wave and QRS complex) as shown in Fig. 14. The use of this sophisticated ECG based biometric mechanism for patient identification will create a seamless patient authentication mechanism in wireless telecardiology applications.
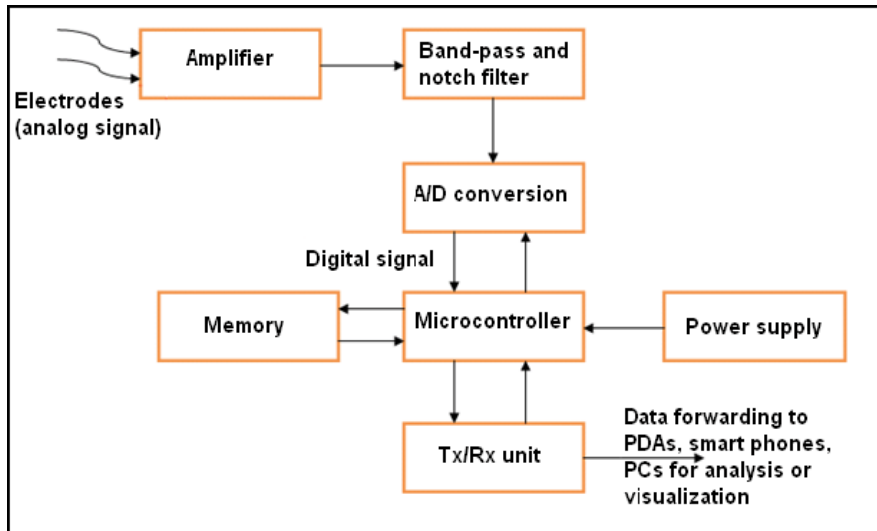
Fig. 12. Architecture of an ECG acquisition and remote monitoring system. Reprinted from Proceedings of 2010 15th IEEE Mediterranean Electrotechnical Conference, Plesnik et al., 2010, ECG Signal Acquisition and Analysis for Telemonitoring, pp. 1350-1355, with permission from IEEE

Fig. 13. Username and password based authentication mechanism for mobile phone dependent remote telecardiology. Reprinted from Proceedings of 2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Sufi & Khalil, 2008, An Automated Patient Authentication System for Remote Telecardiology, pp. 279-284, with permission from IEEE

In the proposed system, the patient's ECG signal is acquired by a portable heart monitoring device, which is capable of transmitting ECG signals via Bluetooth to the patient's mobile

phone. The mobile phone directly transmits the compressed and encrypted ECG signal to the medical server using GPRS, HTTP, 3G, MMS or even SMS. Upon receiving the compressed ECG, the original ECG of the patient is retrieved on the medical server through decompression and decryption. Then the medical server performs extraction of ECG feature template and matches the template against the ECG biometric database. The patient identification is achieved after the closest match is determined.
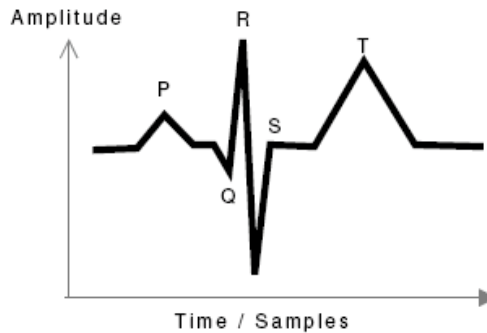
Fig. 14. Typical ECG feature waves (Sufi et al., 2010a)

In a later research (Sufi and Khalil, 2011), a novel polynomial based ECG biometric authentication system (as shown in Fig. 15) was proposed to perform faster biometric matching directly from compressed ECG, which requires less storage for storing ECG feature template. The system also lowered computational requirement to perform one-to-many matching of biometric entity.
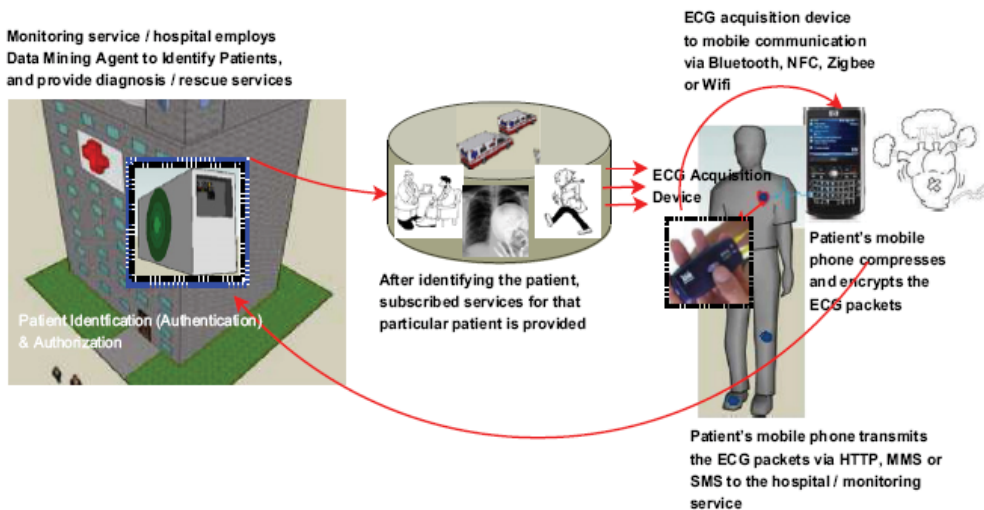
Fig. 15. Architecture of patient identification from compressed ECG based on data mining. Reprinted from Journal of Network and Computer Applications, Vol. 34, Sufi & Khalil, 2011, Faster Person Identification Using Compressed ECG in Time Critical Wireless Telecardiology Applications, pp. 282–293, with permission from Elsevier

With this new ECG biometric authentication mechanism in place, the CVD patients log into the medical server and then have access to the monitoring facility without human intervention and associated delays, making the telecardiology application faster than existing authentication approaches, which eventually leads to faster patient care for saving life.

Challenges for this ECG based biometric system involve the security of transmitting ECG from the patient to the medical server for privacy protection and the pertinence of ectopic beats, the presence of which either with the enrolment ECG or the recognition ECG could result in possible false non-match for a patient.

2.6 Summary and discussion on different systems

There are many more types of biometrics that could be implemented on mobile phones in addition to the above systems introduced in this section. Generally, several key factors should be considered when implementing such biometrics within a mobile phone. These factors will include user preference, accuracy and the intrusiveness of the application process. Table 1 illustrates how these factors vary for different types of biometrics.

Biometric technique	User preference from survey	Sample acquisition capability as standard?	Accuracy	Non-intrusive?
Ear shape recognition	NA	✘	High	✔
Facial recognition	Medium	✔	High	✔
Fingerprint recognition	High	✘	Very high	✘
Hand geometry	Medium	✘	Very high	✘
Handwriting recognition	NA	✔	Medium	✔
Iris scanning	Medium	✘	Very high	✘
Keystroke analysis	Low	✔	Medium	✔
Service utilization	NA	✔	Low	✔
Voiceprint verification	High	✔	High	✔

Table 1. Comparison of different biometric techniques for mobile phone. Reprinted from Computers & Security, Vol. 24, Clarke & Furnell, Authentication of Users on Mobile Telephones - A Survey of Attitudes and Practices, pp. 519-527, 2005 with permission from Elsevier

The user preference is investigated in a survey (Clarke et al., 2003). The assigned accuracy category is based upon reports by the International Biometric Group (IBG, 2005) and National Physical Laboratory (Mansfield et al., 2001). The judgement of intrusiveness is performed according to whether or not the biometrics could be applied transparently.

It could be seen that apparent disparity exists between high authentication security and transparent authentication process. Biometric approaches that have the highest accuracy are also the more intrusive techniques. When implementing biometrics on mobile phones, a compromise between security and the convenience to the user is required.

3. Open issues with biometrics on mobile phone

Biometrics on mobile phone, as an emerging frontier, is very promising while still holding many technical problems to be well addressed in order to be widely and ideally adopted. Issues worth pursuing in future research not only involve biometrics and mobile phones alone, but also come with the applications and styles of implementation i.e. scenarios in which specific biometrics are used.

3.1 Issues with biometrics

The most critical issue with biometrics that needs continuous effort to work on is to recognize biometric patterns with higher accuracy. A biometric system does not always make absolutely right decisions, it can make two basic types of errors, the false match and false non-match. Error rates of typical biometrics are shown in Table 2. Correspondingly, Table 3 lists requirements on typical accuracy performance. It is apparent that there is still a large gap between the currently available technology and requirements of performance.

Biometric	FTE %	FNMR %	FMR1 %	FMR2 %	FMR3 %
Face	n/a	*4*	*10*	*40*	12
Finger	*4*	*2*	*2*	*0.001*	<1
Hand	2	1.5	1.5	n/a	n/a
Iris	7	6	<0.001	n/a	n/a
Voice	1	*15*	*3*	n/a	n/a

Table 2. Typical biometric accuracy performance numbers reported in large third party tests. FTE refers to failure to enroll, FNMR is non-match error rate, FMR1 denotes verification match error rate, FMR2 and FMR3 denote (projected) large-scale identification and screening match error rates for database sizes of 1 million and 500 identities, respectively. Reprinted from IEEE publication title: Proceedings of 2004 17th International Conference on Pattern Recognition, Jain et al., 2004, Biometrics: A Grand Challenge, pp. 935-942, with permission from IEEE

Application	FNMR%	FMR%
Authentication	0.1	0.1
Large Scale Identification	0.001	0.0001
Screening	1.0	0.0001

Table 3. Typical intrinsic matcher (1:1) performance requirements. Reprinted from Proceedings of 2004 17th International Conference on Pattern Recognition, Jain et al., 2004, Biometrics: A Grand Challenge, pp. 935-942, with permission from IEEE

Other problems that need to be further studied include assurance of infeasibility of fraudulence and exploration of new features with existing biometrics and novel types of biometrics. Moreover, as computing power of current mobile phones is still very limited, processing methods of biometric patterns need to be adapted for lower burden on computation.

3.2 Challenges to mobile phone

In order to ensure the accuracy and efficiency of biometrics recognition on mobile phones, computing power and storage capacity of mobile phones are still needed to be significantly enhanced. Currently, the implementation of biometrics on mobile phones usually requires the simplification of algorithm used in conventional biometrics in order to be adapted for the relatively small CPU processing power of a cellular phone. This adaption will inevitably reduce the accuracy and security level, which highly limits the performance of mobile phone enabled biometric techniques.

In addition, the essential hardware i.e. biometric sensors embedded on mobile phones are also required to provide better performance, e.g. higher resolution of image acquired with digital cameras on mobile phones, at lower cost while maintaining their miniaturization feature.

3.3 Optimal implementation of biometrics on mobile phone

Reasonable implementation of biometrics on mobile phone is important for wide adoption of this technology as the application of mobile phone based biometrics must work in a non-intrusive manner for the convenience of users. Examples of feasible scenarios are described as keystroke analysis while texting messages, handwriting recognition while using transcriber function and speaker recognition whilst using microphones (Clarke & Furnell, 2005). Another problem needs to be addressed is the compatibility with multiple platforms of mobile phones during the development of algorithms and software.

3.4 Outlook of future development in mobile phone based biometrics

Numerous types of biometrics hold the potentials of being implemented on mobile phones. According to the different types of signals needed to be collected for feature extraction, applicable biometric methods can be classified into the imaging type, mechanical type and electrical type.

The imaging type includes, but is not limited to the recognition of face, teeth and palm print in addition to fingerprint and iris, utilizing images captured by the camera embedded in the mobile phone. The mechanical type involves voice, heart sound using microphones and blood pressure by specific and miniaturized sensors attached to the mobile phone. Not only ECG can be used in mobile biometrics, the electroencephalography (EEG) identification (Paranjape et al., 2001; Nakanishi et al., 2009; Bao et al., 2009) also has applicability in this new area.

The mobile phone based biometrics is also developing towards a multimodal functionality, which combines several biometric recognition methods to provide more reliable and flexible identification and authentication.

Promising applications include personal privacy security, e-commerce, mobile bank transactions, e-health technology, etc. A grand outlook of future development in mobile phone based biometrics is outlined in the diagram below (Fig. 16).
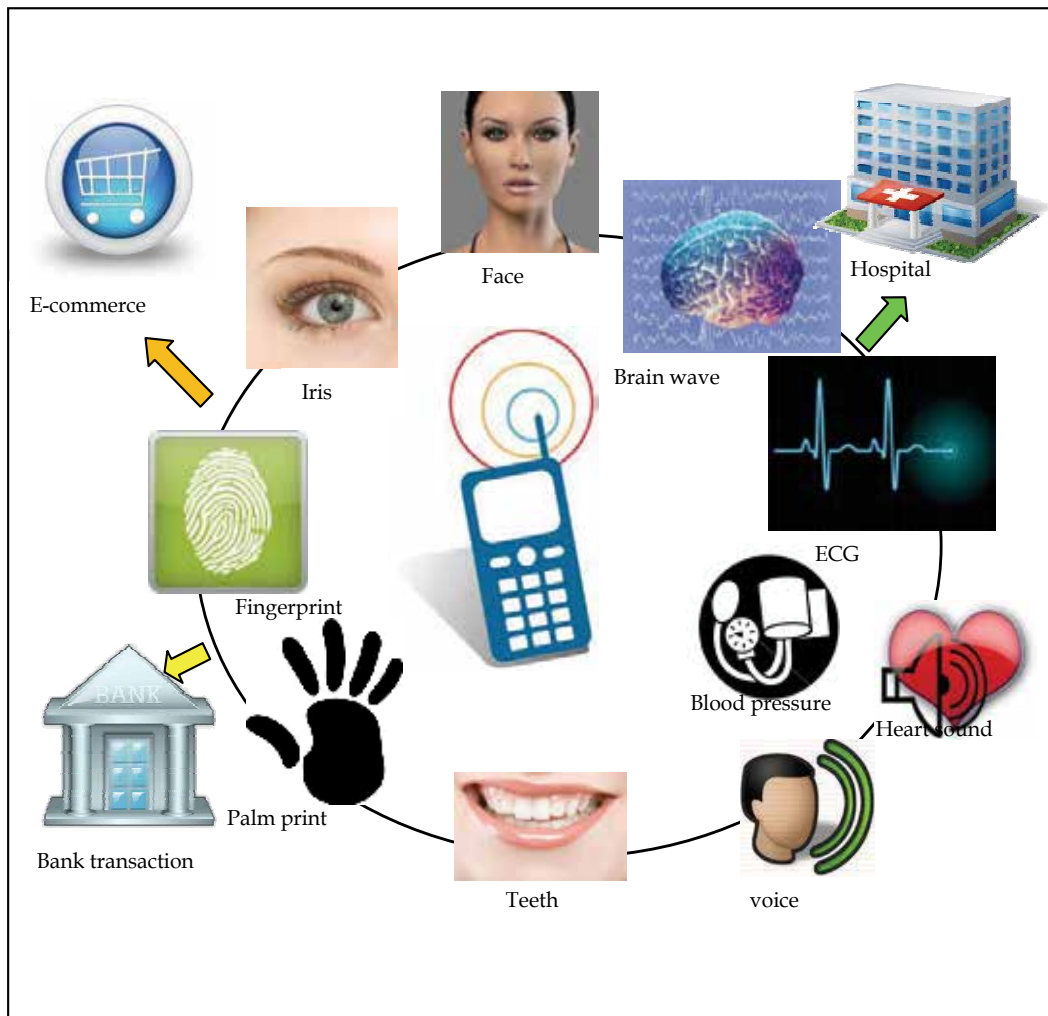
Fig. 16. An outline of future development in biometrics on mobile phone

4. Conclusion

In this chapter, we study how the mobile phone can be used in biometrics. This versatile technique has so far proven to be a unique and promising participant in the areas of biometrics. Not only can mobile phone deliver successful solutions in the traditional biometric arenas of human identification and authentication, it has also been instrumental in securing the resource-constrained body sensor networks for health care applications in an efficient and practical manner. At the same time, there remain many challenges to be addressed and a lot more new technologies to be explored and developed. Before successful consumer-ready products are available, a great deal of research and development is still needed to improve all aspects of the mobile phone based biometric system. With a modicum of expectation, it is hoped that this chapter will play a part in further stimulating the research momentum on the mobile phone based biometrics.

5. Acknowledgement

6. References

Bao, X.; Wang, J. & Hu, J. (2009). Method of Individual Identification based on Electroencephalogram Analysis. *Proceedings of 2009 International Conference on New Trends in Information and Service Science*, pp. 390-393, ISBN 978-0-7695-3687-3, Beijing, P.R.China, June 9-July 2, 2009

Biel, L.; Pettersson, O.; Philipson, L. & Wide, P. (2001). ECG Analysis: A New Approach in Human Identification. *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 3, (June 2001), pp. 808-812, ISSN 0018–9456

Blount, M.; Batra, V.; Capella, A.; Ebling M.; Jerome, W.; Martin, S.; Nidd, M.; Niemi, M. & Wright, S. (2007). Remote Health-Care Monitoring Using Personal Care Connet. *IBM Systems Journal*, Vol. 46, No. 1, (Jan 2007), pp. 95-113, ISSN 0018-8670

Bours, P. & Shrestha, R. (2010). Eigensteps: A giant Leap for Gait Recognition. *Proceedings of 2010 2nd International Workshop on Security and Communication Networks*, pp. 1-6, ISBN 978-1-4244-6938-3, Karlstad, Värmland, Sweden, May 26-28, 2010

Chan, A.; Hamdy, M.; Badre, A. & Badee V. (2008). Wavelet Distance Measure for Person Identification Using Electrocardiograms. *IEEE Transactions on Instrumentation and measurement*, Vol. 57, No. 2, (February 2008), pp. 248-253, ISSN 0018–9456

Chaudhry, S.; Phillips, C.; Stewart, S.; Riegel, B.; Mattera, J.; Jerant, A. & Krumholz, H. (2007). Telemonitoring for Patients With Chronic Heart Failure: A Systematic Review. *Journal of Cardiac Failure*, Vol. 13 No. 1, (February 2007), pp. 56-62, ISSN 1071-9164

Chen, W. & Huang, J. (2009). Speaker Recognition using Spectral Dimension Features. *Proceedings of 2009 4th International Multi-Conference on Computing in the Global Information Technology*, pp. 132-137, ISBN 978-0-7695-3751-1, Cannes, La Bocca, France, August 23-29, 2009

Chen, X.; Tian, J.; Su, Q.; Yang, X. & Wang, F. (2005). A Secured Mobile Phone Based on Embedded Fingerprint Recognition Systems. In: *Intelligence and Security Informatics*, Kantor, P. et al., (Eds.), pp. 549-553, Springer Berlin / Heidelberg, Retrieved from http://dx.doi.org/10.1007/11427995_57

Cho, D.; Park, D. & Rhee, D. (2005). Real-time Iris Localization for Iris Recognition in Cellular Phone. *Proceedings of 2005 6th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and 1st ACIS International Workshop on Self-Assembling Wireless Networks*, pp. 254-259, ISBN 0-7695-2294-7, Towson, Maryland, USA, May 23-25, 2005

Cho, D.; Park, K.; Rhee, D.; Kim, Y. & Yang, J. (2006). Pupil and Iris Localization for Iris Recognition in Mobile Phones. *Proceedings of 2006 7th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, pp. 197-201, ISBN 0-7695-2611-X, Las Vegas, Nevada, USA, June 19-20, 2006

Clarke, N. & Furnell, S. (2005). Authentication of Users on Mobile Telephones - A Survey of Attitudes and Practices. *Computers & Security*, Vol. 24, No. 7, (October 2005), pp. 519-527, ISSN 0167-4048

Daugman, J. (2003). The Importance of Being Random: Statistical Principles of Iris Recognition. *Pattern Recognition*, Vol. 36, No. 2, (February 2003), pp. 279-291, ISSN 0031-3203

Daugman, J. (2004). How Iris Recognition Works. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 14, No. 1, (January 2004), pp. 21-30, ISSN 1051-8215

Fatemian, S. & Hatzinakos, D. (2009). A New ECG Feature Extractor for Biometric Recognition. *Proceedings of 2009 16th International Conference on Digital Signal Processing*, pp. 1-6, ISBN 978-1-4244-3298-1, Santorini-Hellas, Fira, Greece, July 5-7, 2009

Ghofrani, N. & Bostani, R. (2010). Reliable Features for an ECG-based Biometric System. *Proceedings of 2010 17th Iranian Conference of Biomedical Engineering*, pp. 1-5, ISBN 978-1-4244-7484-4, Isfahan, Isfahan, Iran, November 3-4, 2010

Israela,S.; Irvine, J.; Cheng, A.; Wiederhold, M.& Wiederhold, B. (2005). ECG to Identify Individuals. *Pattern Recognition*, Vol. 38, No. 1, (January 2005), pp. 133 – 142, ISSN 0031-3203

Jain, A.; Pankanti, S.; Prabhakar, S.; Hong, L. & Ross, A. (2004). Biometrics: A Grand Challenge. *Proceedings of 2004 17th International Conference on Pattern Recognition*, pp. 935-942, ISBN 0-7695-2128-2, East Lansing, Michigan, USA, August 23-26, 2004

Kang, J. (2010). Mobile Iris Recognition Systems: An Emerging Biometric Technology. *Procedia Computer Science*, Vol. 1, No. 1, (May 2010), pp. 475-484, ISSN 1877-0509

Kinnunen, T.; Karpov, E. & Fränti, P. (2006). Real-Time Speaker Identification and Verification. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 14, No. 1, (January 2006), pp. 277-288, ISSN 1558-7916

Kounoudes, A.; Kekatos, V. & Mavromoustakos, S. (2006). Voice Biometric Authentication for Enhancing Internet Service Security. *Proceedings of 2006 2nd International Conference on Telecommunication Technology and Applications*, pp. 1020-1025, ISBN 0-7803-9521-2, Damascus, Syria, April 24-28, 2006

Lazarus, A. (2007). Remote, Wireless, Ambulatory Monitoring of Implantable Pacemakers, Cardioverter Defibrillators, and Cardiac Resynchronization Therapy Systems: Analysis of a Worldwide Database. *Pacing and Clinical Electrophysiology*, Vol. 30, No. S1, (January 2007), pp. S2-S12, ISSN 1450-8159

Lee, R.; Chen, K.; Hsiao, C. & Tseng, C. (2007). A Mobile Care System With Alert Mechanism. *IEEE Transactions on Information Technology in Biomedicine*, Vol. 11, No. 5, (September 2007), pp. 507-517, ISSN 1089-7771

Louis, A.; Turner, T.; Gretton, M.; Baksh, A. & Cleland, J. (2003). A Systematic Review of Telemonitoring for the Management of Heart Failure. *The European Journal of Heart Failure*, Vol.5 , No. 5, (October 2003), pp. 583–590, ISSN 1388-9842

Mäntyjärvi, J.; Lindholm, M.; Vildjiounaite, E.; Mäkelä, S. & Ailisto, H. (2005). Identifying Users of Portable Devices from Gait Pattern with Accelerometers. *Proceedings of 2005 30th IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 973-976, ISBN 0-7803-8874-7, Philadelphia, Pennsylvania, USA, March 18-23, 2005

Motwani, R.; Dascalu, S. & Harris, F. (2010). Voice Biometric Watermarking of 3D Models. *Proceedings of 2010 2nd International Conference on Computer Engineering and Technology*, pp. 632-636, ISBN 978-1-4244-6347-3, Chengdu, Sichuan, P.R.China, April 16-18, 2010

Nakanishi, I.; Baba, S. & Miyamoto, C. (2009). EEG Based Biometric Authentication Using New Spectral Features. *Proceedings of 2009 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 651-654, ISBN 978-1-4244-5015-2, Kanazawa, Ishikawa, Japan, December 7-9, 2009

Nasri, B.; Guennoun, M. & El-Khatib, K. (2009). Using ECG as a Measure in Biometric Identification Systems. *Proceedings of 2009 IEEE Toronto International Conference - Science and Technology for Humanity*, pp. 28-33, ISBN 978-1-4244-3878-5, Toronto, Ontario, Canada, September 26-27, 2009

OKI Introduces Japan's First Iris Recognition for Camera-equipped Mobile Phones and PDAs, In: *OKI Press Releases*, 27.11.2006, Available from http://www.oki.com/en/press/2006/z06114e.html

Paranjape, R.; Mahovsky, J.; Benedicenti, L. & Koles, Z. (2001). The Electroencephalogram as a Biometric. *Proceedings of 2001 Canadian Conference on Electrical and Computer Engineering*, pp. 1363-1366, ISBN 0-7803-6715-4, Toronto, Ontario, Canada, May 13-16, 2001

Plataniotis, K.; Hatzinakos, D. & Lee, L. (2006). ECG Biometric Recognition Without Fiducial Detection. *Proceedings of 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pp. 1-6, ISBN 978-1-4244-0487-2, Baltimore, Maryland, USA, September 19-21, 2006

Plesnik, E.; Malgina, E.; Tasič, J. & Zajc, M. (2010). ECG Signal Acquisition and Analysis for Telemonitoring. *Proceedings of 2010 15th IEEE Mediterranean Electrotechnical Conference*, pp. 1350-1355, ISBN 978-1-4244-5793-9, Valletta, Malta, April 26-28, 2010

Rohs, M. (2005). Real-World Interaction with Camera Phones, In: *Ubiquitous Computing Systems*, Murakami, H.; Nakashima, H.; Tokuda, H. & Yasumura, M., (Eds.), pp. 74-89, Springer Berlin / Heidelberg, Retrieved from http://dx.doi.org/10.1007/11526858_7

Shen, T.; Tompkinsl, W. & Hu, Y. (2002). One-Lead ECG for Identity Verification. *Proceedings of 2002 2nd Joint Conference of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society*, pp. 62-63, ISBN 0-7803-7612-9, Houston, Texas, USA, October 23-26, 2002

Singh, Y. & Gupta, P. (2009). Biometrics Method for Human Identification Using Electrocardiogram. In: *Advance in Biometrics*, Tistarelli, M. & Nixon, M., (Eds.), pp. 1270-1279, Springer Berlin / Heidelberg, Retrieved from http://dx.doi.org/10.1007/978-3-642-01793-3_128

Sufi, F.; Fang, Q; Mahmoud, S. & Cosic, I. (2006). A Mobile Phone Based Intelligent Telemonitoring Platform. *Proceedings of the 3rd IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors*, pp. 101-104, ISBN 0-7803-9787-8, Boston, USA, September 4-6, 2006

Sufi, F. & Khalil, I. (2008). An Automated Patient Authentication System for Remote Telecardiology. *Proceedings of 2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 279-284, ISBN 978-1-4244-2957-8, Sydney, Australia, December 15-18, 2008

Sufi, F.; Khalil, I. & Tari, Z. (2010a). A Cardiod based Technique to Identify Cardiovascular Diseases using Mobile Phones and Body Sensors. *Proceedings of 2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5500-5503, ISBN 978-1-4244-4123-5, Buenos Aires, Argentina, August 31 - September 4, 2010

Sufi1, F.; Khalil, I. & Habib, I. (2010b). Polynomial Distance Measurement for ECG Based Biometric Authentication. *Security and Communication Networks*, Vol. 3, No. 4, (August 2010), pp. 303 – 319, ISSN 1939-0114

Sufi, F. & Khalil, I. (2011). Faster Person Identification Using Compressed ECG in Time Critical Wireless Telecardiology Applications. *Journal of Network and Computer Applications*, Vol. 34, No. 1, (January 2011), pp. 282–293, ISSN 1084-8045

Tanviruzzaman, M.; Ahamed, S.; Hasan, C. & O'brien, C. (2009). ePet: When Cellular Phone Learns to Recognize Its Owner. *Proceedings of the 2009 2nd ACM Workshop on Assurable and Usable Security Configuration*, pp. 13-17, ISBN 978-1-60558-778-3, Chicago, Illinois, USA, November 9, 2009

Tseng, D.; Mudanyali, O.; Oztoprak, C.; Isikman, S.; Sencan, I.; Yaglidere, O. & Ozcan, A. (2010), Lensfree Microscopy on a Cellphone. *Lab on a Chip*, Vol. 10, No. 14, (July 2010), pp. 1782-1792, ISSN 1473-0197

Wang, H. & Liu, J. (2009). Mobile Phone Based Health Care Technology. *Recent Patents on Biomedical Engineering*, Vol. 2, No. 1, (January 2009), pp. 15-21, ISSN 1874-7647

Woodward, J. & Orlans, N. (2003). Esoteric Biometrics, In: *Biometrics: Identity Assurance in the Information Age*, Gatune, J., (Ed.), pp. 115-136, McGraw-Hill Professional Publishing, ISBN 0-07-222227-1, Berkeley, California, USA

Xie, Q. & Liu, J. (2010). Mobile Phone Based Biomedical Imaging Technology: A Newly Emerging Area. *Recent Patents on Biomedical Engineering*, Vol. 3, No. 1, (January 2010), pp. 41-53, ISSN 1874-7647

Yao, J. & Wan, Y. (2008). A Wavelet Method for Biometric Identification Using Wearable ECG Sensors. *Proceedings of the 2008 5th International Workshop on Wearable and Implantable Body Sensor Networks, in conjunction with The 5th International Summer School and Symposium on Medical Devices and Biosensors*, pp. 297-300, ISBN 978-1-4244-2253-1, Hong Kong, P.R.China, June 1-3, 2008

Real-Time Stress Detection by Means of Physiological Signals

Alberto de Santos Sierra, Carmen Sánchez Ávila, Javier Guerra Casanova
and Gonzalo Bailador del Pozo
Group of Biometrics, Biosignals and Security
Universidad Politécnica de Madrid
Spain

1. Introduction

The incessant demand of security in modern society is requiring a certain effort on providing protected and reliable frames for contemporary scenarios and applications such as bank account access, electronic voting, commerce or border crossing frontiers in airports.

Biometrics is of essential importance due to their capability to identify individuals univocally with low rates in false alarms, aiming to avoid the use of passwords, pin-codes or different tokens for personal identification. Instead, biometrics claim to extract precise and unique information from inidividuals based on whether behavioural or physical characteristics. In fact, there are a wide range of possible techniques for biometric identification, whose enumeration is far beyond the scope of this topic.

However, despite of avoiding the use of pin-codes, biometrics do not consider the case of individuals being forced to provide the biometric data to the corresponding sensor, allowing non-desired accesses. In other words, given a cash withdraw machine in a bank provided with the most sofisticated biometric system able to detect even fake or non-living samples, if a person is forced to present the required biometric data (iris, fingerprint, hand, . . .), the system would let enter that person, as long as the biometric template coincides with the acquired data. Thus, individuals registered or enrolled within the systems could be used as keys to access a complex door.

The presented approach proposes a stress detection system able to cope with this lack of security, based on the fact that former situations take place provoking a huge response in the human stress mechanism. Such response is impossible to disguise, providing a suitable method to detect anomalous situations in where the whole security could be compromised.

This stress detection must provide precise and real-time information on the state-of-mind of the individual, requiring a low number of physiological parameters to keep the acquisition system as less invasive and intrusive as possible. Notice that this fact is an essential concern due to the current misgivings on hygienic considerations.

Therefore, only two phsyiological signals are required, namely Galvanic Skin Response (Skin Conductivity) and Heart Rate, since both provide accurate and precise information on the physiological situation of individuals. The inclusion of adequate sensors for both signals acquisitions require little hardware, being straightforward to include former sensors in current biometric systems.

Besides, this chapter proposes a wide variety of methods for stress detection, in order to elucidate which method is more suitable for implementation and integration in future biometric devices. In addition, methods are oriented for real-time applications, which in most cases provoke a reduction in stress detection accuracy.

Finally, the study comes up with the conclusion that best approach combining accuracy and real-time application is based on fuzzy logic, modelling the behaviour of individuals under different stressing and non-stressing situations, creating a stress template gathering previous physiological information.

The use of the proposed stress template is twofold: On the one hand, to collect and gather the different behaviour of each individual under a variety of situations in order to compare posterior physiological acquisitions. On the other hand, the idea of template implies modelling each individual separately, providing a frame to distinguish to what extent individuals react against stressing situations. This template is based on the idea that human individuals react differently to a same event, and therefore, a stress detection system cannot provide a result based on general parameters but concrete, personal and individualize features.

2. Literature review

The problem of stress detection has been tackled with different approaches. However, former works can be divided into two different groups, depending on the use of physiological signals or other behavioural characteristics.

For example, the work presented by Andren & Funk (2005) provides a system able to compute the stress level of an individual by the manner and rhythm in which a person types characters on a keyboard or keypad. In addition, Dinges et al. (2007) provides a study of stress detection based on facial recognition. Both approaches are related to behavioural human characteristics. On the other hand, there exist many previous works related to stress detection based on physiological signals. The essay presented by Begum et al. (2006) presents a study of stress detection only based on Finger Temperature (FT), together with Fuzzy Logic Zadeh (1996), and Case-Based Reasoning Andren & Funk (2005).

Focusing on stress detection by means of physiological signals, it is necessary to describe which possible signals can be related to stress and their extent.

It is not common to focus only on one certain physiological feature, but on many of them, in order to obtain further and more precise information about the state of mind. Considering this multimodal approach, there are several articles which study a variety of parameters and signals, as well as the combination among them.

Heart Rate variability (HR) has been considered as an earlier stress marker in human body, being widely studied and analyzed. Several authors consider this signal in their reports: Jovanov et al. (2003) presented a stress monitoring system based on a distributed wireless architecture implemented on intelligent sensors, recording HR along different positions in individual body by means of sensors located beneath clothes.

In additoin, the research provided in Angus et al. (2005); Zhai et al. (2005) proposes a system considering Finger Temperature (FT), Galvanic Skin Response (GSR) and Blood Volume Pulse (BVP). The main characteristic of this system lies on the fact that signals are acquired in a non-intrusive manner and furthermore, these previous physiological signals provide a predictable relation with stress variation.

There exist physiological signals of different nature like Pupil Dilation (PD) and Eyetracking (ET) providing very precise information about frame stress. When an individual is

Physiological Signals	References
BVP (Blood Volume Pressure)	Barreto & Zhai (2006); Picard & Healey (2000)
	Lin et al. (2005); Zhai et al. (2005)
GSR (Galvanic Skin Response)	Barreto & Zhai (2006); Picard & Healey (2000)
	Lin et al. (2005); Moore & Dua (2004); Zhai et al. (2005)
PD (Pupil Dilation)	Barreto & Zhai (2006); Lin et al. (2005); Zhai et al. (2005)
ST (Skin Temperature)	Barreto & Zhai (2006); Zhai & Barreto (2006)
ECG, EKG (Electrocardiogram)	Picard & Healey (2000)
Breath (RR)	Picard & Healey (2000)
EMG (Electromyogram)	Picard & Healey (2000)
EEG (Electroencephalogram)	Picard & Healey (2000)

Table 1. Literature Review on physiological signals involved in stress detection.

under stress, PD is wider and the eye movement is faster. The article presented in Prendinger & Ishizuka (2007), not only consider PD and ET, but also GSR, BVP and FT. The main purpose of this approach is to recognize emotions, interest and attention from emotion recognition, a very remarkable conclusion for future computer applications and for the improvement of Human Computer Interaction (HCI) Kim & Ande (2008); Sarkar (2002a). In summary, stress can be detected through many different manners, as stated in Sarkar (2002a), where a wide study is carried out regarding previous physiological signals together with others related to stress (Positron Emission Technology (PET) Healey & Picard (2005); Sarkar (2002a), Functional Magnetic Resonance Imaging (fMRI) Picard & Healey (2000); Sarkar (2002b), Electroencephalography (EEG) Li & hua Chen (2006); Sarkar (2002a), likewise Electromyograms (EMG) Chin & Barreto (2006a); Li & hua Chen (2006); Shin et al. (1998) or Respiratory Rate (RR) Shin et al. (2004)). Nonetheless, these other signals lack of future integrity because they involve more invasive acquisition procedures.

Table 1 gathers a summary on the signals involved in stress detection within literature.

Together with signal processing and feature extraction, the comparison algorithms to elucidate the stress level of an individual are of great importance. There are some previous work considering several approaches for stress detection. The work presented by N. Sarkar Sarkar (2002a) proposes fuzzy logic (as M. Jiang and Z. Wang Jiang & Wang (2009)) to elucidate to what extent a user is under stress. On the other hand, the research presented by A. de Santos et al. de Santos Sierra et al. (2011) proposes the creation of a fuzzy stress template to which subsequent physiological acquisitions could be compared and contrasted. Other approaches have been proposed, based on different techniques like, SVM, *k*-NN, Bayes classifier. In order to extend excesively the document, Table 2 contains a summary of previous approaches within literature.

Finally, a matter of importance are both how stress is induced in individuals and the number of samples to evaluate former approaches. Table 3 and Table 4 briefly show which experiments have been involved for provoking stress and which populations were required in order to validate stress detection algorithms. More extensiveley, the research by Lisetti & Nasoz (2004) provides a complete study on emotion recognition including a deep literature review on the experiments carried out to provoke emotions considering populations, algorithms, approaches and so forth.

Moreover, special mention deserves the work presented by Healey & Picard (2005), since they are considered pioneers on stress detection field.

Algorithms	References
SVM (Support Vector Machines)	Barreto & Zhai (2006); Zhai et al. (2005)
ANOVA Analysis	Lin et al. (2005)
Bayes classifier	Zhai & Barreto (2006)
Fisher Analysis	de Santos Sierra et al. (2010); Picard & Healey (2000)
k-NN	de Santos Sierra et al. (2010)
Fuzzy Logic	Begum et al. (2006); Sarkar (2002b)
	de Santos Sierra et al. (2011); Sarkar (2002a)

Table 2. Literature Review on algorithms applied to stress detection.

Experiments	References
Stroop Test	Barreto & Zhai (2006); Zhai et al. (2005)
	Zhai & Barreto (2006)
Videogames	Lin et al. (2005)
Driver and Pilot Simulation	Picard & Healey (2000)
Hyperventilation and Talk Preparation	de Santos Sierra et al. (2011; 2010)

Table 3. Literature Review on experiment layouts oriented to provoke stress.

Populations	References
6 male individuals	Zhai et al. (2005)
42 adults with ALS[1]	Moore & Dua (2004)
14 males and 4 females	Lin et al. (2005)
32 individuals	Barreto & Zhai (2006); Zhai & Barreto (2006)
3 experienced drivers	Healey & Picard (2005)
10 pilots (with and without experience)	Healey & Picard (2005)

Table 4. Literature Review on populations involved in stress detection evaluation.

3. Physiological signals

Although several possible signals have been considered within the literature to detect stress (Section 2), this paper proposes the use of two signals: Galvanic Skin Response (GSR), also known as Skin Conductance (SC), and Heart Rate (HR). These two signals were selected based on their properties regarding non-invasivity when being acquired and because their variation is strongly related to stress stimuli Barreto & Zhai (2006); Healey & Picard (2005); Prendinger & Ishizuka (2007).

Galvanic Skin Response (GSR), known also as electrodermal activity (EDA), is an indicator of skin conductance Barreto & Zhai (2006); Shi et al. (2007). More in detail, glands in the skin produce ionic sweat, provoking alterations on electric conductivity. First experiment dates back to 1907, when Carl Jung first described some relation between emotions and the response of this parameter Angus et al. (2005); Zhai et al. (2005).

GSR can be obtained by different methods, but the device proposed to acquire signals (Section 4.1) is based on an exosomatic acquisition. In other words, extracting skin conductivity requires a small current passing through the skin. GSR is typically acquired in hand fingers and its measure units are μSiemens ($\mu\Omega^{-1}$) Angus et al. (2005).

Main parameters of GSR like basis threshold, peaks or frequency variation vary enormously among different individuals and thus, no general features can be extracted from GSR signals

Fig. 1. A GSR (Galvanic Skin Response signal) sample during the four stages: First Base Line (BL1), Talk Preparation (TP), Hyperventilation (HV) and Second Base Line (BL2). Notice how GSR arousal responds positively to stressing stimuli (HV and TP).

for a global stress detection purpose, since parameters extracted from GSR signals are strongly related to each individual.

Figure 1 shows an original GSR signal, measured during the experiments. Reader may notice the different arousal of this signal, depending on the stressing stimulus. Initials in Figure 1 stands for BL1 (Base Line 1), TP (Talk Preparation), HV (Hyperventilation) and BL2 (Base Line 2) whose meanings are extensively explained in Section 4.5.

On the other hand, Heart Rate (HR) measures the number of heartbeats per unit of time. HR can be obtained at any place on the human body, being an accessible parameter to be easily acquired Choi & Gutierrez-Osuna (2009); Jovanov et al. (2003).

HR describes the heart activity when the Autonomic Nervous System (ANS) attempts to tackle with the human body demands depending on the stimuli received Picard & Healey (2000). Concretely, ANS react against a stressing stimulus provoking an increase in blood volume within the veins, so rest of the body can react properly, increasing the number of heartbeats.

Most common methods for HR extraction consider to measure the frequency of the well-known QRS complex in a electrocardiogram signal Bar-Or et al. (2004); Sharawi et al. (2008). In contrast to ECG biometric properties Israel et al. (2005), HR is not distinctive enough to identify an individual.

Summarizing, both HR and GSR behave differently for each individual, and therefore posterior stress template must gathered properly this unique response in order to obtain an accurate result in stress detection. Figure 2 shows an original HR signal (measured in Beats per Minute, BPM), measured during the experiments. Reader may notice the different arousal of this signal, depending on the stress stimuli. Initials in Figure 2 stands for BL1 (Base Line 1), TP (Talk Preparation), HV (Hyperventilation) and BL2 (Base Line 2) whose meanings are extensively explained in Section 4.5.1.

4. Database acquisition

This section provides an overview of how the dataset was built considering the experimental setup and the characteristics of the database and which psychological tests were carried out to assess in which manner an individual is likely to react against stress situations Yanushkevich et al. (2007).
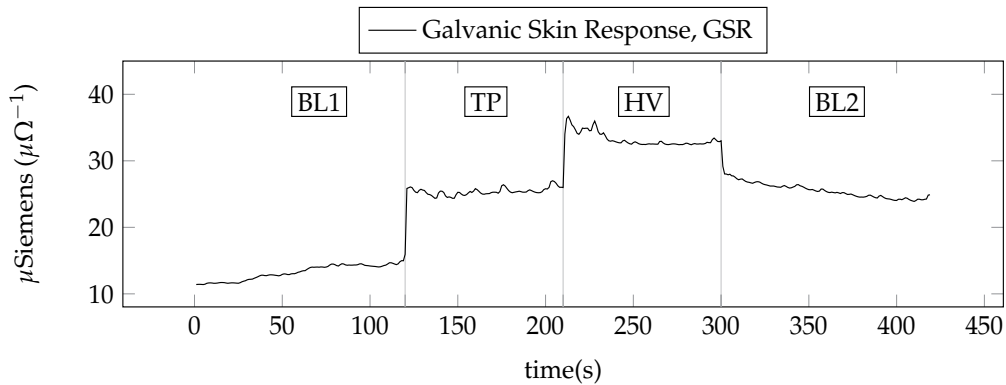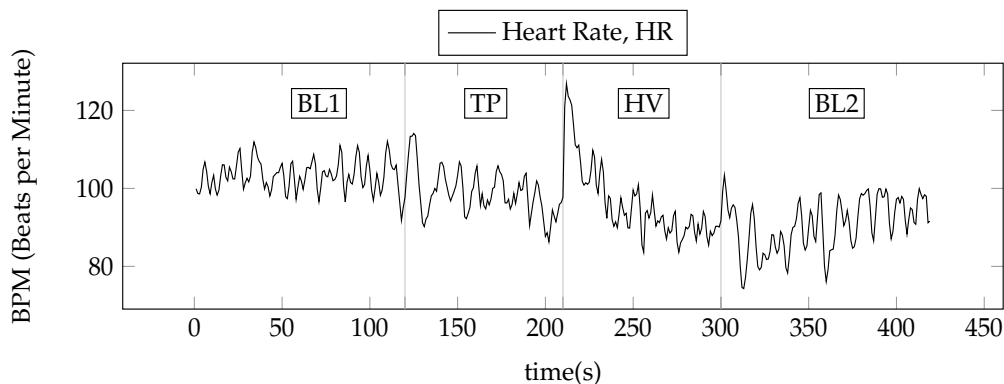
Fig. 2. A HR (Heart Rate signal) sample during the four stages: First Base Line (BL1), Talk Preparation (TP), Hyperventilation (HV) and Second Base Line (BL2).

4.1 Overview

The experiments were carried out in a Faraday room in the Human Psychology Laboratory from Psychology Faculty of Complutense University of Madrid (UCM), endowed with electromagnetic, thermal and acoustic insulation, with the aim of collecting HR and GSR signals from each participant.

The device proposed to carry out these experiments is I-330-C2 PHYSIOLAB (J &J Engineering) able to process and store 6 channels including EMG (Electromyography), ECG (Electrocardiogram), RR (Respiration Rate), HR and GSR. Sensors were attached to hand right (or left, but not both) fingers Cai & Lin (2007), wrist and ankle, in order to acquire both HR and GSR, avoiding sensors detachments, unplugged connectors to analog-to-digital converter and/or software acquisition errors. Moreover, sample acquisition rate is one sample per second for both HR and GSR.

4.2 Participants

The participants were students from Psychology Faculty (UCM) and Social Work (UCM), being a total of 80 female individuals, with ages from 19 to 32 years old, with an average of 21.8 years old and a standard deviation of 2.15. The lack of male individuals is due to the Faculty where the experiment took place, given the fact that the percentage of male students is almost negligible in comparison to the amount of females.

4.3 Task justification

Provoking stress on an individual requires a specific experimental design in order to obtain an adequate arousal to the proposed physiological signal Dinges et al. (2007); Healey & Picard (2005). Concretely, this paper proposes to induce stress by using Hyperventilation and Talk Preparation Cano-Vindel et al. (2007).

Hyperventilation (HV) is defined as a certain kind of breath, which exceeds standard metabolic demands, as a result of excess in respiratory rhythm.

As a consequence, several physiological changes emerge: arterial pressure diminution in blood until a certain level so-called hypocapnea Cano-Vindel et al. (2007); Zvolensky & Eifert (2001), and blood pH increment, known as alkalosis.

However, voluntary hyperventilation does not produce always an actual anxiety reaction Cano-Vindel et al. (2007), and therefore, an additional anxiogenic task is required to ensure

that a positive valence in terms of stress response is provoked. Such a task is Talk Preparation (TP).

Results provided by Cano-Vindel et al. (2007); Zvolensky & Eifert (2001) highlight the fact that hyperventilation produces a physiological reaction similar to that reaction induced by a threatening task of preparing a talk.

As a conclusion, talk preparation and hyperventilation provoke both an alteration in physiological parameters together with different emotional experiences. These previous tasks have been widely studied and evaluated with positive results, and they are very suitable to induce stressing stimuli on individuals.

4.4 Tests

When performing psychological experiments, an important tool to validate results requires the utilization of tests to extract subjective information from the individual.

There exist several tests able to provide information about the predisposition of a certain individual to be affected by anxiety and stress: ISRA test (Inventory of Situations and Responses of Anxiety Miguel-Tobal & Cano-Vindel (2002)), IACTA test (Cognitive Activity Inventory on Anxiety Disorders Cano-Vindel & Miguel-Tobal (2001)) and ASI (Anxiety Sensitivity Index Peterson & Reiss (1992)).

The two former tests were developed by Cano-Vindel (Professor of Faculty of Psychology, Complutense University of Madrid) together with his research group, and have been widely used and accepted by scientific community Cano-Vindel & Miguel-Tobal (2001); Cano-Vindel et al. (2007).

The tests used to detect predisposition to anxiety are described as follows:

- ISRA: Inventory designed according to the model of three response systems which evaluates signs of anxiety in a Cognitive level (C), Physiological level (P), and Motor (M). Furthermore, the addition of this three values provides a general measure of anxiety level, denoted by Total (T). ISRA test also includes to assess tendency to show anxiety in four areas, namely Assessment situations (F1), Interpersonal situations (F2), Phobics situations (F3) and Daily life situations (F4).

  This test provides good psychometrics properties, excellent internal consistency (Cronbach's alpha .99), good test-reset reliability (.81 for T) and good capability in discriminating different samples.

- IACTA: This test is able to measure sub-scales of social phobia, panic attacks and agoraphobia.

- ASI: This index provides information about the fear to anxiety symptom. Three different factors are measured: Physical worries, social worries and thoughts related to mental handicap. Subjects with a high total score in ASI show more elevated levels of anxiety after hyperventilation task compared to subjects with a low score. However, physical worries is the main factor to predict the anxiety level after hyperventilation in both no-clinic subjects and panic disorders patients Cano-Vindel et al. (2007); Zvolensky & Eifert (2001).

4.5 Procedure

The experiments were split into two sessions:

- First session regards a subject selection, applying ISRA, IACTA and ASI test.

- Second session consisted of the HR and GSR sample acquisition under hyperventilation and talk preparation.

4.5.1 First session

Several collective assessments were carried out applying ISRA, IACTA and ASI test. Participants were to fulfill previous tests in the same order as described previously. The average length of this step was deemed to be about 50 minutes.

Finally, two groups (namely Group 1 and Group 2) were created ensuring that the distribution of their respective anxiety levels, measured by psychological tests Cano-Vindel & Miguel-Tobal (2001); Miguel-Tobal & Cano-Vindel (2002), were similar. In other words, this selection seeks to avoid one group containing people which barely react against stress, and other group with people which overreact under stressing conditions. Therefore, both groups must be well-balanced in terms of anxiety levels in order to validate the experiments.

Participants from Group 1 underwent an experimental session using physiological and subjective signals under following conditions: calm state (base line, namely BL1), stimulating task (hyperventilation, HV), threatening task (talk preparation, TP) and base line post-stress (BL2). On the other hand, the order of tasks was swapped for participants from Group 2: calm state (base line), threatening task (talk preparation), stimulating task (hyperventilation) and base line post-stress. Main reason to alter the order consists of making independent the task order from the results obtained Cano-Vindel et al. (2007); Miguel-Tobal & Cano-Vindel (2002). The emotional experience was assessed after each situation (Base line (BL1), threatening (TP) and stimulating task (HV) and base line post-stress (BL2)), using a Likert scale (0-100) Cano-Vindel et al. (2007).

Specifically, individuals were asked to evaluate the following emotional parameters: displeasure, anxiety level, corporal sensations and thoughts lack of control. In fact, control dimension was divided into two variables to explore possible differences among previous facets. Lack of control on observable behavior was not assessed, since the subjects were strongly indicated not to move during the experiment procedure, in order to avoid noise in physiological signal acquisition (Section 4.1).

Furthermore, a precise order was given to avoid ambiguities regarding emotions: 'Assess anxiety level/intensity experimented at this very precise moment'. Usually, participants did not understand the meaning of evaluating their emotions (which emotion?).

After this four parameters evaluation (i.e., displeasure, anxiety level, corporal sensations and thoughts lack of control), next step was carried out, consisting of recording Heart Rate (HR) and Galvanic Skin Response (GSR), together with new evaluations of previous four parameters (displeasure, anxiety level, corporal sensations and thoughts lack of control) regarding emotional experience.

The experimental session consisted of the following steps for participants from Group 1:

- Sensors location and adaptation time. After this adaptation time (variable time), base line of physiological signals (heart rate and skin conductivity) were taken under rest situation, during 2 minutes. Once the signals were recorded, a new assessment of emotional parameters was carried out. Besides, this step is so-called BL1.
- Hyperventilation task (HV), consisting of deep and fast breathes each 3 seconds, conducted by a sound produced by the experimenter. This task was performed till the individual realized clearly of changes in his or her corporal sensations. The participant let the experimenter know that moment, by the use of a simple word. The experimenter made the participant to breath deeply three times more, recording after that moment physiological signals (90 seconds), followed by a new evaluation of the emotional parameters.

- Talk preparation task facing an audience (TP). The experimenter asked the individual to prepare mentally a talk of few minutes on a certain topic explained during the lectures that these students attended, in order to be recorded by a video-camera. After three minutes, new appraisal of emotional parameters were inquired, recording again HR and GSR (90 seconds), informing the participant that the talk was not necessary anymore.

- Rest period (BL2). The experiment comes to an end, and post-stress base line is recorded during 2 minutes, acquiring HR and GSR signals, together with a new subjective evaluation of the emotional parameters.

Obviously, BL1 implies no stressing stimuli on the individual in contrast to HV and TP. However, nothing can be assured in relation to BL2, since it cannot be considered neither as a stressing nor as a relaxing state Cano-Vindel & Miguel-Tobal (2001); Cano-Vindel et al. (2007); Miguel-Tobal & Cano-Vindel (2002).

Individuals from Group 1 carried out the talk preparation task (TP), after base line state (BL1), followed by the hyperventilation task (HV), ending with post-stress base line (BL2). On the other hand, individuals from Group 2 performed the experiment in the following order: BL1, TP, HV and BL2.

4.6 Database discussion

A biometric database based on a certain physical characteristic, e. g. Iris, consists of different samples taken from a wide range of users during different sessions separated by a lapse of time of days, weeks or even months. On the contrary, the database gathered in these experiments does not verify any of the previous points described above.

This stress database consists of a unique sample of a very specific set of individuals, namely female students with ages on the interval 19 to 32 years old, with an average of 21.8 years old and a standard deviation of 2.15 (Section 4.2).

However, there exists justification for this drawback. A psychological experiment is far from being repeatable, since the specific tasks previously described (Section 4.3) require a component of surprise and unexpectedness. In other words, if an individual carries out again the same tasks, even after an undefined period of time, such person would be prepared to come through the task, and what is more, the response of her or his physiological signals will not be certainly the same Cano-Vindel et al. (2007).

Obviously, a third session similar to second one (first session consisted of answering tests ISRA, IACTA and ASI, second session attempted to register the physiological signals) could have taken place with different tasks. Though, this option was rejected, since different tasks provoke different stress responses Chin & Barreto (2006b); Fairclough (2009); Kim & Ande (2008); L. et al. (2007), and therefore, the signals registered in both sessions would not correspond to same degrees of stress.

The physiological response to a stressing agent is strongly related to each individual and such a response is similar, independently of the time during the stressing stimulus provoked the response Yerkes & Dodson (1908). As an overview, the stress mechanism could be considered as a linear temporal invariant system, which provides certain outputs depending on the inputs. Therefore, same inputs produce same outputs. Moreover, stress mechanism extracts some information from the stimuli, so that if such stressing agent appears again, the human body is able to react faster and better, compared to first time Lin et al. (2005); Rosch (1996). This characteristic makes useless to repeat same tasks after a certain period of time, and furthermore makes unnecessary a third session with different tasks, since the response will not be the same, as the stimuli provided by different task, provoke different responses. Then,

this implies that these experiments assure that a final stress detection system is able to detect stress in real applications, despite of being train with the presented database.

Finally, it is difficult, even for expert psychologists, to state whether the response among female and male individuals differs as much as previous response varies within female individuals Sapolsky (1988). Several researchers support the idea that male and female individuals suffer different responses when stress agent endures through time, (e.g., a great amount of work at job, a bad economical situation, and so forth), but they have similar responses when stress stimuli consist of specific actions in a very short period of time, e.g. an accident, an armed robbery and the like Lin et al. (2005).

Thereby, it is justified to extend the results obtained with this database to a wider population, even containing males and females. Nonetheless, the algorithm responsible for detecting stress has been implemented independently of this likely drawback, since it considers how an individual behaves under both stress and relax situation. This procedure provides in theory more independence from database population.

5. Template extraction

Before describing the different approaches compared within this paper, how the template is created is explained, combining both the physiological signals and the different stressing and non-stressing tasks.

In other words, the template extraction is required so that the system could create a profile in order to contrast, whether a user is actually under stress. This template is based on specific characteristics extracted from individual concerning parameters from the physiological signals HR and GSR.

On the other hand, once the user is associated to a template, the individual is able to access the system, and therefore a template comparison is required. Both steps are described in following sections.

First step consists of extracting a stress template from the user. Such a template describes the behavior of HR and GSR signals in both situations calm state (relax) and excited state (stress). As stated in Section 4.1, HR and GSR signals are recorded by I-330-C2 PHYSIOLAB (J &J Engineering) able to process and store 6 channels including EMG, ECG, RR, HR and GSR.

The main aim of a stress detection system consists of elucidating calm state (relax) or excited state (stress). Therefore, the system must know how both signals (HR and GSR) behave in both situations. Since these states cannot be controlled easily by an individual, calm state and excited state must be induced while HR and GSR are recorded.

Each user must undergo the experiments described previously (Section 4.5). Briefly, as an overview, there exist four stages in the experiments:

- First stage (BL1): Elicit a relax state by suggesting the individual to sit comfortably.

- Second stage (HV): Hyperventilation, i.e., deep and fast respirations.

- Third stage (TP): Speech/talk preparation.

- Forth stage (BL2): Relax state.

Three states emerge from previous stages (Section 4.5): Calm state (First stage, BL1), Excited state (Second and Third stage, HV and TP) and post-excited state (Forth stage, BL2). This latter state regards the fact that after a stress short period of time, HV and GSR require more time to achieve a calm state. Thereby, forth stage and first stage diverge in terms of HR and GSR despite of corresponding to the same instructions in the experiment.
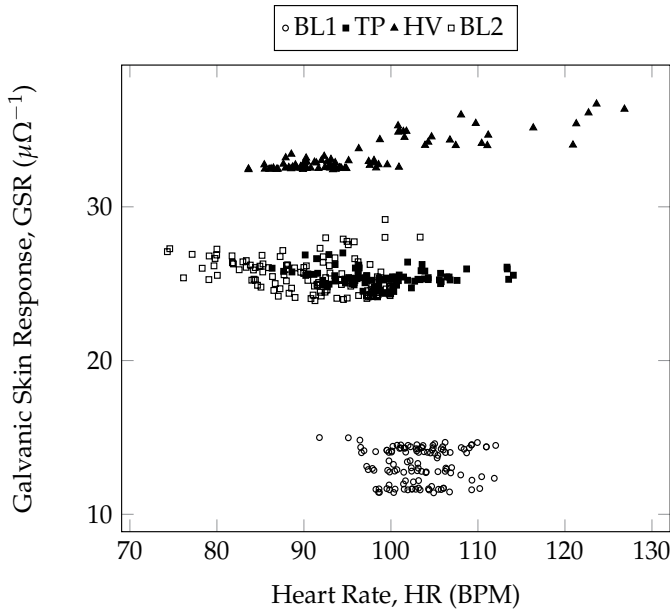
Fig. 3. Graphic representation of γ. Notice how the relation between HR and GSR varies depending on the stressing stimuli (BL1, TP, HV and BL2).

The experiments involve Hyperventilation (HV) and Talk Preparation (TP), as presented in Section 4.3. However, any task involving a considerable cognitive load (such mathematical operations, color distinction, Stroop test and so forth) may come in useful for inducing stress in a similar manner as previous task met such a required goal Cano-Vindel et al. (2007); Conway et al. (2000); Healey & Picard (2005); Kim & Ande (2008).

Mathematically, both HR and GSR are considered as stochastic signals. Therefore, $\mathcal{H}$ represents the space of HR possible signals and $\mathcal{G}$ represents the space of GSR possible signals. Each stage will come up with a pair of signals $h \in \mathcal{H}$ and $g \in \mathcal{G}$ according to the experimental task conducted in each situation. Thus, a template extraction requires four pair of signals, namely $\gamma = [(h_1, g_1), (h_2, g_2), (h_3, g_3), (h_4, g_4)] \in \mathcal{H} \times \mathcal{G}$ corresponding to how the individual behaves under different states. Notice that signals h_i and g_i are not normalized, in contrast to previous approaches Angus et al. (2005); Healey & Picard (2005) . The decision to avoid normalization was done based on the experience, since data without normalization provided more accurate results in terms of stress detection.

Once γ is obtained, for each pair of signals, (h_i, g_i), $i = \{1, 2, 3, 4\}$, a mean vector is obtained together with the deviation for each pair. In other words, four parameters are obtained: $\zeta_{h_i} = \bar{h}_i$ and $\zeta_{g_i} = \bar{g}_i$, which represent the mean of signals h_i and g_i in addition to σ_{h_i} and σ_{g_i} related to the dispersion for each pair. Finally, stress template, namely $\mathcal{T}$ is described by $\mathcal{T} = (\zeta_{h_i}, \zeta_{g_i}, \sigma_{h_i}, \sigma_{g_i})$, $i = \{1, 2, 3, 4\}$.

Figure 3 provides a visual example of a scattering representation of each pair of signals γ. Notice how non-stressing stimuli provokes a low excitation in GSR (Figure 3, ∘), and on the contrary, the evidence of an arousal when undergoing on stressing tasks like Talk Preparation (Figure 3, ▪) and Hyperventilation (Figure 3, ▴).

The aim of this action is to described the information in HR and GSR by four Gaussian distributions, centered in $(\zeta_{h_i}, \zeta_{g_i})$ and with standard deviation σ_{h_i} and σ_{g_i}. This approach
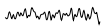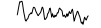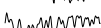
Task	HR	GSR	ζ_h	ζ_g	σ_h	σ_g
BL1			103.2	13.22	3.84	1.13
TP			98.9	25.28	5.74	1.13
HV			96.14	33.18	9.5	1.31
BL2			90.67	25.60	6.69	1.32

Table 5. Parameters extracted from GSR and HR signals in relation to experimental task (BL1, TP, HV and BL2). Those pieces of signals (columns HR and GSR) have been extracted from Figure 2 and Figure 1 respectively.

will facilitate the implementation of fuzzy antecedent membership functions by Gaussian distributions in a posterior fuzzy decision algorithm. This approach will facilitate a system access (section 6) implementation based on fuzzy logic, able to provide a more accurate decision on the degree of stress of a certain individual.

Furthermore, Table 5 provides a visual example of how previous parameters ζ_{h_i}, ζ_{g_i}, σ_{h_i} and σ_{g_i}, $i = \{1, 2, 3, 4\}$ are extracted from signals HR and GSR (Figure 1 and Figure 2) depending on the experimental task (BL1, TP, HV and BL2).

Let $t_\mathcal{T}$ be the time used to acquire both signals in order to extract the stress template. Evidently, the performance of the system depends on this parameter, since the longer $t_\mathcal{T}$, the more information the system obtains, and therefore, the stress template may be more accurate. A study regarding this relation between $t_\mathcal{T}$ and system performance is presented in Section 7.3.

Finally, after template extraction, the template must be stored. The template $\mathcal{T}$ requires 16×32 bits, since each template element (whatever ζ_{h_i}, ζ_{g_i}, σ_{h_i} or σ_{g_i}), is represented by a float element. In other words, 512 bits, i.e. 64 Bytes.

6. Stress detection

After a stress template extraction, $\mathcal{T}$ describes how an individual behaves under stressing and non-stressing situations. This section describes how the stress detection procedure is performed in addition to an overview on the algorithm involved to elucidate on the degree of stress. Once the user is enrolled, and a template is created, describing how such a user behaves under different stressing situations, the subject is able to access the system. In other words, the system is able to decide whether a certain registered user is under stressing stimuli. First requirement for a stress detection system access regards user identification and verification. The individual attempting to access the system, must be firstly identified so that the system can load his/her template, $\mathcal{P}$. This step is indispensable, otherwise, the system could not contrast the information (resulting from a HR and GSR acquisition) presented by the user.

Firstly, the signals GSR and HR must be measured from the individual. This acquisition process lasts a variable time, t_{acq} (acquisition time), responsible for the performance of the overall system, in addition to $t_\mathcal{T}$. In fact, the main difference between t_{acq} and $t_\mathcal{T}$ relies on the fact that $t_\mathcal{T}$ is related to the required time to obtain template $\mathcal{T}$ and t_{acq} regards the time needed to decide to what extent an individual is under stress. Both are measured in seconds, and main aim of posterior expert system consists of obtaining highest accuracy in detecting stress by requiring shortest time of t_{acq} and $t_\mathcal{T}$.

This compromise will be discussed in Section 7.3.

This document proposes five different approaches to solve stress detection, given a template and two physiological signals: GMM McLachlan & Basford (1988), *k*-Nearest Neighbour (*k*-NN) Nilsson (1996), Fisher's linear discriminant analysis Michie et al. (1994), Support vector machines (SVMs) Wang (2009) and Fuzzy Logic Zadeh (1996).

6.1 Gaussian mixture model

Let $\mathbf{x}$ be a two-dimensional observation describing a sample of both GSR and HR. The probability density function of $\mathbf{x}$ in the finite mixture form is expressed in Eq. 1 and Eq. 2,

$$p(\mathbf{x}; \phi_c) = \sum_{i=1}^{K} \pi_i g(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{1}$$

$$g(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{2\pi |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)} \tag{2}$$

where K is the number of mixtures, the parameter $\phi_c = \{\pi_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K}$ consists of the mixture weight π_i ($\Sigma \pi_i = 1$), the mean vector $\boldsymbol{\mu}_i$ and the covariance matrix $\boldsymbol{\Sigma}_i$ of the ith Gaussian component $\forall i = 1, 2, \ldots, K$, in the c class. In fact, Eq. 2 is a specific case with $d = 2$ McLachlan & Basford (1988); Wolfe (1970), since there are only two physiological signals, HR and GSR.

The parameters represented by ϕ_c are estimated by applying the Expectation Maximization (EM) algorithm Wolfe (1970). Let $\{\mathbf{x}^t\}_{t=1}^{N}$ be the training samples, then EM algorithm finds

$$\phi_c^* = \text{argmax} \, \Pi_{t=1}^{N} P(\mathbf{x}^t | \phi_c) \tag{3}$$

6.2 *k*-Nearest neighbour

The *k*-Nearest Neighbour (*k*-NN) is a simple method used for density estimation. The probability of a point x' falling within a volum V centred at a pont x is given by the followinf relation (Equation 4):

$$\theta = \int_{V(x)} p(x) dx \tag{4}$$

where the integral is carried out over the volume V. This integral can be approximated by the relation $\theta \sim p(x)V$ when V is small.

In fact, the probability θ may be approximated by the proportion of samples falling within V, so that $\theta \sim \frac{k}{n}$.

The distance metric used in *k*-NN methods can be described by a simple Euclidean distance. In other words, given two patterns $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$, then the distance is given by $\sqrt{\sum_{j=1}^{n}(x_j - y_j)^2}$.

A deeper understanding of *k*-NN is provided in Nilsson (1996).

6.3 Fisher linear discriminant analysis

Fisher's linear discriminant analysis is an empirical method for classification based purely on attribute vectors Michie et al. (1994). A hyperplane in th *p*-dimensional attribute space is chosen to separate the known classes as accurate as possible. Points are then classifed according to the side of the hyperplane that they fall on.

More precisely, in the case of two classes, let $\bar{x}, \bar{x}_1, \bar{x}_2$ be respectively the means of the attribute vectors overall and for the two classes. Suppose that a coefficient set $a_1, \ldots, a_p$ is given, obtaining the particular linear combination attributes $g(x) = \sum a_j x_j$, which is called the discriminant between the classes.

The criterion proposed by Fisher is the ratio of between-class to within-class variances. Formally, we seek a direction w such that:

$$J_F = \frac{|w^T(m_1 - m_2)|^2}{w^T S_W w}$$

is maximum, where m_1 and m_2 are the group means and S_W is the pooled within-class sample covariance matrix, in its bias-corrected form given by $\frac{1}{n-2}\left(n_1\hat{\Sigma}_1 - n_2\hat{\Sigma}_2\right)$

where $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ are the maximum likelihood estimates of the covariance matrices of classes ω_1 and ω_2 respectively, with n_i samples in class ω_i.

6.4 Support vector machines

Support vector machines (SVMs) have been widely applied to pattern classification problems and non-linear regressions Wang (2009).

After SVM classifiers are trained, they can be used to predict future trends.

In this case, the supervised classification that involves two steps: firstly, an SVM is trained as a classifier with a part of the data in a specific data set. In the second step (i.e., prediction), we use the classifier trained in the first step to classify the rest of the data in the data set.

The SVM is a statistical learning algorithm pioneered by Vapnik Vapnik (1995). The basic idea of the SVM algorithm is to find an optimal hyper-plane that can maximize the margin (a precise definition of margin will be given later) between two groups of samples. The vectors nearest to the optimal hyper-plane are called support vectors.

In comparison with other algorithms, SVMs have shown outstanding capabilities in dealing with classification problems.

6.5 Fuzzy logic

An automatic decision algorithm must elucidate a certain output accordingly to specific inputs. There exist several kinds of decision strategies in relation to the concrete task to solve Andren & Funk (2005); Jiang & Wang (2009); Liao et al. (2005).

The decision algorithm proposed in this article is based on fuzzy logic Zadeh (1996), since it is a very suitable strategy considering the data involved. In other words, a fuzzy decision algorithm provides an indefinite output allowing, however, to achieve a more precise decision than by using crisp decision algorithms Begum et al. (2006); Picard & Healey (2000); Sarkar (2002a;b).

As a main description, a stress fuzzy decision algorithm attempts to elucidate to what extent a certain individual is under stressing stimuli, by capturing his or her physiological HR and GSR signals during t_{acq} seconds, and comparing his or her response with a previous stored template $\mathcal{T}$, obtained by a prior acquisition carried out during $t_{\mathcal{T}}$ seconds. It is on that template comparison where this decision algorithm focuses on.

Any fuzzy decider requires different elements, which are described in subsequent points:

- **Antecedent membership functions**. The antecedent membership functions attempt to represent the information extracted from input, and they constitute in fact the template $\mathcal{T}$ itself.

Two different groups of antecedent membership functions are required by this system. Both groups of functions describes how HR and GSR behave under previous four situations (BL1, TP, HV, BL2).

- **Consequent membership functions**. The aim of these membership functions consists of describing the output of the system. This paper proposes an output on the interval $[0, 1]$, where 0 represents relax state and 1, stress state.

- **Rule description**. The rules describing how to transform the information provided in antecedent membership functions into consequent functions is maybe the most important part of a fuzzy decision system Pedrycz (1994). Rules provide a method to combine properly the information supplied by previous membership functions in order to produce an output stating to what extent an individual is under stress.

7. Results

This section aims at comparing the results provided by former approaches: GMM, k-NN, Fisher Discriminant Analysis, SVM and Fuzzy Logic, considering previous parameters: threshold ρ_{th} and temporal parameters ($t_{\mathcal{T}}$ and t_{acq}).

7.1 Database: Training, validation and testing data
In order to obtain valid results, the database must be divided into three groups:

- Training data: Used to extract the template, i.e., $\mathcal{T} = (\zeta_{h_i}, \zeta_{g_i}, \sigma_{h_i}, \sigma_{g_i})$.

- Validation data: Used to fixed threshold ρ_{th} and temporal parameters ($t_{\mathcal{T}}$ and t_{acq}) in order to maximize the performance of the system.

- Testing data: Used to obtain which implementation and metric is most suitable, and therefore, which is the performance of the whole system.

For each individual, a vector containing $t_{\mathcal{T}}$ seconds of γ (for each task BL1, HV, TP and BL2) was used for training data; a vector of t_{acq} seconds for each task was used to validate the system, and rest of the data was used as testing data. This latter testing data will be split in slots of t_{acq} seconds. Notice that one second corresponds to one sample in HR and GSR one-dimensional signals (Section 4.1). This assignment is done randomly. Notice that this validation scheme is similar to a K-fold cross-validation.

The justification for this division is based on the research carried out by Picard & Healey (2000), where several physiological signals (not only HR and GSR) were recorded during a period of time of thirty-two days in a same person. Eight emotions were provoked during thirty minutes per day, and no substantial changes were appreciated during that period in each emotions. In other words, physiological signals behave similarly in each task through time, and therefore h and g signals can be divided into smaller parts, considering each segment as an independent acquisition.

7.2 Stress evaluation parameters
A stress detection system must reach a compromise between detecting properly which individuals are under stress situations, and which individuals are in a relax state.

Thereby, two assessment parameters are defined:

- True Stress Detection rate (TSD): When the system properly detects stress when an individual is under stress stimuli. This TSD factor corresponds to the sensitivity statistical measure, since TSD can be described as follows in Eq. 5:

$$\text{TSD} = \frac{\#\text{True Positives}}{\#\text{True Positives} + \#\text{False Negatives}} \tag{5}$$

   where a True Positive means classifying as stressed an individual which is indeed under stress, and False Negative means classifying as relaxed an individual which is under stressing situations.

- True Non-Stress Detection rate (TNSD): When the system correctly detects no stress in an individual and the subject is indeed not under stressing situations. This TNSD factor corresponds to the specificity statistical measure, since TNSD can be described by Eq. 6:

$$\text{TNSD} = \frac{\#\text{True Negatives}}{\#\text{True Negatives} + \#\text{False Positives}} \tag{6}$$

   where a True Negative means classifying as non-stressed an individual which is not under stress, and False Positive means classifying as stressed an individual which is calm and relaxed.

Obviously, TSD and TNSD depend strongly on threshold ρ_{th}. If $\rho_{th} \rightarrow 0$ then the system considers every output as a stress stimuli (TSD decreases, TNSD increases) and vice versa.

Therefore, a compromise must be achieved by finding a threshold ρ_{th} where TSD equals TNSD. This threshold is defined as True Equal Stress Detection rate (TESD).

Notice that the higher TESD, the more accurate the performance of the system.

At this point, one question arises: Which is the best indicator (TESD,TSD or TNSD) to provide an evaluation on the performance of a stress detection system? TESD is obtained with validation data, and therefore threshold ρ_{th} and temporal parameters $t_{\mathcal{T}}$ and t_{acq} are fixed to maximized TESD. These parameters are set *a posteriori* Yanushkevich et al. (2007). On the other hand, TSD and TNSD are obtained with testing data, i.e. TSD and TNSD give an understanding on how the system behaves with real data. Notice that previous parameters (ρ_{th}, $t_{\mathcal{T}}$ and t_{acq}) have been already fixed and adapted with validation data, and therefore the performance of the system might be barely unbalanced. In other words, TSD and TNSD will not be equal at TESD, but TSD could increase in expense of TNSD or vice versa.

This suggests that TESD is a fine system performance indicator, since it provides an approximation based on validation data. However, TSD and TNSD provides a real rate of the performance. Obviously, TESD cannot be always calculated in previous schemes (Section **??**), as TESD requires both stressing and non-stressing data during the training and the validation data.

7.3 Temporal parameters

The performance of the system (TSD and TNSD) not only depends on previous threshold ρ_{th} but also on two temporal parameters: Template time ($t_{\mathcal{T}}$) and Acquisition time (t_{acq}). The former time regards the required time to obtain the template, and the latter is related to the time demanded to acquire stress information from an individual.

Evidently, the longer $t_{\mathcal{T}}$ and t_{acq}, the more accurate the system is. However, in real applications, time is the most valued asset, and therefore, a balance among $t_{\mathcal{T}}$, t_{acq}, TSD and TNSD must be achieved.

These temporal parameters are fixed during the validation step, and remain constant during the testing stage.

Figure 4 provides information about how TSD varies in relation to different values of $t_{\mathcal{T}}$ and t_{acq}.
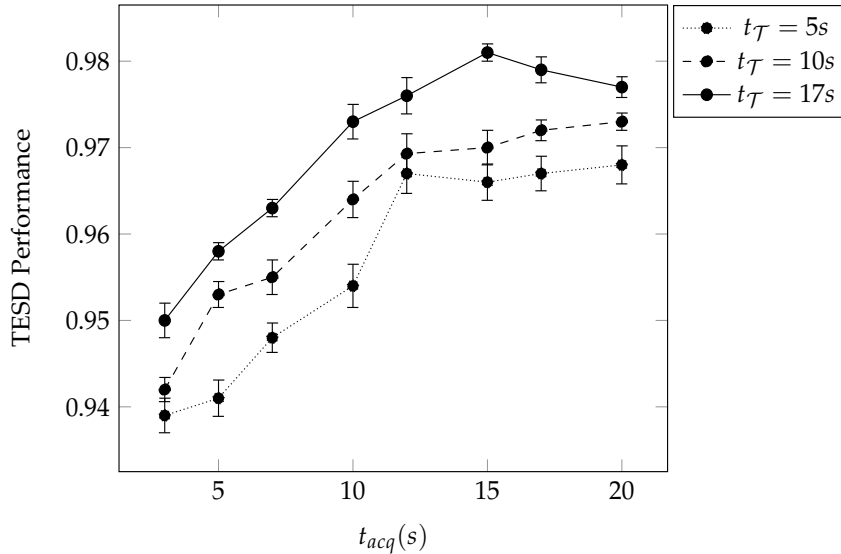
Fig. 4. Relation between TESD Performance and time to obtain template ($t_{\mathcal{T}}$) and acquisition time (t_{acq}). These values were obtained with $\rho_{th} = 0.27$.

Reader can notice how the performance in detecting stress (TSD) increases as $t_{\mathcal{T}}$ and t_{acq} do so. In fact, a TESD= 99.16% (Figure 4, —●—) is obtained with $t_{\mathcal{T}} = 17s$ and $t_{acq} = 17s$, which means that physiological signals GSR and HR from an individual are measured during $t_{\mathcal{T}} = 17s$, and furthermore, that in subsequent accesses, such an individual must present their physiological signals during $t_{acq} = 17s$, so the system can decide to what extent is under stress.

As before, the results obtained in this section are achieved for a given scheme, concretely Fuzzy Logic. In order to obtain the best combination of parameters, this procedure must be repeated for each scheme and implementation.

7.4 Evaluation performance

In order to evaluate the performance of the proposed approaches, the procedure presented previously was carried out for every system. A detailed description of these results is far beyond the scope of this section and therefore, Table 6 is presented to compared former methods when detecting stress. Reader may notice that the performance of the proposed methods depends on the temporal parameters $t_{\mathcal{T}}$ and t_{acq}, all of them measured in seconds.

The best result in every scheme is achieved with scheme BL1+HV which means that for an accurate stress detection, only two tasks are required: a relaxing situation and a stressing

	GMM ($t_\mathcal{T} = 5$, $t_{\text{acq}} = 10$)	k-NN ($t_\mathcal{T} = 5$, $t_{\text{acq}} = 5$)	Disc. Anal. ($t_\mathcal{T} = 7$, $t_{\text{acq}} = 10$)	SVM ($t_\mathcal{T} = 5$, $t_{\text{acq}} = 10$)	Fuzzy Logic ($t_\mathcal{T} = 7$, $t_{\text{acq}} = 10$)
TSD	95.1 ± 0.2	92.8 ± 0.4	95.6 ± 0.3	95.6 ± 0.4	99.5 ± 0.3
TNSD	86.3 ± 0.4	97.3 ± 1.3	96.7 ± 0.4	96.7 ± 0.3	97.4 ± 0.2

Table 6. Comparative stress detection performances. Best result is achieved with fuzzy logic, although the rest of the results are competitive when compared to those obtained within literature. Temporal parameters are provided in seconds and rates in percentage (%).

Reference	Stress Detection Rate (%)	Physiological Signals	Population
Healey & Picard (2005)	97.4%	ECG, EMG, RR, GSR	Not provided
Wagner et al. (2005)	79.5-96.6%	ECG, EMG, RR, GSR	1 subject
Cai & Lin (2007)	85-96%	BVP, ST, RR, GSR	Not provided
Guang-yuan & Min (2009)	75-85%	ECG, EMG, RR, GSR	1 subject
Kulic & Croft (2005)	76%	ECG, EMG, GSR	8 subjects
Sharawi et al. (2008)	60-78%	ST, GSR	35 subjects
Best of our proposed methods	99.5%	HR, GSR	80 subjects

Table 7. A comparison between approaches comparing stress detection rates, physiological signals and population involved. The initials ST stand for Skin Temperature.

situation. An outstanding result, since it allows to decrease (in terms of time) the template extraction step, among other aspects discussed in posterior Section 8.

The conclusion is that stress can be detected by means of fuzzy logic with an accuracy of 99.5% recording the signal of the user during 10 seconds to create the template and 7 seconds for stress detection. These results highlight the improvement achieved in comparison to other approaches, showed in Table 7, providing the following parameters to be compared: Stress Detection rate (TSD), the physiological signals involved and the population used to evaluate the proposed approach. This improvement is achieved not only in terms of accuracy in stress detection, but also in relation to the number of physiological signals (only HR and GSR) and the population.

8. Conclusions and future work

The proposed stress detection systems are able to detect stress by using only two physiological signals (HR and GSR) providing a precise output indicating to what extent a user is under a stressing stimulus.

In addition, HR and GSR allows a plausible future integration of former proposed systems on current biometric systems, achieving and increase in the overall security.

Main characteristics of the proposed systems regard non-invasiveness, fast-oriented implementation and an outstanding accuracy in detecting stress when compared to previous approaches in literature.

In other words, the system can detect stress almost instantly, allowing a possible integration in real-time systems. Notice that only two physiological signals are involved in contrast to the amount of features required to elucidate on the stress degree provided by previous approaches.

An individualization of not only the template $\mathcal{T}$, but also ρ_{th}, $t_\mathcal{T}$ and t_{acq} must be adapted for each individual, so that the overall performance can be increased.

These parameters (ρ_{th}, $t_\mathcal{T}$ and t_{acq}) have been fixed for the whole database within this work, and therefore, if a different version of these parameters is considered for each individual, then the accuracy of the system could be increased. This implementation remains as future work.

The database acquisition was based on psychological experiments carried out by expert psychologists. These experiments ensure that stressing situations are provoked on an individual, validating posterior HR and GSR acquisitions.

This paper provides a decision system able to detect stress with an accuracy of 99.5% using fuzzy logic and 10 seconds to extract the stress template and 7 seconds to detect stress on an individual using two physiological signals HR and GSR measured only during two tasks: a stressing task and a relaxing stage.

The rest of the approaches are also competitives in terms of computational cost and performance. This results are achieved due to the fact that the stress template provides precise information on the state of mind of individuals, coming up with an innovative concept in stress detection. A combination of approaches is regarded as future work.

Finally, these systems may be applicable in scenarios related to aliveness detection (e.g., detecting if an individual is accessing a biometric system with an amputated finger), civil applications (e.g., driver control), withdrawing money from a cash dispenser, electronic voting (e.g., someone is forced to emit a certain vote) and so forth. Moreover, future research entails an integration in mobile devices.

9. References

Andren, J. & Funk, P. (2005). A case-based approach using behavioural biometrics to determine a user's stress level, *ICCBR Workshops*, pp. 9–17.

Angus, F., Zhai, J. & Barreto, A. (2005). Front-end analog pre-processing for real-time psychophysiological stress measurements, *Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 05)*, pp. 218–221.

Bar-Or, A., Healey, J., Kontothanassis, L. & Van Thong, J. (2004). Biostream: a system architecture for real-time processing of physiological signals, *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, Vol. 2, pp. 3101–3104.

Barreto, A. & Zhai, J. (2006). Stress detection in computer users based on digital signal processing of noninvasive physiological variables, *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBS '06.*, pp. 1355–1358.

Begum, S., Ahmed, M. U., Funk, P., Xiong, N. & von Schï£¡ele, B. (2006). Using calibration and fuzzification of cases for improved diagnosis and treatment of stress, *in* M. Minor (ed.), *8th European Conference on Case-based Reasoning workshop proceedings*, pp. 113–122.

Cai, H. & Lin, Y. (2007). *An experiment to non-intrusively collect physiological parameters towards driver state detection*, Academic Press.

Cano-Vindel, A. & Miguel-Tobal, J. J. (2001). Iacta: Cognitive activity inventory on anxiety disorders.

Cano-Vindel, A., Miguel-Tobal, J. J., Gonzalez-Ordi, H. & Iruarrizaga-Diez, I. (2007). Hyperventilation and anxiety experience, *Anxiety and stress* 13(2–3): 291–302.

Chin, C. & Barreto, A. (2006a). Electromyograms as physiological inputs that provide efficient computer cursor control, *Proceedings of the 2006 WSEAS International Conference on Mathematical Biology and Ecology (MABE'06)*, pp. 221–226.

Chin, C. & Barreto, A. (2006b). Stress performance comparison of electromyogram-based computer cursor control systems, *WSEAS Transactions on Biology and Biomedicine*, Vol. 3, p. 118.

Choi, J. & Gutierrez-Osuna, R. (2009). Using heart rate monitors to detect mental stress, *BSN '09: Proceedings of the 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, IEEE Computer Society, Washington, DC, USA, pp. 219–223.

Conway, J. C. D., Fernandes, A. O., Coelho, C. J. N. J., Andrade, L. C. G., da Silva, D. C., J. & Carvalho, H. S. (2000). Wearable computer as a multi-parametric monitor for physiologicalsignals, *IEEE International Symposium on Bio-Informatics and Biomedical Engineering, 2000. Proceedings.*, Arlington, VA, USA, pp. 236–242.

de Santos Sierra, A., Sanchez Avila, C., Bailador del Pozo, G. & Guerra Casanova, J. (2011). A stress detection system based on physiological signals and fuzzy logic, *Industrial Electronics, IEEE Transactions on* PP(99).

de Santos Sierra, A., Sanchez-Avila, C., Guerra Casanova, J., Bailador del Pozo, G. & Jara Vera, V. (2010). Two stress detection schemes based on physiological signals for real-time applications, *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on*, pp. 364 –367.

Dinges, D. F., Venkataraman, S., McGlinchey, E. L. & Metaxas, D. N. (2007). Monitoring of facial stress during space flight: Optical computer recognition combining discriminative and generative methods, *Acta Astronautica* 60(4-7): 341 – 350. Benefits of human presence in space - historical, scientific, medical, cultural and political aspects. A selection of papers presented at the 15th IAA Humans in Space Symposium, Graz, Austria, 2005.

Fairclough, S. H. (2009). Fundamentals of physiological computing, *Interact. Comput.* 21(1-2): 133–145.

Guang-yuan, L. & Min, H. (2009). Emotion recognition of physiological signals based on adaptive hierarchical genetic algorithm, *World Congress on Computer Science and Information Engineering* 4: 670–674.

Healey, J. A. & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors, *IEEE Transactions on Intelligent Transportation Systems* 6(2): 156–166.

Israel, S. A., Irvine, J. M., Cheng, A., Wiederhold, M. D. & Wiederhold, B. K. (2005). Ecg to identify individuals, *Pattern Recognition* 38(1): 133 – 142.

Jiang, M. & Wang, Z. (2009). A method for stress detection based on FCM algorithm, *CISP'09. 2nd International Congress onImage and Signal Processing, 2009.*, Tianjin, pp. 1–5.

Jovanov, E., O'Donnell Lords, A., Raskovic, D., Cox, P. G., Adhami, R. & Andrasik, F. (2003). Stress monitoring using a distributed wireless intelligent sensor system, *Engineering in Medicine and Biology Magazine, IEEE* 22(3): 49–55.

Kim, J. & Ande, E. (2008). Emotion recognition based on physiological changes in music listening, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(12): 2067–2083.

Kulic, D. & Croft, E. (2005). Anxiety detection during human-robot interaction, *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005. (IROS 2005).*, pp. 616–621.

L., M., V., S. & E., V. (2007). Sensitive and accurate measurement environment for continuous biomedical monitoring using microelectrodes, *MEASUREMENT SCIENCE REVIEW* 7(2): 20–24.

Li, L. & hua Chen, J. (2006). Emotion recognition using physiological signals, *Advances in Artificial Reality and Tele-Existence*, Vol. 4282 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 437–446.

Liao, W., Zhang, W., Zhu, Z. & Ji, Q. (2005). A real-time human stress monitoring system using dynamic bayesian network, *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, IEEE Computer Society, Washington, DC, USA, pp. 70–78.

Lin, T., Omata, M., Hu, W. & Imamiya, A. (2005). Do physiological data relate to traditional usability indexes?, *OZCHI'05: Proceedings of the 17th Australia conference on Computer-Human Interaction*, Computer-Human Interaction Special Interest Group (CHISIG) of Australia, Narrabundah, Australia, Australia, pp. 1–10.

Lisetti, C. L. & Nasoz, F. (2004). Using noninvasivewearable computers to recognize human emotions from physiological signals, *EURASIP Journal on Applied Signal Processing:* 11: 1672ï£¡1687.

McLachlan, G. J. & Basford, K. E. (1988). *Mixture models. Inference and applications to clustering*, Statistics: Textbooks and Monographs, New York: Dekker.

Michie, D., Spiegelhalter, D. J., Taylor, C. C. & Campbell, J. (eds) (1994). *Machine learning, neural and statistical classification*, Ellis Horwood, Upper Saddle River, NJ, USA.

Miguel-Tobal, J. J. & Cano-Vindel, A. (2002). Isra: Inventory of situations and responses of anxiety.

Moore, M. M. & Dua, U. (2004). A galvanic skin response interface for people with severe motor disabilities, *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2004* pp. 48–54.

Nilsson, N. J. (1996). *Introduction to Machine Learning*.

Pedrycz, W. (1994). Why triangular membership functions?, *Fuzzy Sets Syst.* 64(1): 21–30.

Peterson, R. A. & Reiss, S. (1992). Anxiety sensitivity index manual, *International Diagnostic Systems* .

Picard, W. & Healey, J. A. (2000). Wearable and automotive systems for affect recognition from physiology, *Technical report*, MIT.

Prendinger, H. & Ishizuka, M. (2007). Symmetric multmodality revisited: Unveiling users' physiological activity, *IEEE Transactions on Industrial Electronics* 54(2): 692–698.

Rosch, P. J. (1996). *Handbook of stress, medicine, and health*, CRC Press, Inc., pp. 27–60.

Sapolsky, R. M. (1988). Individual differences and the stress response: studies of a wild primate, *Adv. Exp. Med. Biol.* pp. 399–411.

Sarkar, N. (2002a). A novel interface system for seamlessly integrating human-robot cooperative activities in space, *Technical report*, NASA Institute for Advanced Concepts.

Sarkar, N. (2002b). Psychophysiological control architecture for human-robot coordination-concepts and initial experiments, *IEEE International Conference on Robotics and Automation, 2002. Proceedings. ICRA'02.*, Vol. 4, pp. 3719–3724.

Sharawi, M. S., Shibli, M. & Sharawi, M. I. (2008). Design and implementation of a human stress detection system: A biomechanics approach, *5th International Symposium on Mechatronics and Its Applications, 2008. ISMA 2008.*, Amman, pp. 1–5.

Shi, Y., Ruiz, N., Taib, R., Choi, E. & Chen, F. (2007). Galvanic skin response (GSR) as an index of cognitive load, *CHI'07 extended abstracts on Human factors in computing systems*, ACM, New York, NY, USA, pp. 2651–2656.

Shin, J., Seongo, H., Cha, D., Yoon, Y. & Yoon, H. (1998). Estimation of stress status using biosignal and fuzzy theory, *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, Vol. 3, pp. 1393–1394.

Shin, J. W., Lee, J. S., Sul, A. R., Lee, C. G., Yoon, Y. R., Takeuch, H. & Minamitani, H. (2004). Chronic stress evaluation using neuro-fuzzy, *26th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, 2004. IEMBS'04*, Vol. 1, San Francisco, CA, pp. 373–374.

Vapnik, V. N. (1995). *The nature of statistical learning theory*, Springer-Verlag New York, Inc., New York, NY, USA.

Wagner, J., Kim, N. J. & Andre, E. (2005). From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification, *IEEE Computer Society* pp. 940–943.

Wang, L. (2009). *Data Mining with Computational Intelligence*, Springer-Verlag, Berlin, Heidelberg.

Wolfe, J. H. (1970). Pattern clustering by multivariate mixture analysis, *Multivariate Behavioral Research* 5: 329–350.

Yanushkevich, S., Wang, P., Gavrilova, M., Nixon, M. & Srihari, S. (2007). *Image pattern recognition: synthesis and analysis in biometrics*, World Scientific Pub Co Inc.

Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation, *Journal of Comparative Neurology and Psychology* pp. 459–482.

Zadeh, L. A. (1996). Fuzzy logic = computing with words, *IEEE Transactions on Fuzzy Systems* 4(2): 103–111.

Zhai, J. & Barreto, A. (2006). Stress detection in computer users through non-invasive monitoring of physiological signals, *Biomedical Science Instrumentation* 42: 495–500.

Zhai, J., Barreto, A., Chin, C. & Li, C. (2005). Realization of stress detection using psychophysiological signals for improvement of human-computer interactions, *SoutheastCon, 2005. Proceedings.*, pp. 415–420.

Zvolensky, M. J. & Eifert, G. H. (2001). A review of psychological factors/processess affecting anxious responding during voluntary hyperventilation and inhalations of carbon dioxide-enriched, *Clinical Psychological Review* 21: 375–400.

Automatic Personal Identification System for Security in Critical Services: Two Case Studies Based on a Wireless Biometric Badge

Stefano Tennina[1], Luigi Pomante[2], Francesco Tarquini[2], Roberto Alesii[2],
Fabio Graziosi[2], Fortunato Santucci[2] and Marco Di Renzo[3]
[1]*CISTER Research Unit, ISEP/IPP Rua Dr. António Bernardino de Almeida, Porto*
[2]*Dept. of Electrical and Information Engineering and Center of Excellence in Research DEWS, University of L'Aquila, Poggio di Roio, L'Aquila (AQ)*
[3]*Laboratory of Signals and Systems (L2S), CNRS - SUPELEC - Univ Paris-Sud 3 rue Joliot-Curie, Gif-sur-Yvette (Paris)*
[1]*Portugal*
[2]*Italy*
[3]*France*

1. Introduction

Due to the relevant innovations in the ICT domain, today, a lot of services is being provided with a self–service approach to an even more great number of people simplifying several tasks of everyday life (e.g., cash retrieval by ATM, remote-banking, etc.). The key aspect to fully enable such services and make them wide accepted by peopole is the possibility to reliably count on biometric identification mechanisms (Adeoye, 2010; BTAM, 2010; Elliott et al., 2007; Li & Zhang, 2010; Sonkamble et al., 2010). Some of them are already exploited in several real–life scenarios, like the Access Control–Border Management in Hong Kong or the Access Control–Restricted Area Access by Canadian Air Transport Authority (BTAM, 2010). However, some of such services could be very critical and so, their provisioning, should be managed very carefully in order to avoid the possibility of malicious operations. So, the best way to support the evolution of automatic services providing is to develop a system, also automatic, which is able to trust in a secure and flexible way the identity of people that need to access such services. Such a system should be of easy integration in several scenarios, especially with respect to existing infrastructures, and should be designed to respect all the relevant privacy issues while providing to the users all the feelings (about reliability, safety and usability) needed to make the system acceptable.

In such a context, this book chapter aims at presenting an automatic personal identification system developed by WEST Aquila (WESTAquila, 2010). The system, described in detail later, exploits the recent advances in the biometric and heterogeneous wireless networks fields to provide a true authentication platform supporting several services encompassing physical access (e.g., to restricted areas or vehicles) as well as logical access (e.g., to personal services like e-banking) management. This is realized by maintaining a full control over critical data (biometric) that are used for the authentication. In fact, the main component of the system is

a novel biometric badge, i.e., a smartcard equipped with a biometric reader (i.e., a fingerprint reader) and a short–medium range wireless transceiver which allow the identification of both the card and the card owner. In other words, it constitutes a system–on–badge: when required, the card owner is identified through an on–system biometric matching and only the result of such a matching is sent (appropriately ciphered) through the wireless interface towards the rest of the system. Therefore, personal biometric data is always under the full control of its owner, leading to high levels of security and privacy protection.

The intended content of this chapter will be to illustrate the badge as a biometric system and its usage in two case studies: (i) physical access of authorized people in a restricted area, which involves also physical positioning of the badge owner and (ii) logical access of authorized people in an e-banking–like scenario.

2. System architecture description

The proposed system is composed by a set of elements (Fig. 1) enabling high level of flexibility and reliability, needed to make this proposal as a reference in the field of personal automatic identification systems, where particular emphasis is on aspects like robustness, ease–of–use and privacy.
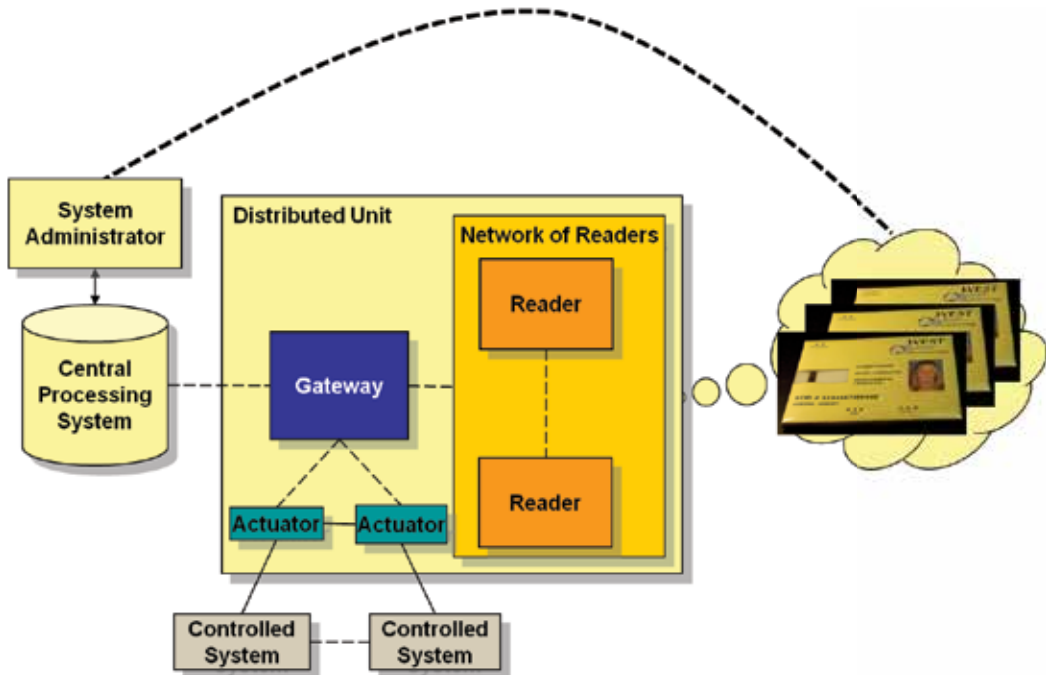
Fig. 1. Logical System Architecture

2.1 Biometric Badge (BB)

The key point of the proposed system is our embedded biometric badge (Fig. 2), which is a "system–on–badge" performing four main tasks: (i) enable the localization of its owner using distributed positioning techniques, (ii) scan and verify fingerprints of people, (iii) check if an user is the badge's owner based on fingerprint matching, and (iv) send related outcomes

wirelessly to the rest of the system (e.g., the DU which interconnects the badge to the rest of the infrastructure), without the need to transmit the owners' biometric data over the wireless medium (so, in a secure way from the point of view of transmitting critical data of the users).

Fig. 2. WEST Aquila's Biometric Badge – components

The badge is equipped with:

- the Texas Instruments' SoC CC2430 (TI, 2009), which embeds a 8051 micro controller and a CC2420 radio transceiver, compliant with the IEEE 802.15.4 (IEEE, 2006) standard. It is used for wireless medium-range communications, as well as localization operations;

- a fingerprint sensor reader with its embedded "companion chip" provided by UPEK (UPEK, 2009). This chip is the key element for handling biometric data: it allows to authenticate people based on fingerprint information, as well as store data in a memory protected even from physical external attacks. Moreover, only this chip and the gateway are aware on how to decode the messages they send to each other;

- a RFID tag based on the ISO15693 standard and its companion chip provided by Montalbano Technology (Montalbano, 2009), which allow the microcontroller to get access to data stored in the tag;

- a rechargeable battery, its driver to monitor the charge status, and a user interface with 8 leds and a push-button.

Such features allow the badge to support a full range of applications (Fig. 3), mainly due to the embedded fingerprint reader enabling both civil and military use of such a technology in a secure and safe way.

2.2 Distributed Unit (DU)
Every DU is logically constituted by a "Gateway" (GW), one or more "Readers" (RDs) and one or more "Actuators" (ATs) communicating one another using wireless or wired technologies. The whole system can rely on several DUs, networked through secure communications with a Central Processing Station (CPS) and related controlled systems (Fig. 1). Each DU maintains synchronization of data with the central system and allows secure communications among the system components.
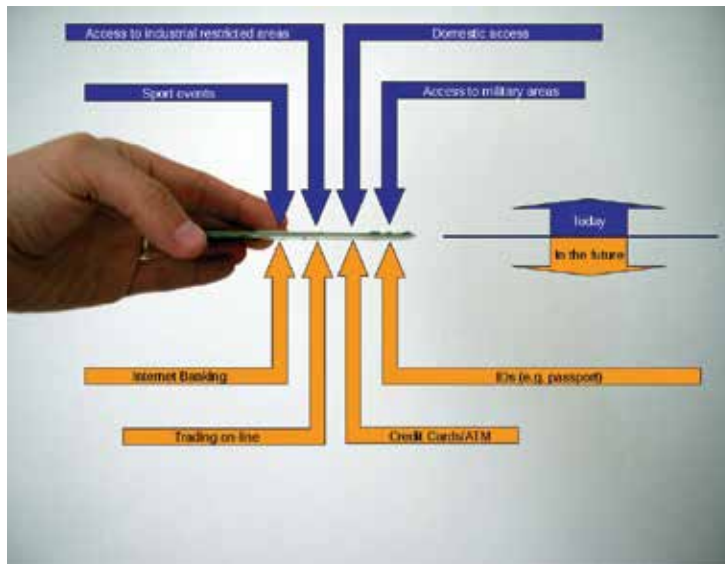
Fig. 3. WEST Aquila's Biometric Badge – features

2.3 Gateway (GW)

The Gateway is the central element of each DU. It communicates in a secure way with the RDs, with the ATs and back with the Central Processing Station (CPS). Its main function is to provide an interface between the BBs and the CPS, both upwards and downwards. In the upwards flow, the GW collects data from BBs through its associated RDs, interprets the information contained within each packet and sends it to the CPS. In the downwards flow, the GW receives instructions from the CPS and controls accordingly the ATs to grant or deny the access to the BB.

2.4 Reader (RD)

The Reader is the element for interfacing the DU with the BBs. Its role is to communicate wirelessly with the BBs, and it uses two mechanisms: the IEEE 802.15.4 communication standard and a proximity–based RFID technology. As a logical component of the system, it is not allowed to locally decode the data, but it simply forwards it towards the GW over a secure communication channel. The communication between the RD and the GW can be either wired or wireless. In the latter case, it can be direct, when they are in the communication range of each other, or indirect, i.e., multi-hop through intermediate RDs. When it is wireless and indirect, then the RDs constitutes a network of readers (NRD), organized into a ZigBee Cluster Tree (ZigBee, 2008), (Koubâa et. al, 2008).

2.5 Actuator (AT)

The Actuator is a device in direct contact with the Controlled System (Fig. 1) and it constitutes the interface with the GW so that the operations needed to provide the requested services to the BB owner are executed, when he/she has passed the authentication process, or the safety procedures when the authentication fails are applied. Similarly to the NRD, multiple ATs might form a multi–hop wireless network (NAT) to reach the GW.

2.6 Central Processing System (CPS)

The Central Processing System contains all data related to the whole system configuration. This implies that it stores and handles all the data related to the UDs and the BBs that have grants with the different UDs, as well as it handles the services that BB's owners can use when they request for them after a successful authentication. CPS communicates in a secure way with UDs and it is the interface with "System Administrator" (SA). The SA is in charge of two main tasks: (i) deliver the BBs to the people having rights of owning one of them and (ii) add and update in the CPS all data related to the system configuration, i.e., the association between the BB's ID and the services to which it can grant the access to its owner.

Table 1 summarizes the acronyms used in this book chapter.

Acronym	Meaning
AT	Actuator
BB	Biometric Badge
CPS	Central Processing Unit
DU	Distributed Unit
FP	Fingerprint scanner/sensor
GW	Gateway
NAT	Network of Actuators
NRD	Network of Readers
RD	Reader
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Indicator
SA	System Administrator
SoC	System-on-Chip
SW	The application module running on the GW for communicating with the companion chip on the BB
uC	Micro Controller

Table 1. Acronyms

2.7 System security framework

It is of paramount importance to clearly state that the security of the system is based on a novel framework of network security built on top of the framework provided by UPEK (UPEK, 2009) for its chips.

The "companion chip" is embedded in the fingerprint reader on board of each badge and is the element able to handle all the biometric aspects. When the authentication process is running, this chip is in communication with the GW over a wireless secure channel, where data travels ciphered by that chip. On the GW a UPEK–made application runs: it is the only component in the whole system able to decode these messages. This leads to an interesting aspect of the system in terms of security: the microcontroller on board of the badge is definitely not able to communicate with the companion chip. It can only switch it on or off or ask the GW to activate the procedures. In other words, the biometric part is usable if and only if the proposed system is able to establish a secure connection between the companion chip on the badge and the application running on the GW (Fig. 4). The security of this communication is granted by the fact that it is ciphered using a symmetric key based mechanism. These keys are unique for each badge and provided during the so called "key provisioning" performed
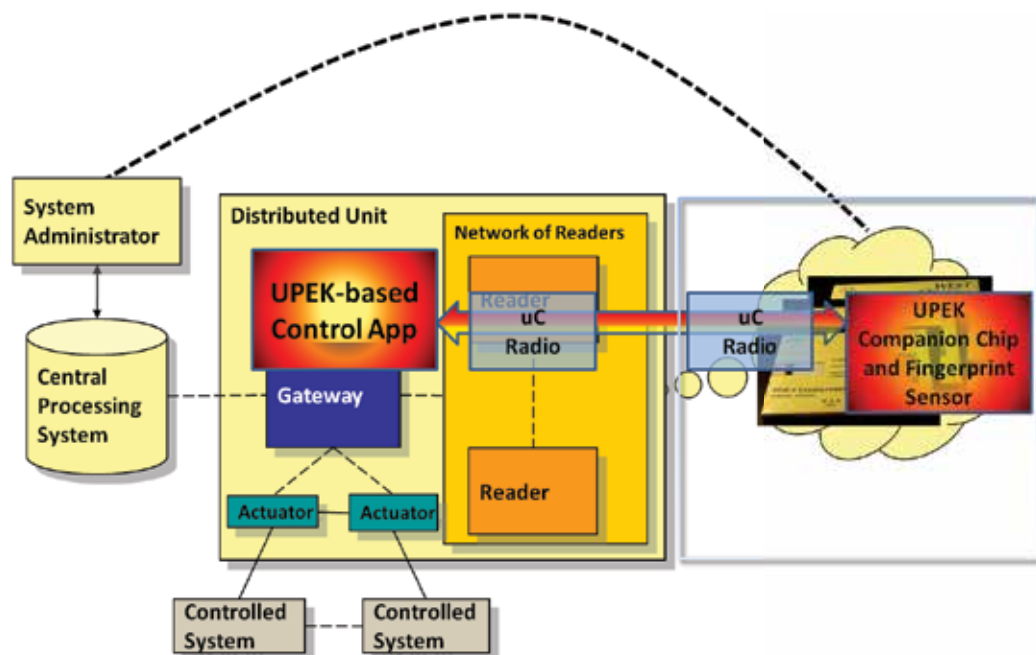
Fig. 4. Secure connection for biometric operations

when the badge is registered into the system for the first time. They are stored both in the gateway and in the "secure memory" of the companion chip. This memory is designed by UPEK to resist also to HW attacks and contains the fingerprints template. Furthermore, the communication is ciphered using a random component that modifies the content of the message so that its eventual sniffing doesn't provide useful information. The intermediate software and hardware elements between the companion chip and the application running on the GW during the authentication process simply act as packets' forwarders.

Since the BB is also equipped with RFID communication capabilities, there is the possibility to introduce another level of system security, granted by a novel mechanism we called "ring check". When the BB is really close to a RD, the RFID technology is activated. Then, RD writes in the BB's tag memory a ciphered code to allow the BB to recognize it as a qualified reader. The uC on the badge gets this code and checks its consistency to identify if it has not been altered. If it recognizes it, then on the BB side there is a confidence to be communicating with a verified reader. Then BB builds a packet with a "reader ok" status field set and that code, then sends this message to the GW, using the IEEE 802.15.4 transceiver. By this way, the BB can authenticate the system with which it is currently communicating and the system can check if the BB is not corrupted, by checking the integrity of the code returned back to it. In other words, the system checks if the RFID and IEEE 802.15.4 transceivers and related storage areas are both working.

2.8 System configurations

Starting from the logical architecture shown in Fig. 1, it is possible to derive several physical configurations that could be applied depending on the different needs. In particular, based on the possibilities offered by the allocation of the different components of the DU and the CPS onto a single HW or multiple communicating devices, several combinations of alternative

configurations can be identified. As an example, Fig. 5 shows 2 different scenarios. The red elements, FP and SW, represent the components dedicated to manage the biometrics operations. i.e., SW is the only component able to decode and manage information about the result of biometric verification coming from FP.

(a)

(b)
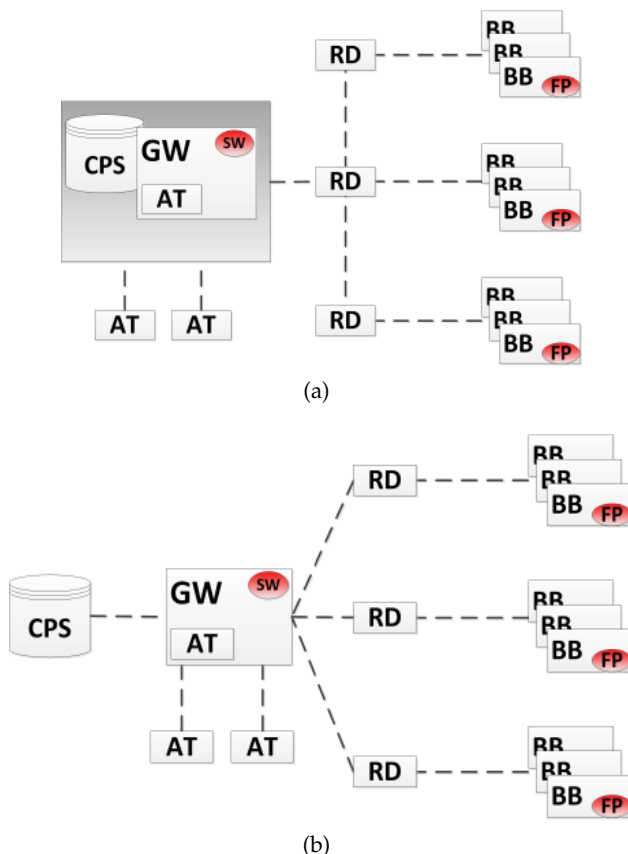
Fig. 5. Two different configurations

Fig. 5a refers to a scenario where CPS and GW are allocated on a single physical machine. A wireless interface is used to communicate with the RD units, organized into a network of readers. Fig. 5b shows a configuration where the system is fully distributed, i.e., CPS, GW and RDs are mapped onto different machines. The communication between these components might be based on IP protocols, like over Internet.

Although several other combinations are possible, these two scenarios are our reference to the case studies described in the following sections.

3. Case Study 1 – Physical access to a critical area

Let us assume that a SA has released a number of BBs to a number of authorized people by means of an enrollment procedure, then each one of them will have an enabled BB storing their own fingerprint, while the central system will be aware of the basic access rights of BB and related persons.
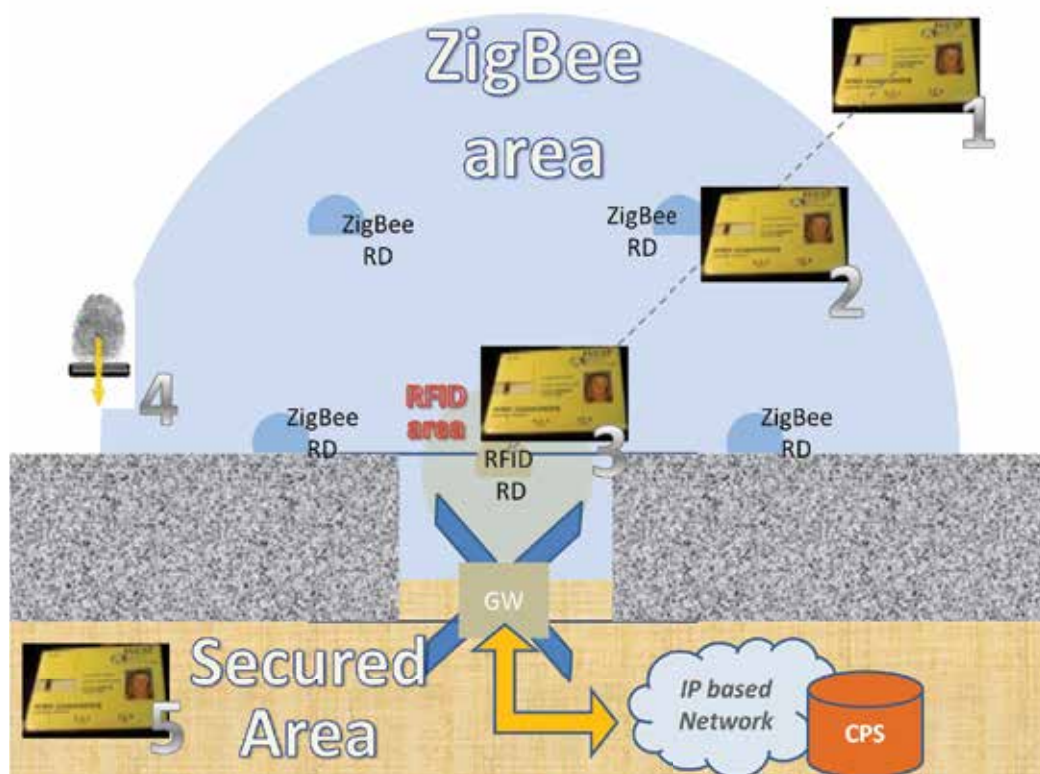
Fig. 6. Case Study 1 – Physical access to a critical area

The access to a critical area towards a controlled gate will be performed by means of the following steps (Fig. 6):

1. The BB is in stand–by mode, i.e., it is waiting for a beacon sent by one of the ZigBee RD forming a Cluster Tree topology (Hauer et al., 2011; Jurcik et al., 2010).

2. When the BB enters the ZigBee area, it is able to ear the beacons sent by the RDs and to communicate with the control unit on the gate in order to communicate its arrival. In this context, the badge implements the positioning solution as described in (Tennina, Di Renzo, Santucci & Graziosi, 2009) and summarized in the next subsections. In such a way, the DU is able to communicate with the central system in order to make in advance any control related to the badge identification (i.e., to check if it is allowed to access the gate it is approaching).

3. The BB is allowed to pass the gate, the DU will wait for the proximity of the BB.

4. When the BB is close to the gate then DU will request to the BB to start the personal identification, i.e., the BB will ask the owner to scan his/her fingerprint, it compares that scan with the stored one, and the result of such a verification is sent back to the DU, being the only unit able to decode that information.

5. If the identification is successful the DU will open the gate, otherwise proper actions defined by the system administrator will be taken.

Fig. 7 shows the setup used for the experimental testbed, where a Notebook plays the role of the GW: it has a RFID reader (on the right–hand side) and a ZigBee reader (on the left–hand side); furthermore it switches on a lamp to confirm the successful authentication, as well as feedbacks the user in a demo GUI (Fig. 8).
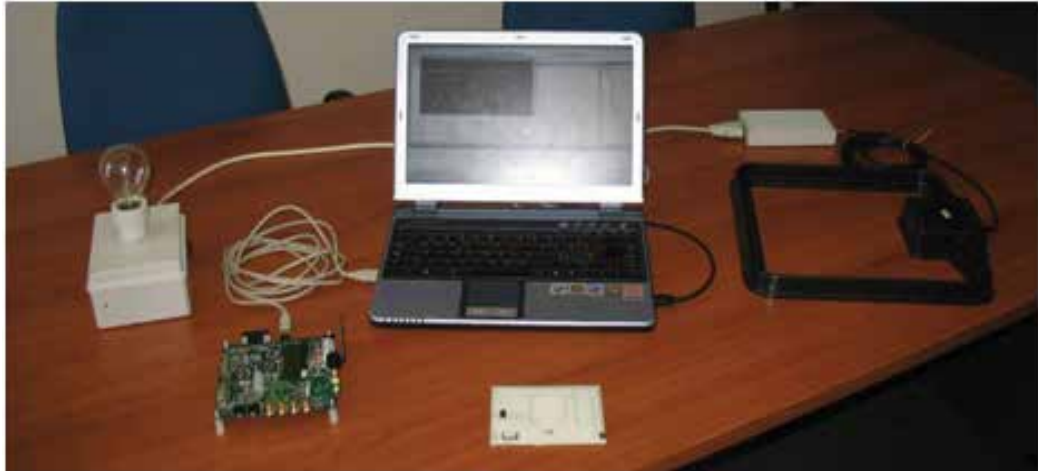
Fig. 7. Biometric Authentication – Setup

Fig. 8. Biometric Authentication – Successful verification

3.1 ESD: an improved optimization SW routine for positioning

The badge is equipped with a novel distributed localization algorithm, which is called ESD (Enhanced Steepest Descent) (Tennina, Di Renzo, Santucci & Graziosi, 2009). In particular, since this method represents an improved version of the well-known Steepest Descent (SD), the latter one is briefly summarized as well.

Let us consider N_A wireless nodes $\{A_i\}_{i=1}^{N_A}$ distributed in the region of interest, whose exact locations in the considered scenario are known, based on a predefined and common reference system of coordinates. These nodes are called *reference* or *anchors* nodes. Let us assume

also that N_U wireless nodes $\left\{ U_j \right\}_{j=1}^{N_U}$ with unknown location are present in the same area. These nodes are called *unknown* or *blind* nodes. Both these wireless nodes have a simple radio interface to communicate, which allows not only data exchange but also distance measurements. The main goal of a positioning system is to use the anchor nodes to estimate the position of the blind nodes in the specified coordinate system. In particular, position estimation algorithms require a minimum of either three or four reference nodes in a two– and tree–dimensional coordinate system, respectively (Perkins et al., 2006). In our context it is obviously assumed that Ai are the ZigBee Readers, while U_j are the Biometric Badges.

The following notation will be used to describe the algorithm: (i) bold symbols are used to denote vectors and matrices, (ii) $(\cdot)^T$ denotes transpose operation, (iii) $\nabla(\cdot)$ is the gradient operator, (iv) $\|\cdot\|$ is the Euclidean distance, (v) $\angle(\cdot, \cdot)$ is the phase angle between two vectors, (vi) $\hat{\mathbf{u}}_\mathbf{j} = [\mathbf{u}_{\mathbf{j},\mathbf{x}}, \mathbf{u}_{\mathbf{j},\mathbf{y}}, \mathbf{u}_{\mathbf{j},\mathbf{z}}]^\mathbf{T}$ with j=1,.., N_U denotes the estimated position of the unknown node U_j, (vii) $\mathbf{u}_j = [u_{j,x}, u_{j,y}, u_{j,z}]^T$ is the trial solution of the optimization algorithm for the unknown node U_j, (viii) $\mathbf{a}_i = [x_i, y_i, z_i]^T$ with i=1,.., N_A are the positions of the anchor/reference nodes A_i, and (ix) $d_{j,i}$ denotes the estimated (via ranging measurements) distance between reference node A_i and the unknown node U_j.

3.1.1 Multilateration methods

Both SD and ESD algorithms belong to the family of the multilateration methods. In particular, in such methods the position of an unknown node U_j is obtained by minimizing the error cost function $F(\cdot)$ defined as in Equation 1:

$$F\left(u_j \right) = \sum_{i=1}^{N_A} \left(d_{j,i} - \left\| \mathbf{u}_j - \mathbf{a}_i \right\| \right)^2 \tag{1}$$

The minimization of the error cost function can be realized using a variety of numerical optimization techniques, each one having its own advantages and disadvantages in terms of accuracy, robustness, speed, complexity, and storage requirements (Nocedal & Wright, 2006). Since optimization methods are iterative by nature, we will denote by the index k the k–th iteration of the algorithm, and with $F(\mathbf{u}_j(k))$ and $\mathbf{u}_j(k)$ the error cost function and the estimated position at the k–th iteration, respectively.

Steepest Descent (SD) The SD is an iterative line search method that allows to find the (local) minimum of the cost function in Equation 1 at step $k+1$ as follows (Nocedal & Wright, 2006, pp. 22, sec. 2.2):

$$\mathbf{u}_j(k+1) = \mathbf{u}_j(k) + \alpha_k \cdot \mathbf{p}(k) \tag{2}$$

where α_k is a step length factor, which can be chosen as described in (Nocedal & Wright, 2006, pp. 36, ch. 3), and $\mathbf{p}(k) = -\nabla(F(\mathbf{u}_j(k)))$ is the search direction of the algorithm. In particular, when the optimization problem is linear, some expressions exist to compute the optimal step length in order to improve the convergence speed of the algorithm. On the other hand, when the optimization problem is non-linear, as considered for positioning problems, a fixed and small step value is in general preferred in order to reduce the oscillatory effect when the algorithm approaches a solution. In such a case, we have $\alpha_k = 0.5\mu$ (Santucci et al., 2006), where μ is the learning speed.

Enhanced Steepest Descent (ESD) The SD method provides, in general, a good accuracy in estimating the final solution. However, it often requires a large number of iterations, which may result in a too slow convergence speed for mobile ad–hoc wireless networks. The

proposed ESD algorithm aims at improving the convergence speed of the SD algorithm, while trying to maintain its good accuracy for position estimation. The basic idea behind the ESD algorithm is to adjust the step length value α_k as a function of the current and previous search directions $\mathbf{p}(k)$ and $\mathbf{p}(k-1)$, respectively. In particular, α_k is adjusted as shown in Equation 3, where $\theta_k = \angle(\mathbf{p}(k), \mathbf{p}(k-1))$, $0 < \gamma < 1$ is a linear increment factor, $\delta > 1$ is a multiplicative decrement factor, and θ_{min} and θ_{max} are two threshold values which control the step length update.

$$\begin{cases} \alpha_k = \alpha_{k-1} + \gamma & if \ \theta_k < \theta_{\min} \\ \alpha_k = \alpha_{k-1}/\delta & if \ \theta_k > \theta_{\max} \\ \alpha_k = \alpha_{k-1} & otherwise \end{cases} \tag{3}$$

By using the four degrees of freedom $\gamma, \delta, \theta_{min}$ and θ_{max}, the convergence rate of the algorithm, and the oscillatory phenomenon when approaching the final solution can be simultaneously controlled in a simple way and without appreciably increasing the complexity of the algorithm when compared to the SD method. Basically, the main advantage of the ESD algorithm is the adaptive optimization of the step length factor α_k at run time, which allows to dynamically either accelerate or decelerate the convergence speed of the algorithm as a function of the actual value of the function to be optimized

3.1.2 Positioning system validation

Localization is performed in a fully RSSI–based distributed and decentralized fashion for the blind node. In other words, each blind node receives data from the fixed anchor/reference nodes (see Fig. 9a) and convert the RSSI measurements of each packet into an estimation of distance. It is well known that RSSI is as simple as really inaccurate, but this distance estimation accuracy has been improved on the blind node side, by allowing anchor nodes to perform an innovative on–line radio signal propagation characteristics estimation (Tennina et al., 2008).

In order to validate the proposed solutions, and have a sound understanding of the performance of the ESD algorithm in realistic scenarios, we have conducted a campaign of measurements during the opening ceremony day of the NCSlab on March 27, 2008 (Fig. 9b). The event was characterized by a half-day kick–off conference during which the past, present, and future activities of the laboratory were presented. The kick–off conference was attended by several people, and offered a good occasion to test the performance of the deployed network, and, in particular, to test the achievable performance in a realistic GPS-denied environment, where the propagation characteristics of the radio channel changed appreciably during the event due to the people's movement inside the room (i.e., dynamic indoor environment). The duration of the event was approximately three hours and forty minutes, thus providing enough statistical data to well support our findings and conclusions. This ceremony was characterized by four main phases, well describing the dynamic nature of the event and, as a consequence, the dynamic nature of the propagation environment to be analyzed. In what follows there is a brief description of each phase.

1. The first phase, which took place before the starting of the opening ceremony, is characterized by a progressive increase of the number of people inside the room, some of them very close and in motion around the blind node to be localized (i.e., the dot point in Fig. 9c).

(a) Battery powered anchor node   (b) Plan of the NCSlab
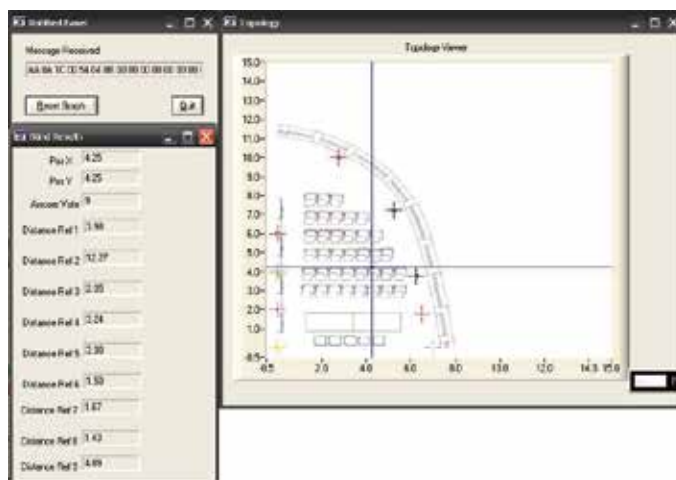
(c) Host Application

Fig. 9. ESD Positioning Validation – Experimental Setup

2. The second phase, which took place during the development of the ceremony, is characterized by several people (staying either seated or stand) inside the room, and some people coming in and going out the room.

3. The third phase, which took place at the end of the ceremony, is characterized by the vast majority of people staying stand and leaving the conference room.

4. The fourth phase corresponds to the scenario with no people in the room, thus giving a virtually static indoor scenario with almost fixed propagation characteristics.

Fig. 9c the host application interface with anchor (cross points) and blind (dot point) nodes deployed during the field tests and available to the user to analyze the behavior of localization and tracking operations.

The setup was characterized by the following main settings: (i) 9 anchor nodes, distributed on the room's perimeter, and 1 blind node have been considered, (ii) all the nodes were placed on the top of wood supports, (iii) the anchor nodes broadcasted their ID and position every 800 milliseconds as well as estimated the radio signal propagation characteristics as described in (Tennina et al., 2008), and (iv) every RSSI used by the blind node was obtained by averaging 10 RSSIs (Average RSSI) per anchor.
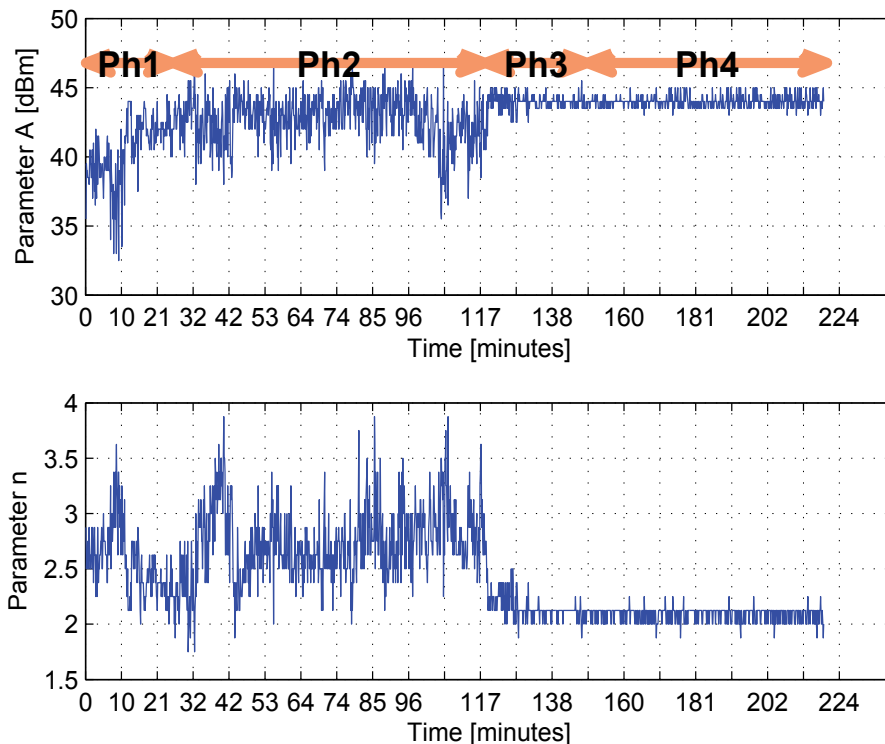
Fig. 10. Estimated propagation parameters during the NCSlab's opening ceremony

In Fig. 10, the estimated propagation parameters A and n (Tennina et al., 2008) are reported as a function of time. We can readily figure out that there is a significant fluctuation of these parameters during the progress of the conference, and, as expected, the variation gets large during Phase 1 and 3, and, in particular, during Phase 2, while they are almost constant during Phase 4, which represents a virtually static reference scenario. This figure qualitatively suggests that using an outdated estimate for the channel parameters may certainly yields less accurate estimates of the distances and thus of the position of the blind node.

Finally, Fig. 11 shows the positioning accuracy of the proposed ESD algorithm running on the blind node when it can resort on the estimations of the propagation parameters updated on–line by the anchor nodes. As we can see, even if the dynamic of the environment might change dramatically the propagation conditions, the positioning accuracy is good enough, i.e., with an average error generally less than 2 meters, and stable, i.e., no major fluctuations.

Similar accuracies have been obtained in recent experimental trials (Tennina, Pomante, Graziosi, Di Renzo, Alesii & Santucci, 2009).
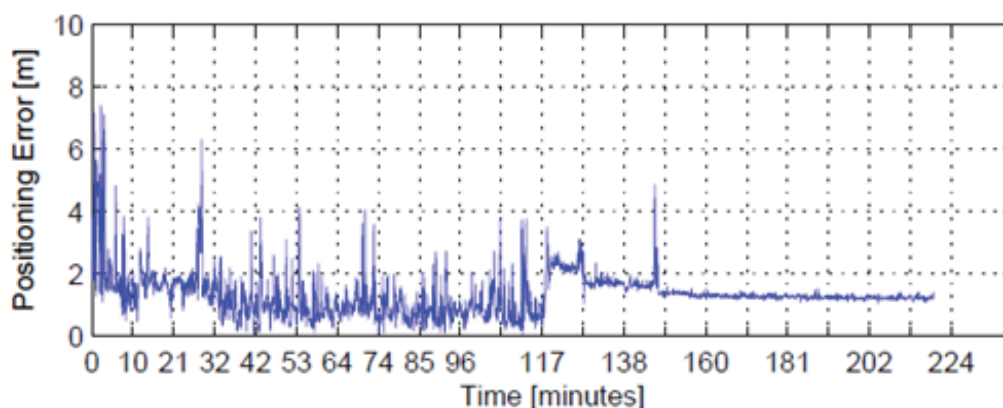
Fig. 11. ESD positioning accuracy with on-line dynamic propagation parameters estimation

4. Case Study 2 – Logical access to a critical area

The security framework described in Section 2.7 allows the exploitation of the badge virtually everywhere. A typical scenario is the home banking, where users access remotely to their bank account. Nowadays they usually receive a one–time password generator, which is used when the bank's web page ask it. The idea is to grant access to such services by relying on the higher security levels guaranteed by the usage of biometric–based authentication. By simply using a PC with an IEEE 802.15.4 radio interface (today it is available as an external USB dongle, but in the near future it will be probably integrated into the PC's motherboards as for IEEE 802.11 radio interfaces) and a classical Internet connection, the badge is able to establish a secure connection between its on-board companion chip and the management SW by means of the PC and the Internet that are used to reach the gateway. Basically, the PC acts as a RD.

Fig. 12 shows an example of such a configuration where the access to a web site is authorized only when a verification operations is correctly performed by means of the biometric badge. In such a scenario, the web server acts as the GW, so managing the companion chip of the badge by means of a secure connection that exploits the Internet and the connection from the PC to the badge. The web server asks the badge for the authentication of its owner and based on the result (that, in this case, only the web server is able to decode) it grants or denies the access to the web site.

In order to clarify the whole procedure, let assume that the system is used to manage the access to an online bank account. As for the case study 1, the SA has released a number of BB to a number of authorized people by means of an enrollment procedure. Some of these BB are enabled to identify the user in order to allow him/her to access to the bank account. Finally, the user has a personal computer with an IEEE 802.15.4 transceiver in order to communicate with the BB (like the one shown in Fig. 13).

The access to the bank account will be performed through a web interface and by means of the following steps:

1. The user tries to access the bank account on the server and he/she is accordingly redirected to an identification page where some credentials (i.e., username and password) are requested (Fig. 14).

Fig. 12. Case Study 2 – Logical Access to a Critical Area

(a) USB IEEE 802.15.4 dongle and the badge   (b) Personal web site
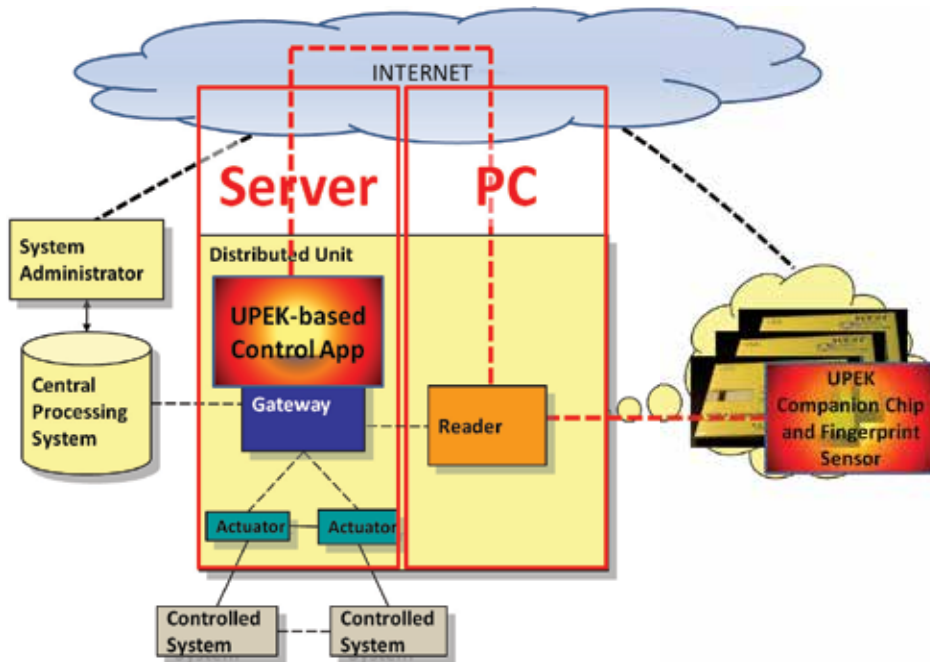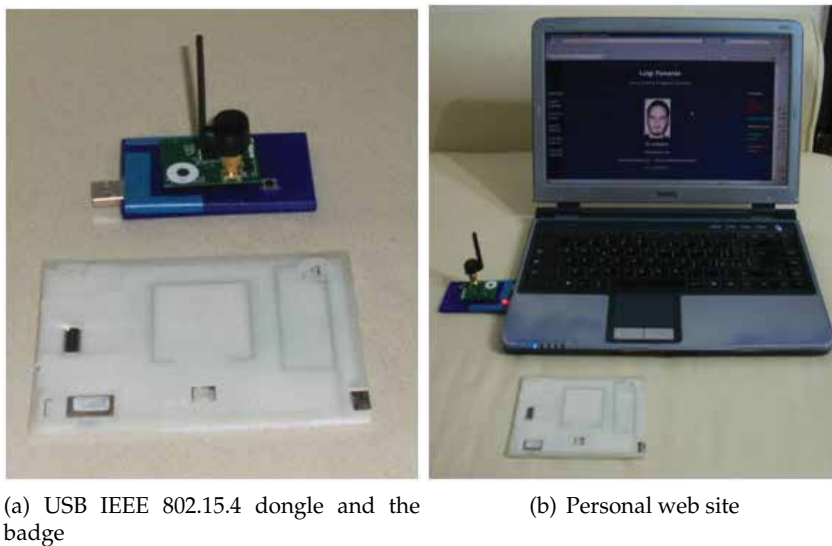
Fig. 13. Case Study 2 – Logical Access to a Critical Area – Prototypes

2. Then the system asks the user to activates the badge. The BB is then able to establish a secure connection with the web server and the SW running therein starts communicating with the companion chip.

Fig. 14. Case Study 2 – Logical Access to a Critical Area – Login

Fig. 15. Case Study 2 – Logical Access to a Critical Area – Biometric Authentication

3. The web server asks the BB to start the personal identification, i.e., the user will scan his/her fingerprint, while the BB compares it with the stored one, and the result of such a verification is sent back to the web server that is the only unit able to decode such an information (Fig. 15).

4. If the identification is successful, the web server grants the access to the bank account web site (Fig. 16), otherwise the proper actions defined by the system administrator are taken.

It is worth noting that the user's PC and the server can be in different location everywhere in the world, they only need an Internet connection to communicate.

5. Conclusions

The system presented in this chapter is an innovative solution to the problem of automatic and secure advanced services provision, and it is able to guarantee users' identity in a easy and safe way. Although this system is flexible enough to be used in several domains, it meets

Fig. 16. Case Study 2 – Logical Access to a Critical Area – Success

the more rigid laws about the users' privacy without compromising its ease of use, that is the main factor to make it accepted and widely used. The key component is the innovative biometric badge which implements the concept of system–on–badge: a system that is able to automatically perform and check fingerprint scans, in order to verify if the badge owner is actually the person to whom the badge was delivered. Furthermore, it is able to communicate only the results of this verification to the remaining part of the system without the need to share sensible data, i.e., users' biometric information. The use of mature, reliable and low-cost technologies, accurately integrated, is the basis to truly achieve high level of pervasiveness of the biometric techniques in order to support the most important human activities.

6. References

Adeoye, O. S. (2010). A survey of emerging biometric technologies, *International Journal of Computer Applications* 9(10): 1–5. Published By Foundation of Computer Science.

BTAM (2010). Biometric technology application manual.
URL: *www.nationalbiometric.org*

Elliott, S., Massie, S. & Sutton, M. (2007). The perception of biometric technology: A survey, *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on*, pp. 259 –264.

Hauer, J.-H., Daidone, R., Severino, R., Busch, J., Tiloca, M. & Tennina, S. (2011). An open-source ieee 802.15.4 mac implementation for tinyos 2.1. Poster Session at 8th European Conference on Wireless Sensor Networks.
URL: *http://www.nes.uni-due.de/ewsn2011*

IEEE (2006). Standard for information technology part 15.4: Wireless medium access control (mac) and physical layer (phy) specifications for low rate wireless personal area networks (lr-wpans), LAN/MAN Standards Committee of the IEEE Computer Society Std.
URL: *http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf*

Jurcik, P., Severino, R., Koubaa, A., Alves, M. & Tovar, E. (2010). Dimensioning and worst-case analysis of cluster-tree sensor networks, *ACM Transactions on Sensor Networks* 7(2).

Li, P. & Zhang, R. (2010). The evolution of biometrics, *Anti-Counterfeiting Security and Identification in Communication (ASID), 2010 International Conference on*, pp. 253 –256.

Montalbano (2009). Mtsens iso 15693 compatible 13.56 mhz rfid tag.
    URL: *http://www.montalbanotechnology.com*

Nocedal, J. & Wright, S. (2006). *Numerical Optimization*, second edn, Springer.

Perkins, D., Tumati, R., Wu, H. & Ajbar, I. (2006). Localization in wireless ad hoc networks, 16: 507–542.

Santucci, F., Graziosi, F. & Tennina, S. (2006). Service design and simulation in ad-hoc wireless sensor networks, *International Journal on Mobile Networks Design and Innovation* 1: 208–214.

Sonkamble, S., Thool, R. & Sonkamble, B. (2010). Survey of biometric recognition systems and their applications, *Journal of Theoretical and Applied Information Technology* 11(1).

Tennina, S., Di Renzo, M., Graziosi, F. & Santucci, F. (2008). Locating zigbee nodes using the ti's cc2431 location engine: a testbed platform and new solutions for positioning estimation of wsns in dynamic indoor environments, *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, International Conference on Mobile Computing and Networking, San Francisco, California, USA, pp. 37–42. SESSION: Radio/RSSI based methods.

Tennina, S., Di Renzo, M., Santucci, F. & Graziosi, F. (2009). Esd: A novel optimization algorithm for positioning estimation of wsns in gps–denied environments – from simulation to experimentation, *International Journal of Sensor Networks* 6(3/4): 131–156.

Tennina, S., Pomante, L., Graziosi, F., Di Renzo, M., Alesii, R. & Santucci, F. (2009). Localization, tracking, and automatic personal identification in gps-denied environments a solution based on a wireless biometric badge, *Tridentcom, 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops*, pp. 1–3.

TI (2009). A true system-on-chip solution for 2.4 ghz ieee 802.15.4 / zigbee(tm). Datasheet, Rev. F.
    URL: *http://focus.ti.com/docs/prod/folders/print/cc2430.html*

UPEK (2009). Chipset tcs3-tcd42, touchstripő fingerprint authentication solution. TouchStrip Fingerprint Sensor (TCS3) and the Digital ID Hardware Engine.
    URL: *http://www.upek.com/solutions/portable/chipset.asp*

WESTAquila (2010). Wireless embedded systems technologies – l'aquila.
    URL: *http://www.westaquila.com*

Part 2

Application of Cancelable Biometrics

An Overview on Privacy Preserving Biometrics

Rima Belguechi, Vincent Alimi, Estelle Cherrier, Patrick Lacharme
and Christophe Rosenberger
Université de Caen Basse-Normandie, UMR 6072 GREYC, F-14032 Caen
ENSICAEN, UMR 6072 GREYC, F-14050 Caen
CNRS, UMR 6072 GREYC, F-14032 Caen
France

1. Introduction

The Internet has consolidated itself as a very powerful platform that has changed the communication and business way. Nowadays, the number of users navigating through Internet is about 1,552 millions according to Internet World Stats. This large audience demands online commerce, e-government, knowledge sharing, social networks, online gaming ... which grew exponentially over the past few years. The security of these transactions is very important considering the number of information that could be intercepted by an attacker. Within this context, authentication is one of the most important challenges in computer security. Indeed, the authentication step is often considered as the weakest link in the security of electronic transactions. In general, the protection of the message content is achieved by using cryptographic protocols that are well known and established. The well-known ID/password is far the most used authentication method, it is widely spread despite its obvious lack of security. This is mainly due to its implementation ease and to its ergonomic feature: the users are used to this system, which enhances its acceptance and deployment. Many more sophisticated solutions exist in the state of the art to secure logical access control (one time passwords tokens, certificates ...) but none of them are used by a large community of users for a lack of simplicity usage (O'Gorman, 2003).

Among the different authentication methods of an individual, biometrics is often presented as a promising solution. Few people know that biometrics has been used for ages for identification or signature purposes. Fingerprints were already used as a signature for commercial exchanges in Babylon (-3000 before JC). Alphonse Bertillon proposed in 1879 to use anthropometric information for police investigation. Nowadays, all police forces in the world use this kind of information to solve crimes. The first prototypes of terminals providing an automatic processing of the voice and digital fingerprints have been defined in the middle of the years 1970. Today, a large number of biometric systems are used for logical and physical access control applications. This technology possesses many favorable properties. First, there is a strong link between the user and its authenticator. As for example, it is not possible to loose its fingerprint as it could be the case for a token. Second, this solution is very usable: indeed, it is very convenient for a user to authenticate himself/herself by putting his/her finger on a sensor or making a capture of the face. Last, biometrics is an interesting candidate to be a unique user's authenticator. A study done by NTA group in 2002 (Monitor, 2002) on 500 users showed that there was approximately 21 passwords per user, 81% of them use

common passwords and 30% of them write their passwords in a file. The uniqueness inherent to any biometric information is a helpful property to solve the aforementioned problems.

Of course, some drawbacks are also inherent to this technology (Bolle et al., 2002). Whereas the uniqueness can be considered as an advantage, it could also allow an attacker to trace operations done by an user through the logging of authentication sessions. Then the biometric verification step ensures with a high probability that the user is the genuine one but there is still some possibilities the user is an attacker. This is a far different approach than checking if a password is correct or not. One of the biggest drawbacks of biometrics is the impossibility to revoke the biometric data of a user if they are compromised (Galbally et al., 2008). This point is related to the users acceptance that need to be sure that their privacy will be respected: how can people be sure that their personal data collected during the enrollment step will not be stolen or diverted and used for other purposes ? This pregnant issue limits the operational use of biometrics for many applications. As for example, in France, it is forbidden to establish a centralized database of biometric data because it is considered too dangerous from a privacy point of view.

The objective of this chapter is to realize an overview on the existing methods to enhance the privacy of biometrics. Section 2 is dedicated to a study of the threats involving privacy in biometric systems, and the ensuing requirements. We present in section 3 some biometrics based secure storage and template protection schemes. Section 4 deals with the remaining challenges in this domain. We conclude this chapter in section 5.

2. Privacy: threats and properties

We present in this section privacy issues concerning authentication schemes and biometric ones.

2.1 Privacy and personal data

The word *privacy* means different things to different people; hence the reference (Solove, 2009) has indicated the complexity of defining privacy. Instead of proposing an overview of different conceptual definitions, we formalize a core definition in which privacy means not only keeping a secret but also covering information and activities involving each person. Referring to some jurisdictions like the European Data Protection Directive (Dir95/46/EU), we give the following definitions:

Definition 1. *Personal data is any information relating to an identified or identifiable natural person (data subject).*

Definition 2. *An identifiable person is someone who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.*

It is clear that biometric systems (detailed in the next section) are designed to identify individuals. So, to examine the implication of privacy using biometric data, it is first necessary to define what is a biometric system and to study to what extent biometric data concerns/threats privacy. Then, we will be able to examine whether this data is personal and to measure the amount of sensitive information it reveals.

2.2 On biometric systems
We begin with a theoretical definition.

Definition 3. *A biometric system can be viewed as a signal detection system with a pattern recognition architecture that senses a raw biometric signal, processes this signal to extract a salient set of features called biometric identifier or template and compares these features against the ones stored in the database (Jain et al., 2006).*

More precisely, all biometric systems involve two steps.

- *Enrollment step*
  Biometric data (fingerprint, face, iris. . .) are *captured*, *transformed* into a template linked to the individual and *stored* as a reference.

- *Verification step*
  A new template is *issued* from a new capture, and *compared* to the stored reference template.

Given any biometric modality (fingerprint, face, iris. . .), its representation is not unique. As an illustration, consider fingerprint as a subject of study. Then, there are four different wide-spread fingerprint representations:

- Image-based representation
  Two fingerprints are superimposed and the correlation between corresponding pixels is computed for different alignments.

- Minutiae-based representation
  It is the most popular and widely used technique. A fingerprint appears as a surface alternating parallel ridges and valleys in most regions. Minutiae represent local discontinuities and mark positions where the ridge ends or splits. Minutiae-based matching consists in finding the alignment that results in the maximum number of minutiae pairings.

- Ridge feature-based approach
  Other features of the fingerprint ridge pattern (e.g., local orientation and frequency, ridge shape, texture information) may be extracted more reliably than minutiae in low-quality images.

- Pores-based representation
  With the current development of high resolution sensors, fine fingerprint features such as sweat pores can be considered.

To study privacy issues involved in biometric systems, we explore now how a biometric identifier is personal and sensitive. Two types of errors are present at the verification step:

- *false match*: the verification process outcome is that biometric measurements from two different persons are from the same person

- *false non-match*: the verification process outcome is that two biometric measurements from the same person are from two different persons

These two types of errors are quantified by the *false acceptance rate* and the *false rejection rate*, respectively. Figure 1 presents the error rates of four popular biometric modalities.

2.3 Biometrics and privacy

Biometric data, in its raw or template form (like minutiae template), is in most cases personal data. The reference (Pakanti et al., 2002) estimated a probability that two fingerprints will falsely match as $5.5 * 10^{-59}$. This probability is very low and shows that minutiae information can uniquely identify a person. In practice, as deduced from figure 1, deployment of biometric systems does not imply that the recognition is a fully solved problem. The accuracy changes

Biometric trait	Test	False Rejection Rate	False Acceptance Rate
Fingerprint	FVC 2006	2.2%	2.2%
	FpVTE 2003	0.1%	1%
Face	FRVT 2006	0.8-1.6%	0.1%
Voice	NIST 2004	5-10%	2-5%
Iris	ICE 2006	1.1-1.4%	0.1%

Fig. 1. Illustrations of error rates for different biometric modalities (Teoh et al., 2004b)

depending on different factors (the used modality, the population characteristics, the test conditions and the employed sensor to mention a few) but is never perfect. However, the obtained performances are considered sufficient to conclude that biometric data identifiers can recognize persons. Thus, they are *personal* or *very personal*, in the sense that they consist of information collected from an observation of the individual physical itself.

In return, biometric data are generally considered as sensitive data involving ethical and privacy contests.

2.3.1 Privacy threats in biometric systems

We summarize below potential privacy pitfalls arising when using a biometric identifier (fingerprint modality being again focused on):

1. Biometric information (especially raw images) can expose sensitive information such as information about one's health, racial or ethnic origin and this information can then provide a basis for unjustified discrimination of the individual data subjects (Mordini & Massari, 2008).

2. As revealed in (Schneier, 1999), biometric data are unique identifiers but are not secret: fingerprint is leaved on everything we touch, faces can be easily acquired and voice can be simply recorded. Hence, the potential collection and use of biometric data without knowledge of its owner, without his/her consent or personal control make this information very sensitive.

3. Many proponents of biometric systems claim that it is sufficient to store a compact representation of the biometric (template) rather than the raw data to ensure privacy of individuals. They consider that template is not sensitive information because it does not allow the reconstruction of the initial signal. Recently, several research works showed that this reconstruction is possible. For example, fingerprint can, in fact, be reconstructed from a minutiae template (Cappelli et al., 2007), (Feng & Jain, 2009).

4. The linkage problem which means the possibility to cross matched data across different services or applications by comparing biometric references is another privacy concern. The uniqueness of biometric characteristics allows an intruder to link users between different databases, enabling violations as tracking and profiling individuals.

5. A function creep is another privacy risk. Here, the acquired biometric identifiers are later used for purposes different from the intended ones. For example, an application initially intended to prevent misuse of municipal services may gradually be extended to rights to buy property, to travel, or the right to vote without the consent of individuals.

6. The inherent irrevocability of biometric features in case of data misuse like database compromise or identity theft makes biometrics very sensitive.

With the present risks on privacy violation, carefully handling biometric data becomes more important. Considering the implication of personal sensitive data, the use of biometrics falls within the purview of legislation and laws. In reality, regulations and legislation have codified what Judge Samuel Warren and Louis Brandeis summarized in 1890 as the right of the individual to be alone (Warren & Brandeis, 1890) (this reference is considered as the birthplace of privacy rights), and expanded the notion of data protection beyond the fundamental right to privacy. In the sequel, we are interested in the main attack vectors concerning biometric systems.

2.3.2 Biometric attack vectors
Possible attacks points (or attack vectors) in biometric systems have been discussed from different viewpoints. First, we can mention the scheme of figure 2, provided by the international standard ISO /IEC JTC1 SC37 SD11, which identifies where possible attacks can be conducted.
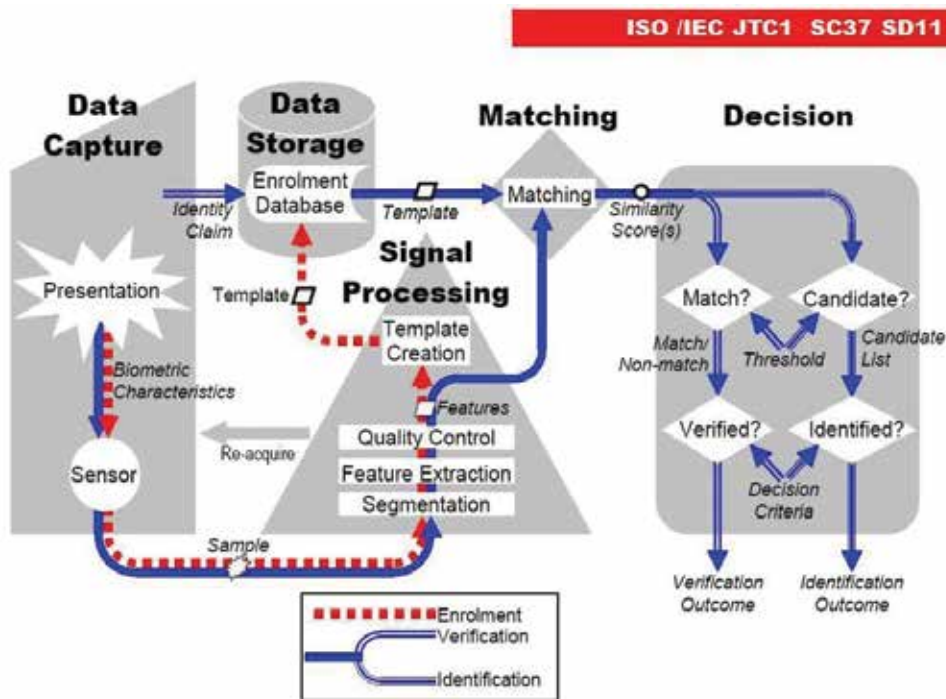
Fig. 2. ISO description of biometric systems

Besides, some of the early works by Ratha, Connell and Bolle (Ratha et al., 2001), (Bolle et al., 2002) identified weak links in each subsystem of a generic authentication system. Eight places where attacks may occur have been identified, as one can see in figure 3.
We do not detail in the present chapter all the types of attacks identified by Ratha. We only focus on attacks concerning privacy. This corresponds to the points 6 and 7 in figure 3. These points are related to attacks violating template protection. Generally, attacks directly threatening biometrics template can be of different types. For instance, an attacker could:
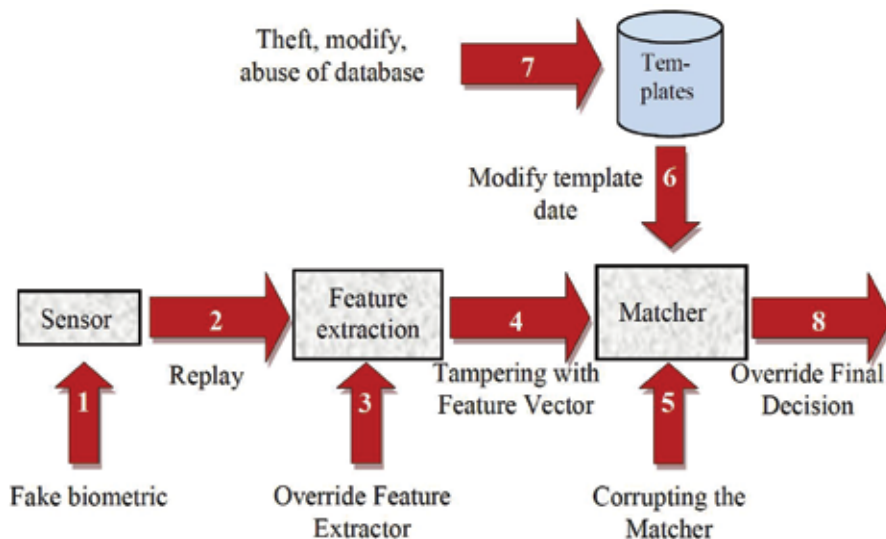
Fig. 3. Ratha's model attack framework

- attempt to capture a reference template
- substitute a template to create a false reference
- tamper the recorded template
- compromise the database by stealing all its records

Such attacks can be very damaging, owing to unavoidable exposure of sensitive personal information and identity theft. Therefore basic requirements that any privacy preserving biometric system should fulfill will be stated in the following.

2.3.3 Requirements of privacy protection
In view of our discussion about biometric systems vulnerabilities and possible threats, a few desirable properties are required, regarding the system safety. A critical issue in the biometric area is the development of a technology to handle both privacy concerns and security goals, see (Jain et al., 2008) for example. We detail now the key considerations for privacy protection.

- All deployments of biometric technology should be implemented with respect to local jurisdictional privacy laws and regulations.

- Today, some legal frameworks introduce the idea of *Privacy by Design*. This new paradigm requires that privacy and data protection should be integrated into the design of information and communication technologies. The application of such principle would emphasize the need to implement Privacy Enhancing Technologies (PET) that we will see after.

As explained in the reference (Adler, 2007), privacy threat is closely related to security weakness. Therefore a particular attention has been paid to privacy enhancing techniques. The aim is to combine privacy and security without any tradeoff between these two basic requirements. Among the techniques related to privacy enhancing, we can mention recent trends:

- Biometric encryption
  Based on cryptographic mechanisms, the ANSI (American National Standards Institute) proposes X9.84 standard as a means to manage biometric information. ANSI X9.84 rules were designed to maintain the integrity and confidentiality of biometric information using encryption algorithms. Even if cryptography has proven its ability to secure data transmission and storage, it becomes inadequate when applied to biometric data. Indeed, owing to the variability over multiple acquisitions of the same biometric trait, one cannot store a biometric template in an encrypted form and then perform matching in the encrypted domain: cryptography is not compatible with intra-user variability. Therefore the comparison is always done in the biometric feature domain which can make it easier for an attacker to obtain the raw biometric data. Since the keys and *a fortiori* the biometric data are controlled by a custodian, most privacy issues related to large databases remain open.

- Template protection schemes
  To solve this problem, recently some algorithms known as template protection schemes have been proposed. These techniques, detailed in section 3.1, are the most promising for template storage protection.

- Anonymous database
  The idea in anonymous data is to verify the membership status of a user without knowing his/her true identity. A key question in anonymous database is the need for secure collaboration between two parties: the biometric server and the user. The techniques presented in sections 3.2 and 3.3 fulfill this requirement.

In this chapter, privacy protection is be considered from two points of view: trusted systems and template protection. Recently, some template protection schemes have been proposed. Ideally, these algorithms aim at providing the following properties as established in the reference (Maltoni et al., 2009), and fulfill the privacy preserving issues 1 to 6 raised at page 4.

- *Non-reversibility*
  It should be computationally infeasible to obtain the unprotected template from the protected one. One of the consequences of this requirement is that the matching needs to be performed in the transformed space of the protected template, which may be very difficult to achieve with high accuracy. This property concerns points 1 to 3.

- *Accuracy*
  Accuracy recognition should be preserved (or degraded smoothly) when protected templates are involved. Indeed, if the accuracy of recognition degrades substantially, it will constitute the weakest link in the security equation. For example, instead of reversing the enrolment template, the hacker may try to cause a false acceptance attack. Thus, it is important that the protection technique does not substantially deteriorate the matching accuracy. This property is a general one.

- *Cancelability and Diversity*
  It should be possible to produce a very large number of protected templates (to be used in different applications) from the same unprotected template. This idea of cancelable biometrics was established for the first time in the pioneering references (Ratha et al., 2001) and (Bolle et al., 2002). To protect privacy, diversity means the impossibility to match protected templates from different applications (this corresponds to the notion of non linkage). Points 4 and 5 are concerned with diversity while point 6 with cancelability.

Compared to (Maltoni et al., 2009), we wish to add an extra property which will be used in the sequel:

- *Randomness*
  The knowledge of multiple revoked templates does not help to predict a following accepted one. This property deals with points 3 and 4.

Since some basic requirements in terms of privacy protection have been stated, biometric techniques fulfilling these requirements are detailed in the next section.

3. Privacy enhancing biometric techniques

In this section, we focus on some recent solutions brought by researchers in biometrics to handle template protection issues. These solutions generally aim at guaranteeing privacy and revocability. First, we detail promising solutions concerned with the storage of biometric templates in secure elements. Then, we emphasize on two approaches dealing with privacy enhancing biometric techniques: the first one is called biometric cryptosystems and the second is known as BioHashing.

3.1 Storage in secure elements

A key question is in relation with the place of storage of data and its security: Is it conserved in a local way (e.g. token)? Or in a central database with different risks of administration, access and misuse of this database? The problem becomes more relevant when dealing with large scale biometric projects such as the biometric passport or the national electronic identity card. Different organisations like the CNIL in France warn against the creation of such databases especially with regard to modality with traces as is the case for fingerprint (it is possible to refer to the central database to find the identity of those who left their traces). The INES debate launched in 2005 is a good illustration of the awareness of such subject (Domnesque, 2004). The use of biometrics may pose significant risks that encourage link ability and tracing of individuals and hence violating the individual liberties. So, the security of the stored biometric data remains challenging and this crucial task is pointed out by experts and legislation authorities. In this section, we study the storage of data in a secure local component: the *Secure Element*.
In (Madlmayr et al., 2007) one can find the following definition:

Definition 4. *The* Secure Element *is a dynamic environment, where applications are downloaded, personalized, managed and removed independently with varying life cycles.*

It is mainly used in smart cards or mobile devices to host sensitive data and applications such as biometrics templates or payment applications. It allows a high level of security and trust (e.g. the required security level for payment applications is set to Common Criteria EAL5+).
The Secure Element can be seen as a set of logical components: a microprocessor, some memory, an operating system and some applications. The secure element operating systems are known as MULTOS, Windows for Smart Cards, ZeitControl, SmartCard .NET and the most widespread: Java Card. Java Card is an adaptation of the well-known Java technology to smart card constraints. Java Card is an open language, which explains its great success. Based on a virtual machine environment, it is very portable (following the famous *Write Once, Run Everywhere*) and allows several applications to be installed and run on the same card.
But some drawbacks are also inherent to this technology: indeed the cohabitation of applications raises some questions. How and when to load the applications? Shall

applications loading be secured ? How to isolate applications from each others ? How long is the life cycle of a single application on the card ? How to determine the privileges of an application ?... Answers to these issues are provided by the GlobalPlatform technology.

3.1.1 GlobalPlatform overview

The GlobalPlatform technology is the fruit of the GlobalPlatform consortium's work. The GlobalPlatform consortium (formerly named Visa Open Platform) is an organization established in 1999 by leading companies from the payment and communications industries, the government sector and the vendor community. The GlobalPlatform specifications cover the entire smart card infrastructure: smart cards, devices and systems. Consistently written, this set of specifications allows developing multi-applications and multi-actors smart cards systems. It specifies the technical models that meet the business models requirements.

The GlobalPlatform card specification (GlobalPlatform, 2006) defines the behavior of a GlobalPlatform Card. As it is shown in figure 4, the GlobalPlatform card architecture comprises security domains and applications. A *Security Domain* acts as the on-card representatives of off-card entities. It allows its owner to control an application in a secure way without sharing any keys nor compromising its security architecture. There are three main types of Security Domain, reflecting the three types of off-card authorities recognized by a card: *Issuer Security Domain*, *Application Provider Security Domain* and *Controlling Authority Security Domain*.

3.1.2 Application to biometric template secure storage

The secure storage of biometric templates on a GlobalPlatform card is realized by an application. This dedicated application is installed, instantiated, selected and pushed a reference biometric template. In the verification case, the minutia are pushed to the application which processes a *match-on-card* verification and returns the result to the outside world.

In order to host this application, an Application Provider Security Domain must previously be created on the card. This security domain is personalized with a set of cryptographic keys and can then provide the application with security services such as cryptographic routines support and secure communications.

The application is involved in both the user enrolment and verification. For those two phases, the application queries the security services of its associated security domain.

During the user enrolment process, a secure communication channel is established between an off-card entity (in the present case a personalization bureau) and the application intended to store the reference biometric template. To this purpose, the security domain handles the handshake between the off-card entity and the application and unwraps the ciphered biometric template prior to forwarding it to the application. Three security levels are available for the secure communication: authentication, integrity and confidentiality.

During the verification process, a secure communication channel is established following the previous scheme. Contrary to the enrolment step, in this phase, the minutiae (and not the reference template) are ciphered and sent to the application. Hence the verification process takes place on card in a secure manner.

We have seen in this section how the Secure Element, a tamper proof component, ensures the secure storage of biometric templates. Indeed, thanks to the GlobalPlatform architecture, the access to the application devoted to the storage of the template and performing the match-on-card verification is secured successively by authentication of the off-card entity, check of data integrity and data ciphering for confidentiality purpose.
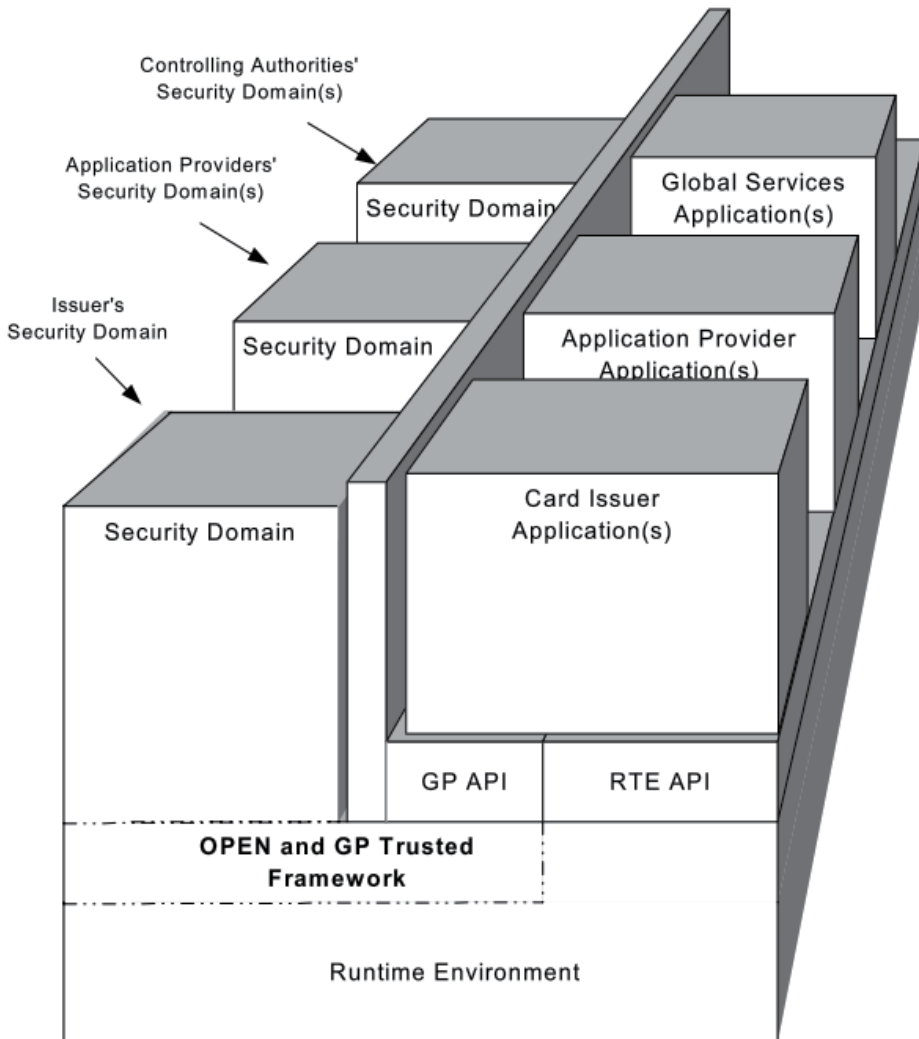
Fig. 4. GlobalPlatform Card Architecture (source: GlobalPlatform)

The next sections are concerned with two algorithmic-based solutions dealing with biometric template protection.

3.2 Cryptographic based solutions

Secure sketches have been introduced by Dodis *et al.* and formalized for a metric space H and the associated distance d, in relation to biometric authentication in (Dodis et al., 2004), (Dodis et al., 2006). A secure sketch considers the problem of error tolerance, existing in biometric authentication context: a template $b \in H$ must be recovered from any sufficiently close template $b' \in H$ and a additional data P. At the same time, the additional data P must not reveal too much information on the original template b. It uses the notion of minimal

entropy $H_\infty(X)$ of a random variable X, defined by the maximal number k such that for all $x \in X$, the probability $P(X = x) \le 2^{-k}$.

Definition 5. *A (H, m, m', t)-secure sketch is a pair of functions SS and Rec such as:*

- *The randomized function SS takes as input a value $b \in H$ and outputs a public sketch in $\{0, 1\}^*$, such that for all random variable B in H, with minimal entropy $H_\infty(B) \ge m$, the conditional minimal entropy $H_\infty(B|SS(B)) \ge m'$.*

- *The deterministic function Rec takes as input a sketch $P = SS(b)$ and a value $b' \in H$ and outputs a value $b'' \in H$ such that $b'' = b$ if the distance $d(b, b') \le t$.*

The first secure sketch was proposed by Juels et Wattenberg in (Juels & Wattenberg, 1999). This scheme is called *fuzzy commitment* and uses error-correcting codes. The *fuzzy vault* scheme of Juels and Sudan (Juels & Sudan, 2001) is also a secure sketch in an other metric.
A binary linear $[n, k, d]$ code C is a vectorial sub-space of $\{0, 1\}^n$ having a dimension k and composed of vectors x having a Hamming weight $w_H(x) \ge d$, where $w_H(x)$ is the number of non-zero coordinates of x. The correction capacity of this code is $t = \lfloor (d-1)/2 \rfloor$. More details on error-correcting codes are given in the book (MacWilliams & Sloane, 1988).
In this construction, the metric space is $\{0, 1\}^n$, with the Hamming distance d_H. Let C be a binary linear code with parameters $[n, k, 2t + 1]$. Then a $(\{0, 1\}^n, m, m - (n - k), t)$-secure sketch is designed as follows:

- The function *SS* takes as input a value $b \in \{0, 1\}^n$ and outputs a sketch $P = c \oplus b$, where $c \in C$ is a randomly chosen codeword.

- The function *Rec* takes as input a sketch P and a value $b' \in \{0, 1\}^n$, decodes $b' \oplus P$ in a codeword c' et returns the value $c' \oplus P$.

The following authentication system is directly related to the previous secure sketch:

Biometric authentication with fuzzy commitment

1. **Enrollment:** the user registers his biometric template b and sends the sketch $P = c \oplus b$, with $H(c)$ to the database, where H is a cryptographic hash function and $c \in C$ is a codeword randomly chosen.

2. **Verification:** the user sends a new biometric template b' to the database which computes $P \oplus b'$. Then the database decodes it in a codeword c' and checks if $H(c) = H(c')$. In cas of equality, the user is authenticated.

According to the minimum distance $2t + 1$ of the code, the new biometric template b' is accepted if and only if the Hamming distance $d_H(b, b') \le t$. The authentication system is based on the following property: if the distance $d_H(b, b') = \epsilon$ is lower than the correction capacity of the code, then it is possible to recover the original codeword c from the word $c \oplus \epsilon$. Applications of this protocol are proposed in face recognition (Kevenaar et al., 2005) or fingerprints (Tuyls et al., 2005), using BCH codes. This fuzzy commitment scheme is also used for iris recognition, where iris templates are encoded by binary vectors of length 2048, called IrisCodes (Daugman, 2004a;b). For example a combination of Hadamard and Reed-Solomon codes is proposed in (Hao et al., 2006), whereas a Reed-Muller based product code is chosen in (Chabanne et al., 2007).
The previous scheme ensures the protection of the biometric template if the size of the code C is sufficient, whereas the loss of entropy of the template is directly connected to the

redundancy of the code. Security of this system is however limited: biometrics templates are not perfectly random and their entropy is difficult to estimate. Moreover, the protection of the biometric template is related to the knowledge of the random codeword c. This codeword is direclty used by the database during the verification phase.

In order to enhance the security of the previous scheme, Bringer *et al.* have proposed a combination of a secure sketch with a probabilistic encryption and a PIR protocol[1] in (Bringer & Chabanne, 2009; Bringer et al., 2007). The following biometric authentication protocol gives a simplified description of their scheme (without PIR protocols) and illustrates nicely the possibilities proposed by homomorphic encryptions for privacy enhancement in biometric authentication.

The Goldwasser-Micali probabilistic encryption scheme is the first probabilistic encryption scheme proven to be secure under cryptographic assumptions (Goldwasser & Micali, 1982; 1984). The semantic security of this scheme is related to the intractability of the quadratic residuosity problem . The Goldwasser-Micali scheme is defined as follows:

Definition 6. *Let p and q be two large prime numbers, N be the product $p.q$ and x be a non-residue with a Jacobi symbol 1. The public key of the crypto-system is $p_k = (x, N)$ and the private key is $s_k = (p, q)$. Let y be randomly chosen in $\mathbf{Z}_n^*$. A message $m \in \{0, 1\}$ is encrypted in c, where $c = Enc(m) = y^2 x^m \bmod n$. The decryption function Dec takes an encrypted message c and returns m, where $m = 0$ if c is a quatratic residue and 1 otherwise.*

This scheme encrypts a message bit by bit. The encryption of a message of n bits $m = (m_1, \dots, m_n)$ with the previous scheme is denoted $Enc(m) = (Enc(m_1), \dots, Enc(m_n))$, where the encryption mechanism is realized with the same key. The Goldwasser-Micali scheme clearly verifies the following property:

$$\text{Dec}(\text{Enc}(m, pk) \times \text{Enc}(m', pk), sk) = m \oplus m'.$$

This homomorphic property is used for the combination of this cryptosystem with the secure sketch construction of Juels and Wattenberg.

The biometric authentication scheme uses the following component: the user U who needs to be authenticated to a service provider SP. The service provider has access to a database where biometrics templates are stored. These templates are encrypted with cryptographic keys, generated and stored by a key manager KM who has no access to the database. For privacy reasons, the service provider SP has never access to the private keys.

For each user U, the key manager KM generates a pair (p_k, s_k) for the Goldwasser-Micali scheme. The public key p_k is published and the private key is stored in a secure way. The biometric authentication system is described as follows:

Biometric authentication with homomorphic encryption

1. **Enrollment:** The user U registers his biometric template b to the service provider. The service provider randomly generates a codeword $c \in C$, computes $H(c)$ where H is a cryptographic hash function and encrypts $Enc(c \oplus b)$ with the Goldwasser-Micali scheme and the public key p_k, and finally stores it in the database.

2. **Verification:** the user U encrypts his biometrics template $Enc(b')$ with p_k and sends it to the service provider. The service provider recovers $Enc(c \oplus b)$ and $H(c)$ from the database, computes and sends the products $Enc(c \oplus b) \times Enc(b')$ to the key manager. The key manager decrypts

---

[1] Private Information Protocol, see (Chor et al., 1998).

$Dec(Enc(c \oplus b) \times Enc(b')) = c \oplus b \oplus b'$ with its private key s_k and sends the result to the service provider who decodes it in a codeword c'. The service provider finally checks if $H(c) = H(c')$.

Homomorphic property of the Goldwasser-Micali scheme ensures that biometrics templates are never decrypted during the verification phase of the authentication protocol. Moreover, the service provider who has access to the encrypted biometric data does not possess the private key to retrieve the biometric templates and the key manager who generates and stores the private keys has never access to the database.

Other encryption schemes with suitable homomorphic property can be used as the Paillier cryptosystem (Paillier, 1999) or the Damgard-Jurik cryptosystem (Damgard & Jurik, 2001). Homomorphic cryptosystems have been recently used for several constructions of privacy-compliant biometric authentication systems. For example, a face identification system is proposed in (Osadchy et al., 2010), whereas iris and fingerprint identification mechanisms are described in (Barni et al., 2010; Blanton & Gasti, 2010).

3.3 BioHashing

The previous cryptosystems represent promising solutions to enhance the privacy. However, the crucial issues of cancelability and diversity seem to be not well addressed by these techniques (Simoens et al., 2009).

Besides biometric cryptosystems design, transformation based approaches seem more suited to ensure the cancelability and diversity requirements and more generally, fulfill the additional points raised page 7: non-reversibility, accuracy and randomness. The principle of transformation based methods can be explained as follows: instead of directly storing the raw original biometric data, it is stored after transformation relying on a non-invertible function. So, the prominent feature shared by these techniques takes place at the verification stage, which is performed in the transformation field, between the stored template and the newly acquired template. Moreover, these techniques are able to cope with the variability inherent to any biometrics template.

The pioneering work (Ratha et al., 2001) introduces a distortion of the biometric signal by a chosen transformation function. Hence, cancelability is ensured: each time a transformed biometric template is compromised, one has just to change the transformation function to generate a new transformed template. The diversity property is also guaranteed, since different transformation functions can be chosen for different applications.

Among the transformation based approaches, we detail in this chapter the principle of BioHashing. BioHashing is a two factor authentication approach which combines pseudo-random number with biometrics to generate a compact code per person. The first work referencing the BioHashing technique is presented on face modality in (Goh & Ngo, 2003). Then the same technique has been declined to different modalities in the references (Teoh et al., 2004c), (Teoh et al., 2004a), (Connie et al., 2004) and more recently (Belguechi, Rosenberger & Aoudia, 2010), to mention just a few.

Now, we detail the general principle of BioHashing.

3.3.1 BioHashing principle

All BioHashing methods share the common principle of generating a unitary BioCode from two data: the biometric one (for example texture or minutiae for fingerprint modality) and a random number which needs to be stored (for example on a usb key, or more generally on a token), called *tokenized random number*. The same scheme (detailed just below) is applied both:

- at the enrollment stage, where only the BioCode is stored, instead of the raw original biometric data
- at the verification stage, where a new BioCode is generated, from the stored random number

Then the verification relies on the computation of the Hamming distance between the reference BioCode and the newly issued one. This principle allows BioCode cancelability and diversity by using different random numbers for different applications.

More precisely, the BioHashing process is illustrated by the figure 5. One can see that it is a two factor authentication protection scheme, in the sense that the transformation function combines a specific random number whose seed is stored in a token with the biometric feature expressed as a fixed-length vector $F = (f_1, \ldots, f_n)$, $F \in \mathbb{R}^n$.
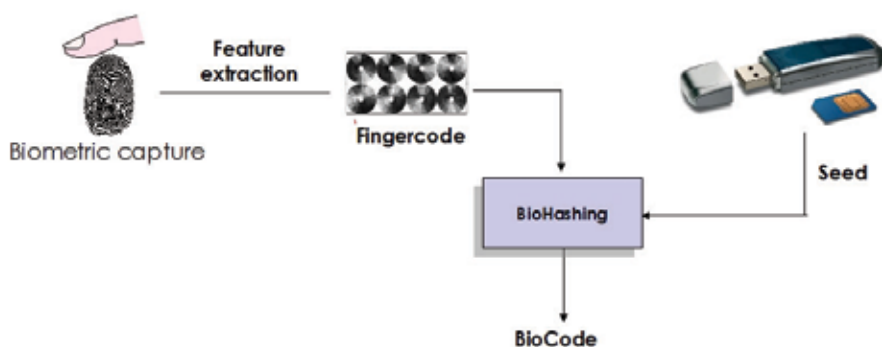
Fig. 5. BioHashing: Ratha's method

Then BioHashing principle, detailed in (Belguechi, Rosenberger & Aoudia, 2010) for example, consists in the projection of the (normalized) biometric data on an orthonormal basis obtained from the random number. This first step somehow amounts to hide the biometric data in some space. The second step relies on a quantization which ensures the non-invertibility of BioHashing: from the final BioCode, it is impossible to obtain the original biometric feature.

Let us give more details on the involved stages: random projection and quantization.

- *Random projection*
  It has been shown in (Kaski, 1998) that random mapping can preserve the distances in the sense that the inner product (which represents a way of measuring the similarity between two vectors from the cosine of the angle between them) between the mapped vectors closely follows the inner product of the initial vectors. One condition is that the involved random matrix R consists of random values and the Euclidean norm of each column is normalized to unity. The reference (Kaski, 1998) proves that the closer to an orthonormal matrix the random matrix R is, the better the statistical characteristics of the feature topology are preserved. As a consequence, in the BioHashing process, the tokenized random number is used as a seed to generate m random vectors $r_i, i = 1, \ldots, m$. After orthonormalization by the Gram-Schmidt method, these vectors are gathered as the column of a matrix $O = \left(O_{ij} \right)_{i,j \in [1,n] \times [1,m]}$.

The following Johnson-Lindenstrauss lemma (1984), studied in (Dasgupta & Gupta, 1999), (Teoh et al., 2008) is at the core of the BioHashing efficiency:

Lemma 1. *For any $0 < \varepsilon < 1$ and any integer k, let m be a positive integer verifying $m \geq \frac{4\log(k)}{\varepsilon^2/2 - \varepsilon^3/3}$. Then, for any set S containing k points in $\mathbb{R}^n$, there exists a map $f : \mathbb{R}^n \to \mathbb{R}^m$ such that:*

$$\forall x, y \in S, \ (1 - \varepsilon)||x - y||^2 \leq ||f(x) - f(y)||^2 \leq (1 + \varepsilon)||x - y||^2 \tag{1}$$

In other words, Johnson-Lindenstrauss lemma states that any n point set in Euclidian space can be embedded in suitably high (logarithmic in k, independent of n) dimension without distorting the pairwise distances by more than a factor of $1 \pm \varepsilon$. As a conclusion of this first step, we can say that the pairwise distances are well conserved by random projection under the previous hypotheses on the random matrix. Notice that this distance conservation becomes better when m increases, therefore one can consider $m = n$.

The resulting vector is denoted $W = (W_1, \ldots, W_m)$, with $W = F.O \in \mathbb{R}^m$, see figure 6.

- *Quantization*
  This step is devoted to the transformation in a binary-valued vector of the previous real-valued vector resulting from the projection of the original biometric data on an orthonormalized random matrix. A reinforcement of the non-invertibility (also relying on the random projection process) of the global BioHashing transformation ensues from this quantization. It requires the specification of a threshold τ to compute the final BioCode $B = (B_1, \ldots, B_m)$ following the formula:

$$B_i = \begin{cases} 0 \text{ if } W_i \leq \tau \\ 1 \text{ if } W_i > \tau \end{cases} \tag{2}$$

In practice, the threshold τ is chosen equal to zero so that half of W_i are larger than the threshold and half smaller. This, in order to maximize the information content of the extracted m bits and to increase the robustness of the resultant template. To this purpose, one may compute the median of the referenced vectors W and use it as a threshold.

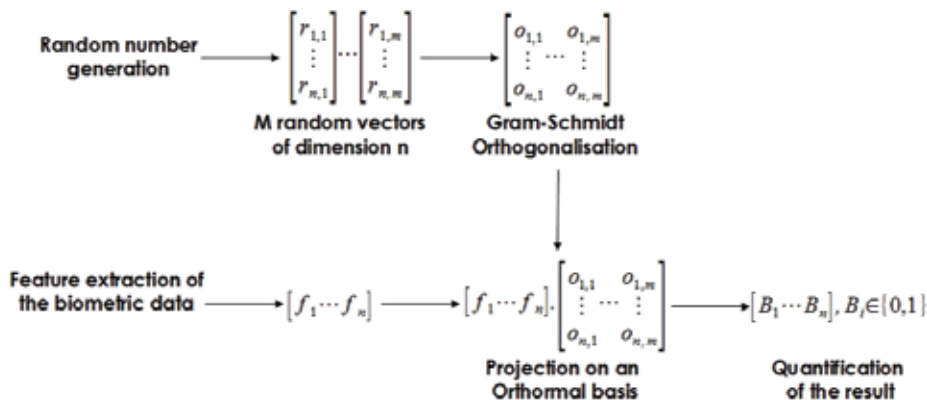These two steps are illustrated by the figure 6.

Fig. 6. BioCode generation

In the literature, one can find that the previous general technique has been applied to several biometric modalities to obtain the biometric template F. In (Teoh et al., 2004a), integrated Wavelet Fourier-Mellin transform is applied to fingerprint raw image. This requires the a

priori detection of the core point and produces a translation and rotation-invariant feature. Besides, integrated Wavelet Fourier-Mellin transform has been applied to face raw image in (Teoh & Ngo, 2005), while Fisher Discriminant Analysis (FDA) for face images is developed in (Teoh et al., 2004b), with a slightly different quantification step. Both Principal Component Analysis (PCA) and Fisher Discriminant Analysis are at the core of PalmHashing developed in (Connie et al., 2004) from the ROI of the raw palmprint image. In two recent papers (Belguechi, Rosenberger & Aoudia, 2010), (Belguechi, Hemery & Rosenberger, 2010), we propose to extract biometric templates from minutiae representation, by using a Gabor filterbank, after a detection process of the region of interest.

3.3.2 Discussion

The conclusion shared by the previously mentioned references is that BioHashing has significant advantages over solely biometrics or token usage, such as extremely clear separation of the genuine and the imposter population and zero EER (Equal Error Rate) level. But, among other papers, the reference (Kong et al., 2005) reveals that the outstanding 0% EER achievement of BioHashing implies the unrealistic assumption that the tokenized random number (TRN) would never be lost, stolen, shared or duplicated. In this paper, it is also pointed out that if this assumption held, the TRN alone could serve as a perfect password, making biometrics useless. The presented results show that the true performance of BioHashing is far from perfect and even can be worst than the basic biometric system. The results of some tests on different modalities are given in (Lumini & Nanni, 2006). For fingerprint texture template for example, the authors have demonstrated that the performance of the system in term of EER moves from 7.3% when no hashing is performed to 10.9% when basic BioHashing is operated under the hypothesis that the token is always stolen, while EER is evaluated to 1.5% in case where no TRN is stolen (FVC2002-DB2). These scores are also discussed in our papers (Belguechi, Rosenberger & Aoudia, 2010), (Belguechi, Hemery & Rosenberger, 2010) where different cases are considered, depending on whether the token is stolen or not.

3.4 Summarization

We saw in the previous sections different solutions to protect the biometric information. On the one hand using a Secure Element to store a biometric template is one convenient solution and is already operational. Even if it is technically possible to cancel a biometric template (by updating the content of the SE), this solution does not give any guarantee about the required cancelability properties. The security of a SE is often evaluated by a certification level (as for example EAL4+ for common criteria) giving some elements about the possibility for an hacker to obtain the biometric template.

On the other hand algorithmic solutions propose nice solutions to protect biometric templates privacy. Cryptography based approaches avoid the transmission of biometric templates but does not solve the non revocability problem. BioHashing reveals itself as a promising solution and seems to respect many privacy properties defined previously. This approach needs to be further studied, especially considering attacks.

4. Research challenges

Even if some solutions exist, there are many challenges to deal with in the future.

How can we evaluate cancelable biometric systems ?

Before proposing new privacy protection schemes for biometric templates, it becomes urgent to define objective evaluation methods for these particular systems. Of course, this type of

biometric systems can be evaluated through existing standards in performance evaluation (see (El-Abed et al., 2010) for example) but it is not sufficient. Computing the EER value or the ROC curve does not give any information on how the system protects the privacy of users. Some researchers try to analyze the security and privacy of these systems by taking into account some scenarios. The robustness to an attack is often quantified as for example by the resulting EER or FRR values when the attacker caught some additional information that he/she was not supposed to have. There is a lot of work on this subject.

How to increase the BioCode length?

In order to prevent brute force attack consisting in testing different values of the BioCode, it is necessary to increase the size of the generated BioCode. Many solutions to this problem exist. First, one simple solution is to use different biometric information. One can generate a BioCode for the fingerprint of each hand finger. There is no reason to have a statistical correlation between information provided by the template of each fingerprint. This solution solves the problem of the size and the associated entropy but it is less usable for an user as he/she has to provide as for example the fingerprint of each hand. Second, it is possible to increase the size of the BioCode by using an adapted representation. As for example, computing a BioCode from minutiae (where 30 are detected in average for a fingerprint) provides smaller BioCode compared to a texture representation. So it is necessary to carefully study biometric data representation.

How many times can we cancel a BioCode ?

The objective of a cancelable biometric information is to be able to generate it again in case of known attack. The question is to quantify the possibility to revoke this data a certain amount of times. Suppose an attacker is listening to the authentication session and can have different values of the BioCode, the problem is to know if he/she is able to predict an authorized BioCode. It is necessary to analyze as for example if some bits keep the same value in the BioCode after regeneration.

To what extent is it usable in an operational context ?

There are some (not so much) publications in this domain but very few industrial solutions (except those using a SE). This domain is not enough mature. Using a secure element to store the biometric data and realizing a match on card is well known. It could be interesting to combine a hardware solution using a secure element and an algorithmic solution. We are currently working on this aspect. The next step is also to be able to make a capture on card with an embedded biometric sensor to limit the transmission of the biometric data.

5. Conclusion

Biometrics is a very attractive technology mainly because of the strong relationship between the user and its authenticator. Unfortunately, many problems are also associated with this authentication solution. The main one concerns the impossibility to revoke a biometric data. Besides there is a major concern for ethical and security reasons. We presented in this chapter the main issues in this field and some solutions in the state of the art based on secure storage of the biometric template or using algorithmic solutions. Even if these methods bring some

improvements, many things need to be done in order to have a totally secure solution. We detailed some trends to work on in the near future.

6. References

Adler, A. (2007). *Biometric system security*, Handbook of biometrics, Springer ed.

Barni, M., Bianchi, T., Catalano, D., Raimondo, M. D., Labati, R. D., Failla, P., Fiore, D., Lazzeretti, R., Piuri, V., Piva, A. & Scotti, F. (2010). A privacy-compliant fingerprint recognition system based on homomorphic encryption and fingercode templates, *BTAS 2010*.

Belguechi, R., Hemery, B. & Rosenberger, C. (2010). Authentification révocable pour la vérification basée texture d'empreintes digitales, *Congrès Francophone en Reconnaissance des Formes et Intelligence Artificielle (RFIA)*.

Belguechi, R., Rosenberger, C. & Aoudia, S. (2010). Biohashing for securing minutiae template, *Proceedings of the 20th International Conference on Pattern Recognition*, Washington, DC, USA, pp. 1168–1171.

Blanton, M. & Gasti, P. (2010). Secure and efficient protocols for iris and fingerprint identification, Cryptology ePrint Archive, Report 2010/627. `http://eprint.iacr.org/`.

Bolle, R., Connell, J. & Ratha, N. (2002). Biometric perils and patches, *Pattern Recognition* 35(12): 2727–2738.

Bringer, J. & Chabanne, H. (2009). An authentication protocol with encrypted biometric data, *AfricaCrypt'09*.

Bringer, J., Chabanne, H., Izabachène, M., Pointcheval, D., Tang, Q. & Zimmer, S. (2007). An application of the Goldwasser-Micali cryptosystem to biometric authentication, *ACISP'07*, Vol. 4586 of *Lecture Notes in Computer Science*, Springer, pp. 96–100.

Cappelli, R., Lumini, A., Maio, D. & Maltoni, D. (2007). Fingerprint image reconstruction from standard templates, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(9): 1489–1503.

Chabanne, H., Bringer, J., Cohen, G., Kindarji, B. & Zemor, G. (2007). Optimal iris fuzzy sketches, *IEEE first conference on biometrics BTAS*.

Chor, B., Kushilevitz, E., Goldreich, O. & Sudan, M. (1998). Private information retrieval, *J. ACM* 45(6): 965–981.

Connie, T., Teoh, A., Goh, M. & Ngo, D. (2004). Palmhashing: a novel approach for dualfactor authentication, *Pattern analysis application* 7: 255–268.

Damgard, I. & Jurik, M. (2001). A generalisation, a simplification and some applications of paillier's probabilistic publickey system, *PKC'01*, Vol. 1992 of *Lecture Notes in Computer Science*, Springer, pp. 119–136.

Dasgupta, S. & Gupta, A. (1999). An elementary proof of the Johnson-Lindenstrauss Lemma. UTechnical Report TR-99-006, International Computer Science Institute, Berkeley, CA.

Daugman, J. (2004a). How iris recognition works, *Circuits and Systems for Video Technology, IEEE Transactions on* 14(1): 21–30.

Daugman, J. (2004b). Iris recognition and anti-spoofing countermeasures, *7-th International Biometrics conference*.

Dodis, Y., Katz, J., Reyzin, L. & Smith, A. (2006). Robust fuzzy extractors and authenticated key agreement from close secrets, *CRYPTO'06*, Vol. 4117 of *Lecture Notes in Computer Science*, Springer, pp. 232–250.

Dodis, Y., Reyzin, L. & Smith, A. (2004). How to generate strong keys from biometrics and other noisy data, *EUROCRYPT'04*, Vol. 3027 of *Lecture Notes in Computer Science*, Springer, pp. 523–540.

Domnesque, V. (2004). Carte d'identité électronique et conservation des données biométriques. Master thesis, Lille university.

El-Abed, M., Giot, R., Hemery, B. & Rosenberger, C. (2010). A study of users' acceptance and satisfaction of biometric systems, *IEEE International Carnahan Conference on Security Technology (ICCST'10)*, pp. 170–178.

Feng, J. & Jain, A. (2009). Fm model based fingerprint reconstruction from minutiae template, *International conference on Biometrics (ICB)*.

Galbally, J., Cappelli, R., Lumini, A., Maltoni, D. & Fiérrez-Aguilar, J. (2008). Fake fingertip generation from a minutiae template, *ICPR*, pp. 1–4.

GlobalPlatform (2006). *GlobalPlatform Card Specification Version 2.2*.

Goh, A. & Ngo, C. (2003). *Computation of Cryptographic Keys from Face Biometrics*, Vol. 2828 of *Lecture Notes in Computer Science*, Springer, Berlin.

Goldwasser, S. & Micali, S. (1982). Probabilistic encryption and how to play mental poker keeping secret all partial information, *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pp. 365–377.

Goldwasser, S. & Micali, S. (1984). Probabilistic encryption, *Journal of Computer and System sciences* 28(2): 270–299.

Hao, F., Anderson, R. & Daugman, J. (2006). Combining crypto with biometrics effectively, *IEEE Transactions on Computers* 55(9): 1081–1088.

Jain, A., Nandakumar, K. & Nagar, A. (2008). Biometric template security, *EURASIP J. Adv. Signal Process 2008*.

Jain, A., Ross, A. & Pankanti, S. (2006). Biometrics: A tool for information security, *IEEE Transactions on Information Forensics and Security* 1(2): 125–143.

Juels, A. & Sudan, M. (2001). A fuzzy vault scheme, *IEEE International Symposium on Information Theory*.

Juels, A. & Wattenberg, M. (1999). A fuzzy commitment scheme, *ACM conference on Computer and communication security*, pp. 28–36.

Kaski, S. (1998). Dimensionality reduction by random mapping: fast similarity computation for clustering, *Proc. of the International Joint Conference on Neural Networks*, Vol. 1, pp. 413–418.

Kevenaar, T., Schrijen, G., van der Veen, M., Akkemans, A. & Zuo, F. (2005). Face recognition with renewable and privacy preserving binary templates, *IEEE workshop on Automatic Identification Advanced Technologies*, pp. 21–26.

Kong, A., Cheung, K., Zhang, D., Kamel, M. & You, J. (2005). An analysis of biohashing and its variants, *Pattern Recognition* 39.

Lumini, A. & Nanni, L. (2006). Empirical tests on biohashing, *NeuroComputing* 69: 2390–2395.

MacWilliams, F. & Sloane, N. (1988). *The Theory of Error-correcting codes*, North-Holland.

Madlmayr, G., Dillinger, O., Langer, J. & Schaffer, C. (2007). The benefit of using sim application toolkit in the context of near field communication applications, *ICMB'07*.

Maltoni, D., Maio, D., Jain, A. & Prabhakar, S. (2009). *Handbook of Fingerprint Recognition*, Springer.

Monitor, N. (2002). 2002 nta monitor password survey.

Mordini, E. & Massari, A. (2008). Body, biometrics and identity, *Bioethics Journal* 22(9): 488–494.

O'Gorman, L. (2003). Comparing passwords, tokens, and biometrics for user authentication, *Proceedings of the IEEE* 91(12): 2021 – 2040.

Osadchy, M., Pinkas, B., Jarrous, A. & Moskovich, B. (2010). Scifi - a system for secure face identification, *IEEE Symposium on Security and Privacy*.

Paillier, P. (1999). Public-key cryptosystems based on composite degree residuosiy classes, *EUROCRYPT'99*, Vol. 1592 of *Lecture Notes in Computer Science*, Springer, pp. 223–238.

Pakanti, S., Prabhakar, S. & Jain, A. K. (2002). On the individuality of fingerprint, *IEEE Trans. Pattern Anal. Machine Intell.* 24(8): 1010–1025.

Ratha, N., Connelle, J. & Bolle, R. (2001). Enhancing security and privacy in biometrics-based authentication system, *IBM Systems J.* 37(11): 2245–2255.

Schneier, B. (1999). Inside risks: the uses and abuses of biometrics, *Commun. ACM* 42: 136.

Simoens, K., Chang, C. & Preneel, B. (2009). Privacy weaknesses in biometric sketches, *30th IEEE Symposium on Security and Privacy*.

Solove, D. (2009). *Understanding privacy*, Harvard university press.

Teoh, A., Kuanb, Y. & Leea, S. (2008). Cancellable biometrics and annotations on biohash, *Pattern recognition* 41: 2034–2044.

Teoh, A. & Ngo, D. (2005). Cancellable biometrics featuring with tokenised random number, *Pattern Recognition Letters* 26: 1454–1460.

Teoh, A., Ngo, D. & Goh, A. (2004a). Biohashing: two factor authentication featuring fingerprint data and tokenised random number, *Pattern recognition* 40.

Teoh, A., Ngo, D. & Goh., A. (2004b). An integrated dual factor authenticator based on the face data and tokenised random number, *1st International conference on biometric authentication (ICBA), Hong Kong*.

Teoh, A., Ngo, D. & Goh, A. (2004c). Personalised cryptographic key generation based on facehashing, *Computers and Security Journal* 23(07): 606–614.

Tuyls, P., Akkemans, A., Kevenaar, T., Schrijen, G., Bazen, A. & Veldhuis, R. (2005). Practical biometric authentication with template protection, *Audio and Video based Personal Authentication*, pp. 436–446.

Warren & Brandeis (1890). The right to privacy. Harvard Law Review (IV).

Protection of the Fingerprint Minutiae

Woo Yong Choi[1], Yongwha Chung[2] and Jin-Won Park[3]
[1]Electronics and Telecommunications Research Institute (ETRI),
[2]Korea University,
[3]Hongik University
Republic of Korea

1. Introduction

With a growing concern regarding security, interest in biometrics is increasing. Since biometrics utilizes a user's physiological or behavioral characteristic, which is unique and immutable, the compromise of biometric templates is a serious problem. Fingerprint authentication system is one of the most widely used biometric authentication systems. In general, in the enrollment procedure, the features are extracted from the enrollment image and are stored as a template. The template is compared to the features extracted from the verification image. Unlike passwords, however, biometrics has no or little substitutions. For example, if one's fingerprint template is compromised, he or she cannot use that fingerprint for any other fingerprint authentication system from then on.

Ratha et al. have introduced cancelable biometrics as a remedy for the problem of compromised templates (Bolle et al., 2002; Ratha et al., 2001). Cancelable biometrics distorts or transforms a user's template using some non-invertible functions to obscure the user's raw physical characteristics, and its matching is performed in a transformed domain. When a template is compromised, a new biometric template is issued (like a new enrollment of a new user) by distorting the biometric traits in a different way using a new instance of the non-invertible function. Ratha et al. proposed the surface folding scheme for cancelable fingerprint templates (Ratha et al., 2007). They proposed a one-way transformation which moves minutia positions using two-dimensional Gaussian functions defined over the feature domain. However, if an attacker obtains two transformed templates and transformation parameters, the original template is recovered by a dictionary attack (Shin et al., 2009).

Fuzzy vault is a crypto-biometric algorithm proposed by Juels et al. (Juels & Sudan, 2002). It gives a promising solution to personal privacy and fingerprint template security problems. Clancy et al. and Uludag et al. suggested the method for applying the fuzzy vault to fingerprint authentication, which is named as *fuzzy fingerprint vault* (Clancy et al, 2003; Uludag et al., 2005). It generates a lot of chaff minutiae and mixes them up with the real minutiae. Then, the real minutiae are projected on a randomly generated polynomial, and the chaff minutiae are projected off the polynomial. The polynomial is successfully reconstructed using either brute-force search or Reed-Solomon code if a sufficient number of real minutiae are chosen. The genuine user can choose a sufficient number of real minutiae by presenting his or her fingerprint while the impostors cannot. Some researchers have implemented the fuzzy vault for fingerprints, and have protected the fingerprint minutiae by adding chaff points into the vault (Chung et al., 2006; Clancy et al., 2003; Dodis et al.,

2004; Kanak & Sogukpinar, 2007; Nandakumar et al., 2007; Uludag et al., 2005). Lagrange interpolation (Hildebrand, 1987) is the most widely used polynomial interpolation method. However, it requires a little much time especially when the degree of polynomial is large. Brute-force search is employed for polynomial reconstruction attempts until the true polynomial is reconstructed, and Lagrange interpolation is used to interpolate the polynomial. Therefore, even if the real minutiae are chosen more than the degree of the polynomial, the brute-force search cannot reconstruct the polynomial in real-time when several chaff minutiae are chosen along with the real minutiae. All the previous results adopted the brute-force search to reconstruct the polynomial or skipped the procedure for polynomial reconstruction because of its difficulty (Li et al., 2006). In this work we propose a fast algorithm for polynomial reconstruction. To reduce the execution time, it determines the candidate sets with chaff points by using the Gaussian elimination and excludes them from the reconstruction trial. Since the Gaussian elimination is a time-consuming process, we have found a recursive formula to perform the Gaussian elimination effectively by using the Consistency Theorem (Anton, 1994). We confirmed that the proposed algorithm can be performed in real time even at the worst case.

There are a few attack methods on the fuzzy vault. Scheirer et al. suggested the methods of attacks against fuzzy vault including the attack via record multiplicity, which is known as the *correlation attack* (Scheirer & Boult, 2007). The correlation attack gives very powerful attack method when two fuzzy vaults obtained from the same fingerprint are available. On the other hand, when only one fuzzy vault is compromised and no additional information is available, brute-force attack can be used. Brute-force attack is employed for polynomial reconstruction attempts using the random combination of points until the true polynomial is reconstructed. In this work we propose a new attack algorithm which applies a fast polynomial reconstruction algorithm based on the Consistency Theorem. Also, we evaluate the proposed attack method, and compare it with the known attack methods such as the brute-force attack and the correlation attack.

This chapter is organized as follows. The conventional fingerprint authentication methods are described in Section 2. Section 3 presents the fuzzy fingerprint vault system and the proposed polynomial reconstruction algorithm followed by experimental results. In Section 4, various attack methods for fuzzy fingerprint vault (brute-force attack, correlation attack, and the fast polynomial reconstruction attack) and the experimental results are explained. Section 5 summarizes the conclusion.

2. Fingerprint authentication

Fingerprint recognition is the most common biometric method for authentication. Since everyone's fingerprint is unique and invariable during life, fingerprint has been used as the evidence of forensic science and the personal authentication method. Modern fingerprint recognition began in 1684, when Nehemiah Grew studied and described ridges, furrows and pores on hand and foot surfaces. In the late 1960s, the Live-Scan system which records fingerprints electronically was developed, which was a turning point of fingerprint recognition. Fingerprint was centralized into database, and the automatic fingerprint recognition system has been developed consistently.

A fingerprint authentication system consists of fingerprint sensor, pre-processing, feature extraction, storage, and matching as shown in Fig. 1. The fingerprint pattern is captured by a fingerprint sensor. A fingerprint sensor takes a snapshot of a fingerprint and saves it into an

image file. From the image, unique features of each fingerprint are extracted and saved in the storage. The storage may be either a central database or a smartcard. Before feature extraction, pre-processing is performed to extract the reliable features from the image. For fingerprint matching, features of an input fingerprint are compared to the features of the enrolled fingerprint data. By comparing similarity between two fingerprint feature sets, it is decided whether the two fingerprints are from the same person or not.

The fingerprint recognition algorithms can be classified into two categories: image-based and minutiae-based. Image-based methods are based on optical correlation and transform based features. The image-based methods lack the ability to track with variations in position, scale, and orientation angle, and hence, they cannot give reliable recognition accuracy. Nevertheless, the image-based methods have been studied continuously because of the following properties (Nanni & Lumini, 2009; Yang & Park, 2008). First, the image-based methods can be combined with the minutiae-based methods to improve the accuracies of the fingerprint authentication systems. Second, fingerprint features can be represented by a fixed length vector, which is suitable for various learning systems.

Minutiae-based methods are more popular matching techniques, which are included in almost all contemporary fingerprint identification and verification systems. They are based on the minutiae, such as ending, bifurcation, and singular points in the fingerprint, which have been known to be effective for fingerprint recognition. The minutiae form a pattern of points, and hence several well-known point pattern matching algorithms have been proposed in the late 80's.
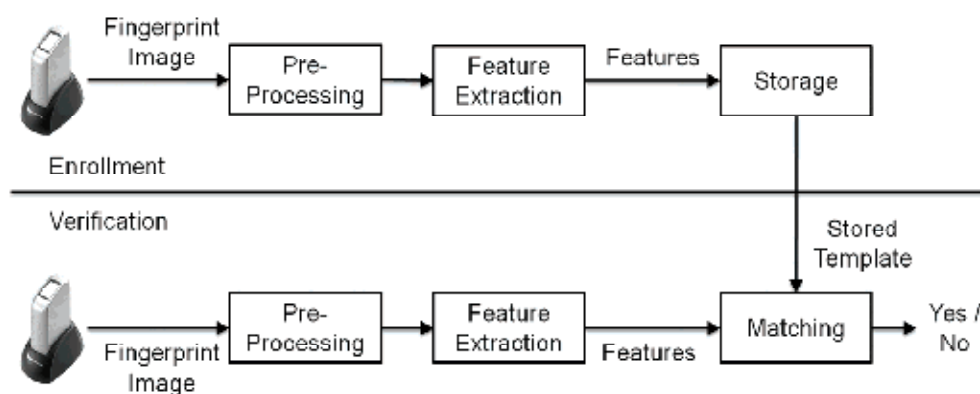
Fig. 1. Block diagram of the fingerprint authentication system (Pan et al., 2003)

A fingerprint authentication system has two phases: *enrollment* and *verification*. In the off-line enrollment phase, an enrolled fingerprint image is preprocessed, and the minutiae are extracted and stored. In the on-line *verification* phase, the similarity between the enrolled minutiae and the input minutiae is examined.

Image preprocessing refers to the refinement of the fingerprint image against the image distortion (poor contrast, flaw, smudge, etc.) obtained from a fingerprint sensor. Minutiae Extraction refers to the extraction of features from the fingerprint image. After this step, some of the minutiae are detected and stored into a pattern file, which includes the position, the orientation, and the type (ridge ending or bifurcation) of the minutiae.

The input fingerprint minutiae are compared with the enrolled fingerprint minutiae. Actually, Minutiae Matching is composed of the *alignment* stage and the *matching* stage. In

order to match two fingerprints captured with unknown direction and position, the differences of the direction and the position between two fingerprints should be detected, and alignment between them needs to be executed. Therefore, in the alignment stage, transformations such as translation and rotation between two fingerprints are estimated, and two minutiae are aligned according to the estimated parameters. If alignment is performed accurately, the matching stage is referred to point matching simply. In the matching stage, two minutiae are compared based on their position, orientation, and type. Then, a matching score is computed.

3. Fuzzy fingerprint vault

3.1 Enrollment procedure

Fig. 2 shows the block diagram of the enrollment procedure of the Fuzzy Fingerprint Vault (FFV) system. Given the fingerprint image to be enrolled, we first extract minutiae from the image to form a locking set of the form.

$$L = \{\mathbf{m}_i \mid 1 \leq i \leq n_e\} \tag{1}$$

where $\mathbf{m}_i = (x_i, y_i, \theta_i, t_i)$ is the i-th enrollment minutia, and n_e is the number of the enrollment minutiae. Then, a number of chaff minutiae are generated and constitute a minutia set along with the real minutiae. After adding the chaff minutiae, the total number of minutiae is n_r. All arithmetic operations are conducted in a finite field of order 2^{20}, namely GF(2^{20}). Thus, each coordinate is scaled to the range [0, 1024] for the purpose of the arithmetic in GF(2^{20}). A finite field (Stallings, 2005) is a field with a finite number of elements, also called a Galois field. All operations performed in the finite field result in an element within that field. For polynomial representation, we define the finite field over the irreducible polynomial $x^{20} + x^3 + 1$. Then, we randomly select $(k + 1)$ elements from GF(2^{20}) and generate a k-degree polynomial as follows.

$$p(u) = a_0 + a_1 u + a_2 u^2 + \cdots + a_k u^k \tag{2}$$
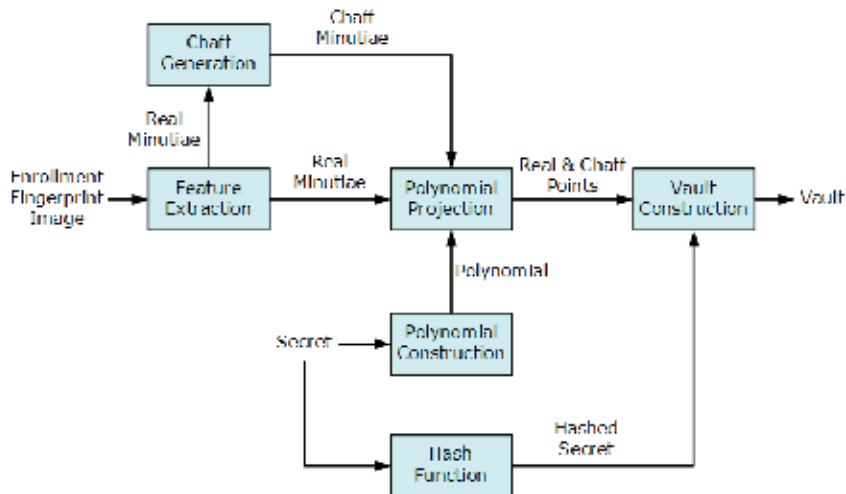
Fig. 2. Block diagram of the enrollment procedure of the FFV system (Choi et al., 2009)

This polynomial becomes the secret to be protected. As in the work of Uludag et al., we concatenate x and y coordinates of a minutia to arrive at the locking/unlocking data unit u. Then, we project the real and the chaff points (i.e., minutiae) on and off the polynomial, respectively. That is,

$$v_i = \begin{cases} p(u_i) & \text{if } u_i \text{ is real} \\ p(u_i) + \delta_i & \text{if } u_i \text{ is chaff} \end{cases} \tag{3}$$

where δ_i is a non-zero element of GF(2^{20}). Finally, the vault is constituted by the real and the chaff points, and the secret. The secret should be stored in a hashed form, instead of in the clear form.

3.2 Verification procedure
Fig. 3 shows the block diagram of the verification procedure of the FFV system. Given the fingerprint image to be verified, the minutiae are first extracted from the image and the verification minutiae set V is denoted by

$$V = \{\tilde{\mathbf{m}}_i \mid 1 \le i \le n_v\} \tag{4}$$

where $\tilde{\mathbf{m}}_i = (\tilde{x}_i, \tilde{y}_i, \tilde{\theta}_i, \tilde{t}_i)$ is the i-th verification minutia, and n_v is the number of the verification minutiae. Then, the verification minutiae are compared with the enrolled minutiae with real and chaff minutiae mixed, and an unlocking set U is finally selected.

$$U = \{\mathbf{m}_i \mid 1 \le i \le n_m\} \tag{5}$$

where n_m is the number of the matched minutiae. The vault can be successfully unlocked only if U overlaps with L to a great extent.
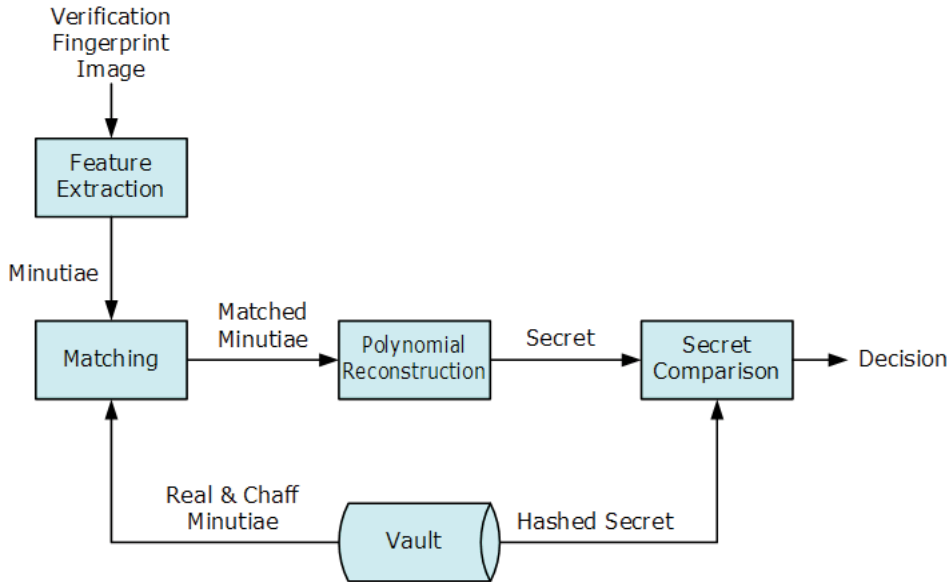

Fig. 3. Block diagram of the verification procedure of the FFV system (Choi et al., 2009)

These n_m points may contain some chaff points as well as the real points even if the user is genuine. Hence, in order to interpolate the k-degree polynomial, we have to select $(k + 1)$ real points from among the n_m points. After the polynomial is interpolated, it is compared with the true polynomial stored in the vault. A decision to accept/reject the user depends on the result of this comparison. If $|U \cap L| \geq (k + 1)$, the k-degree polynomial can be successfully reconstructed by using the brute-force search. The most widely used algorithm for polynomial interpolation is the Lagrange interpolation. The number of cases that select $(k + 1)$ minutiae from n_m minutiae is $C(n_m, k + 1)$. Let n_{real} be the number of real minutiae in set U, then the number of cases that correctly reconstruct the polynomial is $C(n_{real}, k + 1)$. Therefore, the average number of polynomial interpolation is

$$\frac{C(n_m, k+1)}{C(n_{real}, k+1)} \tag{6}$$

Furthermore, when a higher degree of polynomial is used, the Lagrange interpolation needs much more time to reconstruct the polynomial. More precisely, it can be done in $O(k \log^2(k))$ operations (Gathen & Gerhardt, 2003). Hence, it becomes impracticable as n_m and/or k increases. So, Uludag used only 18 minutiae to prevent n_m from being too large (Uludag et al., 2005).

Juels et al. suggested that the chaff points can be removed by means of the Reed-Solomon decoding algorithm (Juels & Sudan, 2002). Among various realizations of the Reed-Solomon code, we used the Gao's algorithm (Gao, 2003), which is one of the fastest Reed-Solomon decoding algorithms. It compensates one chaff point at the cost of one real point, so if there are many chaff points are matched along with real points, it is quite probable that the right user is rejected (i.e., false negative). The situation becomes much more serious when the degree of polynomial increases.

3.3 Polynomial reconstruction

If the matched point set of equation (5) contains more than $(k + 1)$ real point, the true polynomial can be reconstructed by using the brute-force search. Brute-force search chooses $(k + 1)$ points from among n_m points and tries to reconstruct the k-degree polynomial using the Lagrange interpolation, which requires a relatively large number of computations. So, when the matched point set contains many chaff minutiae, the number of the Lagrange interpolation to be performed increases exponentially, and hence, the polynomial reconstruction cannot be performed in real time.

In this section, we introduce a fast algorithm for the polynomial reconstruction (Choi et al., 2008). To begin with, let us consider the following theorem which provides the conditions under which a linear system of m equations in n unknowns is guaranteed to be consistent.

> **Consistency Theorem.** If **Ax = b** is a linear system of m equations with n unknowns, then the followings are equivalent.
> (a) **Ax = b** is consistent.
> (b) **b** is in the column space of **A**.
> (c) The coefficient matrix **A** and the augmented matrix **[A | b]** have the same rank.

Let us consider a linear system of $(k + 2)$ equations with $(k + 1)$ unknowns.

$$\mathbf{U} \cdot \mathbf{a} = \mathbf{v} \tag{7}$$

where

$$\mathbf{U} = \begin{bmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k \\ 1 & u_2 & u_2^2 & \cdots & u_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_{k+2} & u_{k+2}^2 & \cdots & u_{k+2}^k \end{bmatrix}, \ \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{k+2} \end{bmatrix} \tag{8}$$

Then, the corresponding augmented matrix $\mathbf{W}$ is of the form.

$$\mathbf{W} = \begin{bmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k & \vdots & v_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & u_{k+2} & u_{k+2}^2 & \cdots & u_{k+2}^k & \vdots & v_{k+2} \end{bmatrix} \tag{9}$$

It is straightforward that the rank of matrix $\mathbf{U}$ is $(k + 1)$. According to the Consistency Theorem, the rank of the augmented matrix $\mathbf{W}$ must be equal to $(k + 1)$ to guarantee that the linear system has a solution. The Gaussian elimination was used to check whether the augmented matrix had rank $(k + 1)$ or not. The elementary row operations, provided we do not perform the operation of interchanging two rows, were used to reduce the augmented matrix $\mathbf{W}$ into the row-echelon form.

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k & \vdots & v_1 \\ 0 & 1 & u_2^{2(2)} & \cdots & u_2^{k(2)} & \vdots & v_2^{(2)} \\ 0 & 0 & 1 & \cdots & u_3^{k(3)} & \vdots & v_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \vdots & v_{k+1}^{(k+1)} \\ 0 & 0 & 0 & \cdots & 0 & \vdots & v_{k+2}^{(k+2)} \end{bmatrix} \tag{10}$$

where $u_j^{l(i)}$ and $v_j^{(i)}$ are the values of u_j^l and v_j when the j-th row has "leading 1" at the i-th element, respectively. Note that the diagonal elements of equation (10) cannot be zero because the rank of matrix $\mathbf{U}$ is $(k + 1)$. From the parts (a) and (c) of the Consistency Theorem, it follows that if $v_{k+2}^{(k+2)} \neq 0$, the linear system of equation (7) does not have a solution. Hence, there exists at least one chaff point in the set $\{(u_i, v_i) \mid 1 \leq i \leq k + 2\}$, and we need not perform the polynomial reconstruction process. On the contrary, if $v_{k+2}^{(k+2)} = 0$, then all the points are probably the real points. Thus, we try to reconstruct the polynomial with $(k + 1)$ points and compare it with the true polynomial.

Up to this point, we have explained how to reconstruct a k-degree polynomial from $(k + 2)$ matched minutiae. In general, the unlocking set has n_m minutiae, so let us consider a linear system of n_m equations with $(k + 1)$ unknowns as follows.

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k \\ 1 & u_2 & u_2^2 & \cdots & u_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_{n_m} & u_{n_m}^2 & \cdots & u_{n_m}^k \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n_m} \end{bmatrix} \tag{11}$$

where $n_m > (k + 1)$. Clearly, if $n_m < (k + 1)$, then the polynomial cannot be reconstructed. Also, if $n_m = (k + 1)$, we can reconstruct the polynomial with the $(k + 1)$ points. Suppose that we select $(k + 2)$ real points from among n_m points, we can reconstruct the true polynomial. However, if at least one chaff point exists in the $(k + 2)$ selected points the true polynomial cannot be reconstructed. The procedure for the proposed polynomial reconstruction algorithm is as follows.

1. $(k + 1)$ points are selected from among n_m points with real and chaff points mixed, and these points are placed to the top of equation (11).
2. The augmented matrix of equation (11) is obtained, and is reduced into the following row-echelon form.

$$\begin{bmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^k & \vdots & v_1 \\ 0 & 1 & u_2^{2(2)} & \cdots & u_2^{k(2)} & \vdots & v_2^{(2)} \\ 0 & 0 & 1 & \cdots & u_3^{k(3)} & \vdots & v_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \vdots & v_{k+1}^{(k+1)} \\ 0 & 0 & 0 & \cdots & 0 & \vdots & v_{k+2}^{(k+2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \vdots & v_{n_m}^{(k+2)} \end{bmatrix} \tag{12}$$

3. To determine whether the $(k + 1)$ points $(u_1, v_1), \cdots, (u_{k+1}, v_{k+1})$ are valid candidates or not, we will check the values of $v_{k+2}^{(k+2)}, \cdots, v_{n_m}^{(k+2)}$. Therefore, the proposed algorithm needs one more real point than the brute-force search to reconstruct the polynomial. If there is at least one zero among $v_{k+2}^{(k+2)}, \cdots, v_{n_m}^{(k+2)}$, we reconstruct the polynomial with the selected $(k + 1)$ points, and compare it with the true polynomial.
4. Steps (1) ~ (3) are repeated until the true polynomial is reconstructed.

However, the computations of the Gaussian elimination in step (2) take too much time to be implemented in real time. Fortunately, in order to obtain the values of $v_{k+2}^{(k+2)}, \cdots, v_{n_m}^{(k+2)}$, we do not have to apply the Gaussian elimination. We have found the following recursive formula, so the computation time can be considerably reduced.

$$v_j^{(i+1)} = \begin{cases} v_j, & i = 0 \\ \dfrac{v_j^{(i)} - v_i^{(i)}}{u_j - u_i}, & i = 1, \cdots, \min(k + 1, j - 1) \end{cases} \tag{13}$$

This gives exactly the same solution as the Gaussian elimination. The proof can be found on the reference (Choi et al., 2008).

3.4 Experimental results

To evaluate the performance of the proposed polynomial reconstruction algorithm, we used DB1 and DB2 of FVC2002 (Maio et al., 2002). The fingerprint images were obtained in three distinct sessions with at least two week time separating for each session. During the first session, the fingerprint images were obtained by placing the fingerprints with a normal position. During the second session, the fingerprint images were obtained by requesting the individuals to provide their fingerprints with exaggerated displacement and rotation (not to

exceed 35 degrees). During the third session, fingers were alternatively dried and moistened. The characteristics of the databases are listed in Table 1. The databases were obtained from the optical sensors. The size of fingerprint image of DB2 is greater than that of DB1. The resolutions of the databases are about 500 dpi. Each database consists of 100 fingers, and 8 impressions per finger. Each sample was matched against the remaining samples of the same finger to compute the Genuine Acceptance Rate (GAR). Similarly, the first sample of each finger was matched against the first sample of the remaining fingers to compute the False Acceptance Rate (FAR). If the matching g against h is performed, the symmetric one (i.e., h against g) is not executed to avoid the correlation. For each database, the number of genuine tests was 2800, whereas the number of impostor tests was 4950. All experiments were performed on a system with a 3.2 GHz processor.

	DB1	DB2
Sensor Type	Optical	Optical
Image Size	388 × 374(142 Kpixels)	296 × 560(162 Kpixels)
Resolution	500 dpi	569 dpi
Sensor	"TouchView II" by Identix	"FX2000" by Biometrika
No. Fingers	100	100
No. Impressions	8	8

Table 1. Characteristics of FVC2002 Databases

Database	No. Minutiae (n_e)		
	Average	Min	Max
DB1	30.5	5	60
DB2	39.0	7	89

Table 2. The number of minutiae for FVC2002 Databases

Database	No. Chaff Minutiae	Test	No. Matched Minutiae (n_m)			Matching Time (sec)
			Total	Real	Chaff	
DB1	200	Genuine	19.7	18.8	0.9	0.76
		Impostor	9.3	3.3	5.9	0.87
	300	Genuine	20.0	18.6	1.4	1.48
		Impostor	10.9	2.9	8.0	1.73
	400	Genuine	20.4	18.4	2.0	2.45
		Impostor	12.3	2.5	9.8	2.89
DB2	200	Genuine	24.3	23.4	0.9	1.26
		Impostor	10.6	4.3	6.3	1.64
	300	Genuine	24.5	23.2	1.3	2.38
		Impostor	12.4	3.9	8.5	3.17
	400	Genuine	24.9	23.1	1.8	3.87
		Impostor	13.9	3.4	10.5	5.26

Table 3. Average number of matched minutiae and matching time for FVC2002 databases

To examine the effect of the insertion of chaff minutiae on the performance of a fingerprint recognition system, we selected the number of chaff minutiae as 200, 300 and 400. For each database the numbers of minutiae (average, minimum and maximum) are listed in Table 2. Since the images of DB2 are obtained from bigger sensor, more minutiae are extracted from DB2 than from DB1. Also, the average numbers of the matched minutiae according to the number of inserted chaff minutiae are listed in Table 3. The more chaff minutiae are added, the more minutiae are matched. In addition, the number of chaff minutiae increases while the number of real minutiae decreases slightly. Hence, we can predict that both of GAR and FAR will decrease as more chaff minutiae are added. Also, we can find that the matching time increases greatly as more chaff minutiae are added.

Polynomial degree	No. Chaff	FTER	Brute-force			Proposed			Reed-Solomon		
			GAR	FAR	HTER	GAR	FAR	HTER	GAR	FAR	HTER
7	200	0.3	93.1	7.4	7.1	92.2	4.8	6.3	92.0	2.6	5.3
	300		92.3	5.4	6.6	91.3	3.1	5.9	90.8	1.1	5.1
	400		91.0	4.1	6.6	90.2	2.3	6.1	89.2	0.2	5.5
8	200	0.6	90.7	4.1	6.7	89.6	2.2	6.3	89.4	1.3	6.0
	300		90.2	3.1	6.4	89.0	1.5	6.2	88.3	0.5	6.1
	400		89.2	2.1	6.5	88.0	0.9	6.5	86.9	0.1	6.6
9	200	1.3	88.1	2.2	7.0	87.0	1.2	7.1	86.7	0.7	7.0
	300		87.9	1.2	6.7	86.6	0.5	7.0	86.0	0.1	7.1
	400		87.3	0.9	6.8	85.4	0.5	7.6	84.2	0.0	7.9
10	200	1.9	85.0	0.8	7.9	84.0	0.4	8.2	83.5	0.3	8.4
	300		84.7	0.6	8.0	83.5	0.2	8.3	82.6	0.0	8.7
	400		83.8	0.5	8.3	81.3	0.1	9.4	78.2	0.0	10.9
11	200	2.6	82.0	0.2	9.1	80.4	0.1	9.9	79.4	0.1	10.4
	300		81.7	0.3	9.3	79.2	0.1	10.5	76.8	0.0	11.6
	400		81.1	0.1	9.5	78.5	0.1	10.8	75.1	0.0	12.4
12	200	3.1	78.9	0.1	10.6	76.9	0.0	11.6	75.8	0.0	12.1
	300		78.5	0.1	10.8	75.9	0.0	12.1	72.9	0.0	13.5
	400		77.6	0.1	11.2	74.6	0.0	12.7	71.2	0.0	14.4

Table 4. Recognition accuracies of the FFV system of FVC2002-DB1 (unit: %)

To examine the effectiveness of the proposed polynomial reconstruction algorithm, we compare it with both the brute-force search and the Reed-Solomon code. The error rates of the FFV system for DB1 and DB2 are listed in Table 4 and Table 5, respectively. During the enrollment, if the number of the fingerprint minutiae is less than or equal to the degree of the polynomial, the fingerprint is rejected to be enrolled. Failure To Enrollment Rate (FTER) is the ratio of the rejected fingerprints to the total fingerprints. The FTER of DB1 is much higher than that of DB2 since fewer minutiae are extracted from DB1. To compare the recognition accuracies of the three polynomial reconstruction algorithms, Genuine Acceptance Rate (GAR) and False Rejection Rate (FAR) are used. In addition, Half Total Error Rate (HTER) is adopted for the purpose of direct comparison (Poh & Bengio, 2006), which is the average of False Rejection Rate (FRR) and False Acceptance Rate (FAR). The values of GAR, FAR and HTER are obtained by excluding the fingerprints rejected at the enrollment phase. As predicted above, the more chaff minutiae is inserted, the lower both

GAR and FAR become. Since the proposed algorithm needs one more real minutia than the brute-force search, the recognition accuracy of the proposed algorithm using the k-degree polynomial should be exactly the same as that of the brute-force search using the $(k + 1)$-degree polynomial. In practice, however, the polynomial could not always be reconstructed even if the real minutiae are matched more than the degree of polynomial because the polynomial reconstruction process will be stopped after a pre-determined number of iterations to prevent from going into an infinite loop. Furthermore, another reason is the difference in FTER due to the different degree of polynomial. The overall recognition rates of the three algorithms are comparable. The averages of HTER of the brute-force search, the proposed algorithm and the Reed-Solomon code for DB1 are 8.1%, 8.5% and 8.8%, respectively, and 6.1%, 5.8% and 5.4% for DB2. The Reed-Solomon code is better for low degree polynomials, and the brute-force search is better for high degree polynomials.

Polynomial degree	No. Chaff	FTER	Brute-force			Proposed			Reed-Solomon		
			GAR	FAR	HTER	GAR	FAR	HTER	GAR	FAR	HTER
7	200	0.3	96.9	15.4	9.3	96.1	9.3	6.6	95.2	2.5	3.7
	300		96.3	11.3	7.5	95.3	7.0	5.8	94.2	1.1	3.5
	400		95.4	9.3	6.9	94.5	5.7	5.6	93.2	0.5	3.6
8	200	0.3	95.8	9.5	6.9	94.9	5.7	5.4	94.2	1.8	3.8
	300		95.1	6.8	5.9	94.3	4.2	4.9	93.3	0.6	3.7
	400		94.7	5.8	5.5	93.9	3.5	4.8	91.8	0.3	4.3
9	200	0.4	94.3	5.2	5.4	93.2	3.0	4.9	92.3	1.1	4.4
	300		94.2	4.3	5.1	93.2	2.5	4.7	91.7	0.4	4.3
	400		93.4	3.5	5.0	91.8	2.2	5.2	90.0	0.3	5.1
10	200	0.7	92.9	3.1	5.1	91.5	1.6	5.0	91.0	0.7	4.9
	300		92.3	2.5	5.1	90.6	1.5	5.4	89.2	0.3	5.6
	400		92.0	2.0	5.0	90.5	1.2	5.3	88.3	0.1	5.9
11	200	0.9	90.4	1.8	5.7	89.1	1.0	5.9	88.5	0.5	6.0
	300		89.9	1.1	5.6	88.5	0.6	6.1	86.8	0.1	6.7
	400		89.9	1.2	5.6	88.1	0.7	6.3	85.4	0.0	7.3
12	200	1.0	88.4	1.0	6.3	86.9	0.6	6.8	86.3	0.2	7.0
	300		88.0	0.7	6.4	86.1	0.5	7.2	84.6	0.1	7.8
	400		87.3	0.7	6.7	84.8	0.3	7.8	82.2	0.0	8.9

Table 5. Recognition accuracies of the FFV system of FVC2002-DB2 (unit: %)

Fig. 4 shows the execution times (average, min, max) and the average number of the Lagrange interpolations for the brute-force search, the proposed algorithm and the Reed-Solomon code for FVC2002-DB1, respectively. Fig. 5 is the results of FVC2002-DB2. Although the Lagrange interpolation is an efficient technique to interpolate a polynomial, it requires more time. In the case of genuine tests, the average time of success is fast enough to be performed in real time. At the worst case, however, the brute-force search spends too much time because a huge number of the Lagrange interpolations are needed to reconstruct the true polynomial. On the other hand, the proposed algorithm and the Reed-Solomon code spend very little time even at the worst case. In our experiments, the Lagrange interpolation time for a 9-degree polynomial is 0.6 milliseconds, but in a certain case, it takes

more than 284 seconds because 425,415 interpolations are performed to reconstruct the polynomial.
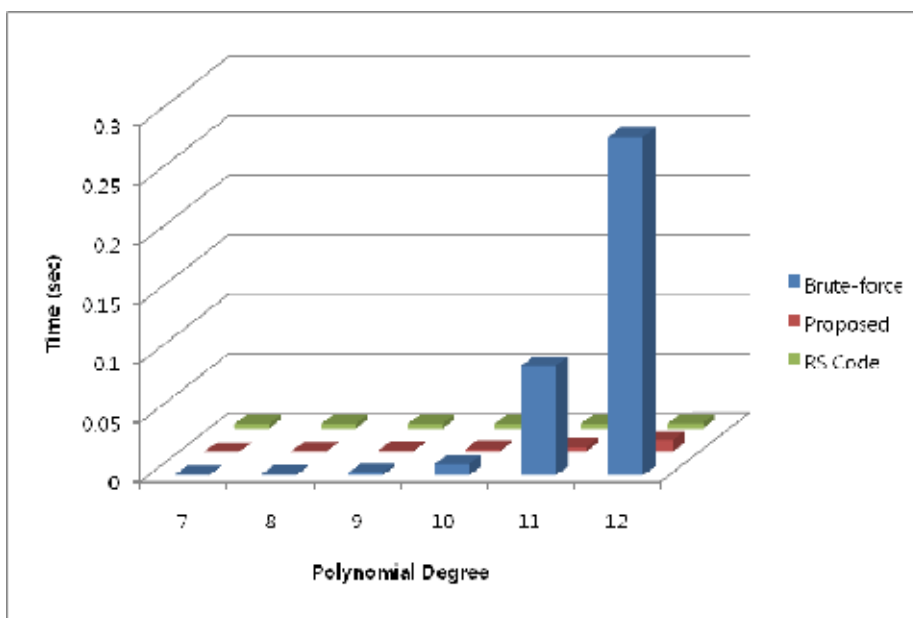

Fig. 4. Comparison of polynomial reconstruction time for the Brute-force search, the proposed algorithm, and the Reed-Solomon code (FVC2002-DB1)
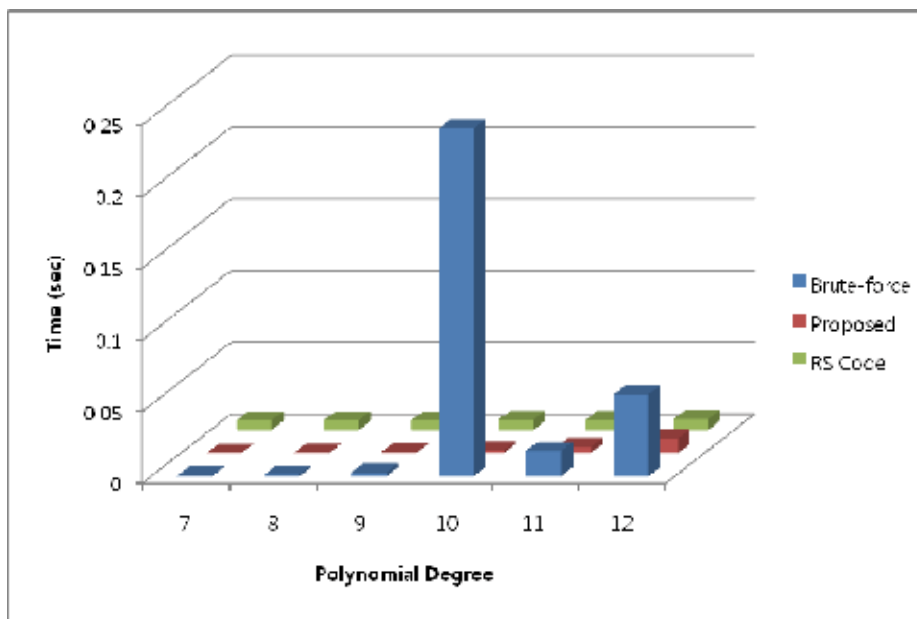

Fig. 5. Comparison of polynomial reconstruction time for the Brute-force search, the proposed algorithm, and the Reed-Solomon code (FVC2002-DB2)

4. Attack methods for fuzzy fingerprint vault

The attack of FFV is selecting real points more than the degree of polynomial. The efficient attack methods are to constitute a minutia set that contains many real points and a little chaff points. If the set contains real points more than the degree of polynomial, the polynomial can be reconstructed by the brute-force search. In addition, if the set contains more chaff points, more time is needed to reconstruct the polynomial. The correlation attack (Kholmatov & Yanikoglu, 2008) is known to be an efficient method that constitutes the minutia set using multiple vaults enrolled for different applications. On the other hand, when multiple vaults cannot be obtained and no information about the minutiae is available, the attacker should select the real minutiae from among the entire points including many chaff points.

4.1 Brute-force attack

The brute-force attack (Paar et al., 2010) is a method used to extract the real minutiae from the vault when no information is available. It tries to reconstruct the polynomial by using all the possible combinations until the correct polynomial is found. Given the $(k + 1)$ points, a k-degree polynomial is uniquely determined, which is computed by the Lagrange interpolation. Lagrange interpolating polynomial $p(x)$ of degree k that passes through $(k + 1)$ points $(x_1, y_1), \cdots, (x_{k+1}, y_{k+1})$ is given by

$$p(x) = \sum_{i=1}^{k+1} p_i(x) \tag{14}$$

where

$$p_i(x) = y_i \prod_{\substack{j=1 \\ j \neq i}}^{k+1} \frac{x - x_j}{x_i - x_j} \tag{15}$$

This is written explicitly by

$$p(x) = \frac{(x-x_2)(x-x_3)\cdots(x-x_{k+1})}{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_{k+1})} y_1 + \frac{(x-x_1)(x-x_3)\cdots(x-x_{k+1})}{(x_2-x_1)(x_2-x_3)\cdots(x_2-x_{k+1})} y_2 + \cdots$$
$$+ \frac{(x-x_1)(x-x_2)\cdots(x-x_k)}{(x_n-x_1)(x_n-x_2)\cdots(x_{k+1}-x_k)} y_{k+1} \tag{16}$$

Recall that there are n_r points in the vault, and among these points, n_e points are real, and we want to reconstruct a k-degree polynomial. Then, the total number of trials which select $(k + 1)$ points from n_r points is $C(n_r, k + 1)$, and $C(n_e, k + 1)$ combinations can reconstruct the true polynomial. If we can randomly select $(k + 1)$ points and exclude the combination from the next selection, we need to try $0.5 \times C(n_r, k + 1) / C(n_e, k + 1)$ times on the average. However, it requires too much memory to check whether the selected combination is used or not. For example, if $n_r = 230$, $n_e = 7$, and char (1 byte) type is used, then $C(230, 8) \approx 172$ Terabytes, which is impossible to be allocated in RAM. Therefore, this attack can be seen as a Bernoulli trial with probability,

$$\frac{C(n_e, k+1)}{C(n_r, k+1)} \tag{17}$$

Hence, let N_L be the number of executions of the Lagrange interpolation until the correct polynomial is reconstructed, then, N_L has the geometric distribution with mean,

$$E(N_L) = \frac{C(n_r, k+1)}{C(n_e, k+1)}$$ (18)

In general, since n_r is much greater than n_e, this attack is time-consuming. For example, if n_r = 230, n_e = 30, and k = 9, then the average number of trials of the Lagrange interpolation is C(230, 10) / C(30, 10) $\approx 3 \times 10^9$. Even though the calculation of the Lagrange interpolation takes 1 millisecond, the attack will take about 36 days on the average.

4.2 Correlation attack

Scheirer et al. suggested the methods of attacks against fuzzy vault including the attack via record multiplicity, which is known as the *correlation attack* (Scheirer & Boult, 2007). Suppose that the attacker can obtain two vaults generated from the same fingerprint (different chaff minutiae and different polynomials), the real minutiae can be revealed by correlating two vaults. If the matching minutiae contain the real minutiae more than the degree of the polynomial, the true polynomial can be successfully reconstructed by the brute-force attack, and hence, all of the real minutiae will be revealed. In addition, even if the number of the real minutiae is larger than the degree of the polynomial, it would be computationally infeasible to reconstruct the polynomial when there are too many chaff minutiae in the matching minutiae with respect to the real minutiae.
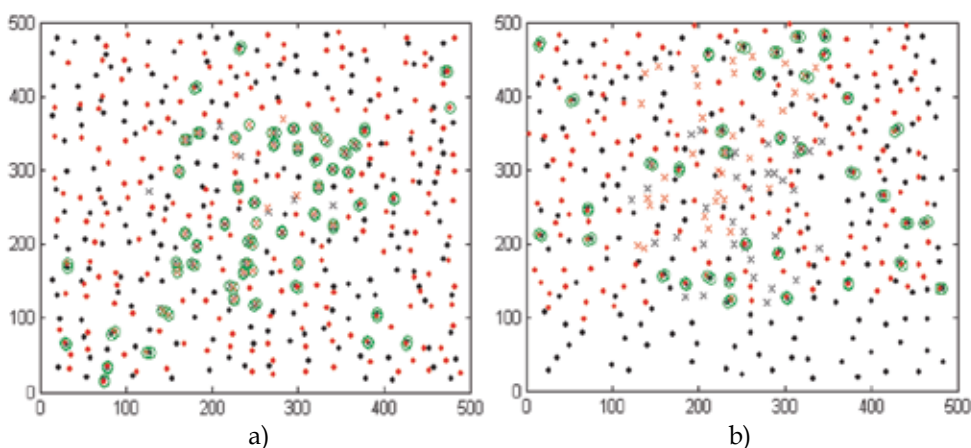
Fig. 6. Example of correlation attack. (a) shows the alignment of two vaults from the same fingerprint, and (b) from different fingerprints. (Kholmatov & Yanikoglu, 2008)

Kholmatov et al. have realized the correlation attack against a database of 400 fuzzy vaults (200 matching pairs) of their own making (Kholmatov & Yanikoglu, 2008). Fig. 6 shows an example of their experimental results. If two vaults which are generated from the same fingerprint are correlated, many real minutiae are obtained, while two vaults from different fingerprints do not have enough real minutiae in their common area to reconstruct the true polynomial. They reported that 59% of them were successfully unlocked with two matching vaults.

4.3 Fast polynomial reconstruction attack

The exhaustive search can be performed much faster than the brute-force search by using the method of the polynomial reconstruction described in Section 3.3. If a vault contains two more real points than the degree of the polynomial (i.e., $n_e \geq k + 2$), the true polynomial can be successfully reconstructed. The polynomial reconstruction time depends on the number of the real points and the number of the chaff points. The more chaff points and the less real points it contains, the more time it takes to reconstruct the polynomial. The process for the Fast Polynomial Reconstruction (FPR) attack algorithm is as follows.

First, k points are selected from n_r points with real and chaff points mixed. Second, the augmented matrix is obtained, and is reduced into the following row-echelon form.

$$
\begin{bmatrix}
1 & u_1 & u_1^2 & \cdots & u_1^{k-1} & u_1^k & | & v_1 \\
0 & 1 & u_2^{2(2)} & \cdots & u_2^{k-1(2)} & u_2^{k(2)} & | & v_2^{(2)} \\
0 & 0 & 1 & \cdots & u_3^{k-1(3)} & u_3^{k(3)} & | & v_3^{(3)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & | & \vdots \\
0 & 0 & 0 & \cdots & 1 & u_k^{k(k)} & | & v_k^{(k)} \\
0 & 0 & 0 & \cdots & 0 & 1 & | & v_{k+1}^{(k+1)} \\
0 & 0 & 0 & \cdots & 0 & 1 & | & v_{k+2}^{(k+1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & | & \vdots \\
0 & 0 & 0 & \cdots & 0 & 1 & | & v_{n_r}^{(k+1)}
\end{bmatrix}
\tag{19}
$$

Third, if there exist the same values among $v_{k+1}^{(k+1)}, \cdots, v_{n_r}^{(k+1)}$, we reconstruct the polynomial with the k points selected and one of the two points which have the same $v^{(k+1)}$ value, and then, compare it with the true polynomial.

As in the brute-force attack, we can estimate the number of the Lagrange interpolations to be performed. First, the $(k + 1)$ points are chosen, and if at least one of the remaining $(n_r - k - 1)$ points lies on the line through the $(k + 1)$ points, the Lagrange interpolation is performed regardless of whether all these points are real points or not. Let us assume that the order of the Galois field is 2^n, then the probability that at least one of the $(n_r - k - 1)$ points lies on that line is

$$
1 - \left(1 - \frac{1}{2^n}\right)^{n_r - k - 1}
\tag{20}
$$

Since the average number of trials until the $(k + 1)$ real points are drawn is $C(n_r, k + 1) / C(n_e, k + 1)$, the average number of the Lagrange interpolations to be performed until the real polynomial is reconstructed is as follows.

$$
E(N_L) = \frac{C(n_r, k+1)}{C(n_e, k+1)} \times \left\{1 - \left(1 - \frac{1}{2^n}\right)^{n_r - k - 1}\right\}
\tag{21}
$$

Since 2^n is a very large number, equation (20) is very small. Hence, the Lagrange interpolation needs to be performed in a limited number of times. In addition, let N_G be the number of executions of the Gaussian elimination of equation (19), then, the expected value of N_G is as follows.

$$E(N_G) = \frac{C(n_r, k)}{C(n_e, k)} \qquad (22)$$

It is smaller than the expected number of the Lagrange interpolation of the brute-force attack in equation (18). Also, the execution time per combination of the FPR attack is much faster than that of the brute-force attack. Therefore, this attack is more efficient than the brute-force attack. Experimental results show the efficiency of this attack method. For example, if $n = 20$, $k = 9$, $n_r = 230$, and $n_e = 30$, the average numbers of executions of Lagrange interpolation and the recursive formula are

$$E(N_L) = \frac{C(230,10)}{C(30,10)} \times \left\{ 1 - \left(1 - \frac{1}{2^{20}}\right)^{220} \right\} \approx 7 \times 10^5 \qquad (23)$$

$$E(N_G) = \frac{C(230,9)}{C(30,9)} \approx 3 \times 10^8 \qquad (24)$$

It is considerable reduction compared to the brute-force attack. The number of the Lagrange interpolation is reduced by the order of 10^4.

4.4 Experimental results

To compare the three attack algorithms mentioned in the previous section (the brute-force attack, the correlation attack, and the FPR attack), we used DB1 of FVC2002 (Maio et al., 2002), which consists of 8 impressions for each of the 100 distinct fingers. From among 8 impressions, the first impressions of each fingerprint were used for the brute-force attack and the FPR attack. For the correlation attack, on the other hand, the first and the second impressions were used to get the correlated minutiae when the correlation reached its peak. Once the matched minutiae are obtained, the polynomial reconstruction is performed by the brute-force search. We chose the number of chaff minutiae as 200, and the degree of the polynomial as 7. The tests are performed to find out whether the attack algorithms can reconstruct the true polynomial or not within 24 hours.

The results of the three attack algorithms are summarized in Table 6. The success rate is defined by the ratio of the number of successes to the total trials. The Correlation attack is known to be very efficient attack method for FFV. However, the brute-force attack and the FPR attack turn out to be much more efficient methods. Especially, the FPR attack cracked 100% of the vaults, and the attack time is only half of that of the brute-force attack.

Attack Method	Brute-force	Correlation	FPR
Success Rate (%)	95	17	100
No. Lagrange (N_L)	3.9×10^7	4.4×10^7	2.0×10^5
$E(N_L)$	6.6×10^7	2.4×10^8	9.6×10^4
No. Gaussian Elimination (N_G)	-	-	9.2×10^6
$E(N_G)$	-	-	2.1×10^7
Time (sec)	4,562	5,089	2,424

Table 6. The summary of the three attack algorithms (brute-force attack, correlation attack, and the FPR attack)

For the experiments of the correlation attack, after correlating two vaults, 23 tests extract more than 8 real minutiae, and perform polynomial reconstruction. The average correlation time is 42 seconds. Among 23 tests, 6 tests cannot reconstruct the true polynomial within 24 hours, so 17% of the vaults are cracked by the correlation attack. On the other hand, 100% of the vaults are cracked by the FPR attack within 24 hours, and the average time is 2,424 seconds, while the average time for the brute-force attack is 4,562 seconds. Since the fixed numbers of chaff minutiae are inserted, the smaller the number of real minutiae is, the more time the polynomial reconstruction requires. For the brute-force attack and the correlation attack, the attack time is proportional to the actual number of the Lagrange interpolations. Also, the time for FPR attack depends mainly on the number of the Gaussian eliminations which is computed by equation (13).

Fig. 7 shows the histogram of the actual number of Lagrange interpolation and its expected value for the case of successful attack. Although the two histograms have some fluctuations, they have similar distributions. Therefore, we can predict the attack time based on the expected number of the Lagrange interpolations when the number of chaff points is more than 200 or the degree of polynomial is greater than 7. In our experiments, the Lagrange interpolation times for the polynomial degree of 7, 8, 9, 10, 11 and 12 are 0.14, 0.28, 0.71, 1.6, 3.9 and 9.8 milliseconds, respectively. Also, the Gaussian elimination times for 200 and 400 chaff points are 0.26 and 0.52 milliseconds, respectively. The expected numbers of the Lagrange interpolations and the Gaussian eliminations can be calculated from the equations (18), (21) and (22). Table 7 shows the predicted time for the brute-force attack and the FPR attack. The security of FFV can be strengthened by adding more chaff points and by increasing the degree of polynomial. From the experimental result in the previous section, however, the more chaff points are added and the higher degree of polynomial is used, the recognition accuracy degrades significantly.
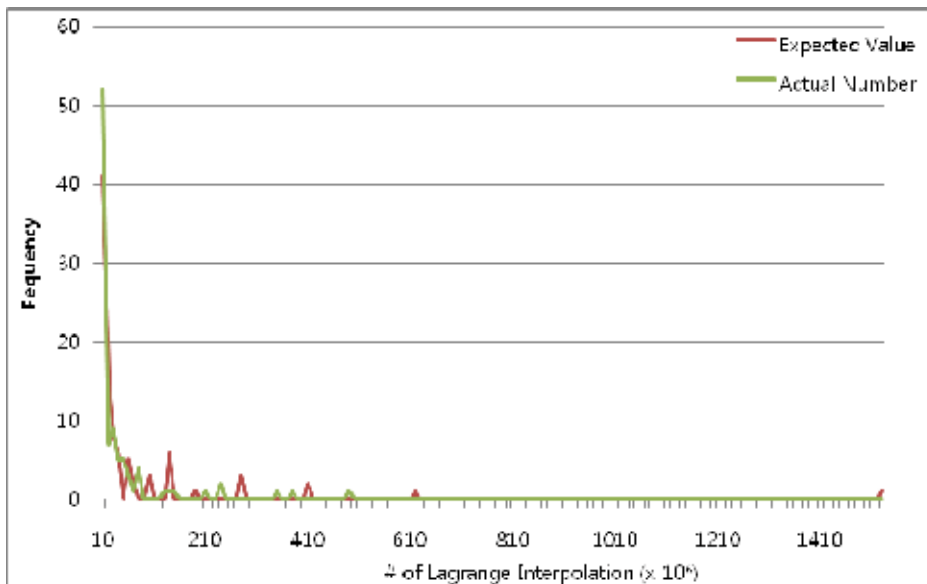

Fig. 7. Histograms of the actual number of the Lagrange interpolation and its expected value

No. Chaff	Polynomial Degree	Brute-force	FPR
	8	10 hours	1 hour
	9	9 days	9 hours
200	10	213 days	4 days
	11	14 years	35 days
	12	369 years	1 year
	7	3 days	18 hours
	8	113 days	13 days
400	9	14 years	214 days
	10	595 years	11 years
	11	26,964 years	200 years
	12	1,365,157 years	4,172 years

Table 7. Predicted time for the brute-force attack and the FPR attack

5. Conclusion

Biometrics is an efficient and convenient method for authenticating persons' identities. However, once the biometric template is compromised, the user's identity is permanently stolen. Hence, many scientists have studied to protect the biometric templates against attack. Fuzzy vault gave a promising solution to user privacy and fingerprint template security problems. Inspired from fuzzy vault, fuzzy fingerprint vault was proposed to securely store fingerprint minutiae in a database. However, there are two problems for the fuzzy fingerprint vault to be used in real world. First, by using brute-force search, the polynomial cannot be reconstructed in real time. Second, fuzzy vault is vulnerable to correlation attack.

In this work, we provided solutions to these problems. First, we proposed a fast polynomial reconstruction algorithm, which speed up the exhaustive search by using the Consistency Theorem. To reduce the execution time, it determines the candidate sets with chaff points by using Gaussian elimination and excludes them from the reconstruction trial. Since Gaussian elimination is a time-consuming process, we have found a recursive formula to perform Gaussian elimination effectively. We confirmed that the proposed algorithm can be performed in real time even at the worst case.

Second, fuzzy vault was found out to be cracked quickly by the correlation attack in 2008. The correlation attack acquires a minutiae set with many real minutiae by correlating two vaults. However, if the minutia set contains a little more chaff minutiae, the attack can hardly crack the vault. In our experiments, brute-force attack was rather more efficient. In addition, the fast polynomial reconstruction algorithm is used to crack the vault. The FPR attack algorithm records 100% attack rate. Therefore, fuzzy fingerprint vault cannot store fingerprint minutiae securely anymore. Furthermore, if we add more chaff points and use higher degree of polynomial to strengthen the security, in return, the recognition accuracy degrades significantly. Therefore, a solution for enhancing security of FFV is required. One possible solution is one-time template (Ueshige & Sakurai, 2006) whose notion is from one-time password. If we can define a one-time template for fingerprint and the corresponding transform, the security of fingerprint authentication system can be enhanced innovatively.

6. Acknowledgement

7. References

Anton, H. (1994). *Elementary Linear Algebra (7th Edition)*, John Wiley & Sons, Inc., ISBN 0-471-30569-3, New York

Bolle, R.; Connell, J. & Ratha, N. (2002). Biometrics Perils and Patches. *Pattern Recognition*, Vol. 35, No. 12, (December 2002), pp. 2727-2738, ISSN 0031-3203

Choi, W.Y.; Lee, S.; Moon, D.; Chung, Y. & Moon, K.Y. (2008). A Fast Algorithm for Polynomial Reconstruction of Fuzzy Fingerprint Vault. *IEICE Electronics Express*, Vol. 5, No. 18, (2008), pp. 725-731, ISSN 1349-2543

Choi, W.Y.; Moon, D.; Moon, K.Y. & Chung, Y. (2009). A New Alignment Algorithm of Fuzzy Fingerprint Vault Without Extra Information, *Proceedings of IASTED International Conference on Artificial Intelligence and Applications*, pp. 197-201, ISBN 978-0-88986-780-2, Innsbruck, Austria, February 2009

Chung, Y.; Moon, D.; Lee, S.; Jung, S.; Kim, T. & Ahn, D. (2006). Automatic Alignment of Fingerprint Features for Fuzzy Fingerprint Vault, *LNCS 3822: Proceedings of 1st SKLOIS Conference on Information Security and Cryptology*, pp. 358-369, ISSN 0302-9743, March 2006

Clancy, T.; Kiyavash, N. & Lin, D. (2003). Secure Smartcard-based Fingerprint Authentication, *Proceedings of ACM SIGMM Workshop on Biometrics Methods and Applications*, pp. 45-52, ISBN 1-58113-779-6, 2003

Dodis, Y.; Ostrovsky, R.; Reyzin, L. & Smith, A. (2004). Fuzzy Extractors: How To Generate Strong Keys from Biometrics and Other Noisy Data, *LNCS 3027: Proceedings of Eurocrypt*, pp. 523-540, ISSN 0302-9743, Interlaken, Switzerland, 2004

Gao, S. (2003). A New Algorithm for Decoding Reed-Solomon Codes, *Communications, Information and Network Security (V. Bhargava, H.V. Poor, V. Tarokh and S. Yoon, Edition)*, pp. 55-68, Kluwer Academic Publishers, ISBN 978-1-4020-7251-2

Gathen, J. von zur & Gerhardt, J. (2003). *Modern Computer Algebra (2nd Edition)*, Cambridge University Press, ISBN 0-521-82646-2

Hildebrand, F. (1987). *Introduction to Numerical Analysis (2nd Edition)*, Dover Publications, ISBN 0-486-65363-3, New York

Juels, A. & Sudan, M. (2002). A Fuzzy Vault Scheme, *Proceedings of IEEE International Symposium on Information Theory*, p. 408, ISBN: 0-7803-7501-7, IEEE Press, Lausanne, Switzerland, 2002

Kanak, A. & Sogukpinar, I. (2007). Fingerprint Hardening with Randomly Selected Chaff Minutiae, *Proceedings of 12th International Conference on Computer Analysis of Images and Patterns*, pp. 383-390, ISBN 978-3-540-74271-5, 2007

Kholmatov, A. & Yanikoglu, B. (2008). Realization of Correlation Attack against the Fuzzy Vault Scheme, *Proceedings of SPIE Symposium on Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, Vol. 6819, pp. 1-7, ISBN 978-0-819-46991-5, 2008

Li, Q.; Liu, Z. & Niu, X. (2006). Analysis and Problems on Fuzzy Vault Scheme, *Proceedings of 2nd International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 244-250, ISBN 0-7695-2745-0, December 2006

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

Maio, D.; Maltoni, D.; Cappelli, R.; Wayman, J. & Jain, A. (2002). FVC2002: Second Fingerprint Verification Competition, *Proceedings of 16th International Conference on Pattern Recognition*, pp. 811-814, ISSN 1051-4651, 2002

Nandakumar, K.; Jain, A. & Pankanti, S. (2007). Fingerprint-based Fuzzy Vault: Implementation and Performance. *IEEE Transactions on Information Forensics and Security*, Vol. 2, No. 4, (2007), pp. 744-757, ISSN 1556-6013

Paar, C.; Pelzl, J. & Preneel, B. (2010). *Understanding Cryptography: A Textbook for Students and Practitioners*, Springer, ISBN 978-3-642-04100-6, New York

Pan, S.B.; Moon, D.; Gil, Y.; Ahn, D. & Chung, Y. (2003). An Ultra-low Memory Fingerprint Matching Algorithm and its Implementation on a 32-bit Smart Card. *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 2, (July 2003), pp. 453-459, ISSN 0098-3063

Poh, N. & Bengio, S. (2006). Database Protocol and Tools for Evaluating Score-Level Fusion Algorithms in Biometric Authentication. *Pattern Recognition*, Vol. 39, No. 2, (2006), pp. 223-233, ISSN 0031-3203

Ratha, N.; Connell, J. & Bolle, R. (2001). Enhancing Security and Privacy in Biometrics-based Authentication Systems. *IBM Systems Journal*, Vol. 40, No. 3, (March 2001), pp. 614-634, ISSN 0018-8670

Ratha, N.; Chikkerur, S.; Connell, J. & Bolle, R. (2007). Generating Cancelable Fingerprint Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 4, (April 2007), pp. 561-572, ISSN 0162-8828

Scheirer, W. & Boult, T. (2007). Cracking Fuzzy Vaults and Biometric Encryption, *Proceedings of Biometrics Symposium*, pp. 1-6, ISBN 978-1-424-41548-9, Baltimore, MD, USA, September 2007

Shin, S.; Lee, M.; Moon, D. & Moon, K. (2009). Dictionary Attack on Functional Transform-Based Cancelable Fingerprint Templates. *ETRI Journal*, Vol. 31, No. 5, (October 2009), pp. 628-630, ISSN 1225-6463

Stallings, W. (2005). *Cryptography and Network Security: Principles and Practices (4th Edition)*, Prentice Hall, ISBN 978-0-131-87316-2

Ueshige, Y. & Sakurai, K. (2006). A Proposal of One-Time Biometric Authentication, *Proceedings of International Conference on Security and Management*, pp. 78-83, ISBN 1-60132-001-9, Las Vegas, Nevada, USA, June 2006

Uludag, U.; Pankanti, S. & Jain, A. (2005). Fuzzy Vault for Fingerprints, *LNCS 3546: Proceedings of 5th International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 310-319, ISSN 0302-9743, New York, USA, July 2005

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

6

Application of Contactless Fingerprinting

S. Mil'shtein, A. Pillai, V. Oliyil Kunnil, M. Baier and P. Bustos
Advanced Electronic Technology Center, ECE Dept., University of Massachusetts, Lowell
USA

1. Introduction

There is an established belief that biometrics, specifically fingerprinting, was invented in 1900 by Sir Francis Galton. But in fact, the Chinese played a large role in biometrics' history. About 400 years B.C. the Chinese emperor was using his fingerprint as an official signature on the imperial documents (Onin.com). There were no cases on record identifying somebody attempting to falsify this unique signature or attempted to construct a decoy. It may well be that respectful handling of the emperor's signature is not an indication of strength of biometric technology rather a proof that copying and cheating were not acceptable in the culture of early Chinese civilization.

The major development of fingerprint technology in the form of wet-ink fingerprinting was initiated and improved for forensic applications by Scotland Yard about 100 years ago. However, the development of new fingerprinting methods has happened in recent years and continues to evolve. Fingerprint recognition technology is an integral part of criminal investigations. It is the basis for the design of numerous security systems in both private and public sector. It is also seen as an important tool for a variety of government organizations including Homeland Security, Immigration, Naturalization Services, and the Armed Forces, where fingerprinting procedures are used for recognition and verification of the identity for employees of federal departments and private contractors. In addition, the growth of the internet has made it necessary to verify the identity of individuals online. The simplest form of individual verification is the use of a password; however, this does not provide high levels of security. In the U.S., where the internet is widely used, an average citizen holds eleven passwords. Individuals tend to choose passwords that are easy to remember which makes them more vulnerable to online attacks. This is exacerbated by the fact that cybercrime is increasing. It is the recognition of this inherent security flaw which amplifies the need to use biometrics in securing network communications.

After the tragic events of September 11, 2001, the need for improved and reliable fingerprint recognition technology drastically increased. We witnessed the replacement of wet ink fingerprinting by digitized contact-based methods. (S. Mil'shtein and U. Doshi, 2004) did a study which emulated the fingerprinting procedure used with computer optical scanners, it was found that on average the distance between ridges decreases about 20% when a finger is positioned on an imaging surface. Using calibrated silicon pressure sensors, the distribution of pressure across a finger was scanned pixel by pixel, and a map of average pressure distribution on a finger during fingerprint acquisition was created. This demonstrated that it is impossible to replicate the same distribution of pressure across a

finger during repeated fingerprinting procedures. Unfortunately, most fingerprints stored in current databases were acquired by contact-based methods using computer scanners and carry at least 20% distortion even if the finger is not pressed forcefully against the scanner. If a large force is applied, as in cases with non-cooperative individuals, the distortion is worse. In order to avoid pressure induced distortion we developed a line of unique contactless fingerprint scanners where rolled fingerprint equivalent images are recorded in less than one second. In (Yang and Park, 2008 as well as Nanni and Lumini, 2009), non-minutiae based methods have been proposed to overcome invariance.

Although fingerprinting is the most widely used biometric technique the common belief of law enforcement officials is that multi-modal biometrics is the future of this technology, i.e. the combination of fingerprinting with other types of biometric data such as face recognition or retina identification as mentioned by (Ross and Jain, 2004). This belief can explain the dynamics of the market's development, as illustrated in Figure 1. The global biometrics market experienced a growth from $1.95 billion in 2006 to around $2.7 billion at the start of 2008 (Techbiometric.com). Due to an annual growth rate of 21.3%, biometrics manufacturers are expected to achieve an impressive figure of $7.1 billion in revenue by 2012.
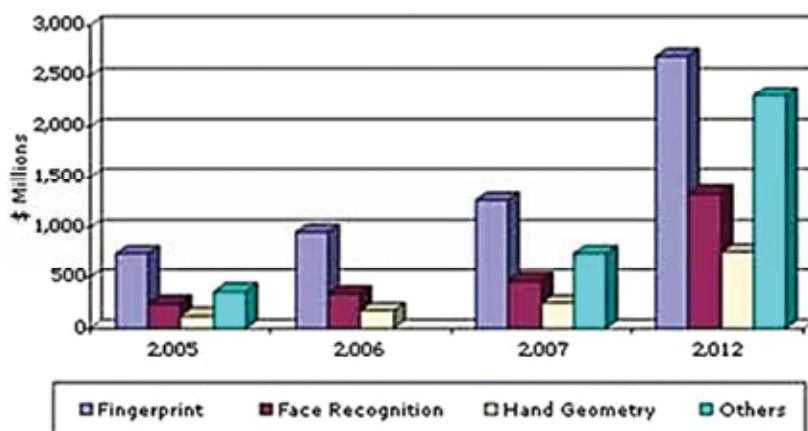

Fig. 1. Global Biometric Market Projections by Technology 2005-2012 ($ Millions)

Biometrics is starting to become a reality not only in the field of forensics, but in banking, internet services, medical services, food distribution, welfare, etc. The most integrated statewide approach to biometrics exists in Israel, where fingerprinting is mandatory and experimental passports with holographically imprinted fingerprints are expected to be issued in 2012.

Despite the known deficiencies and drawbacks of contact-based fingerprinting, this method is deployed in a variety of small mobile devices due to a relatively low cost of production. At the conference of Biometric Consortium held in Tampa, Florida on September 2010, about 17 companies demonstrated various forms of state-of-the-art contact-based hardware. Although contactless methods are known for producing distortion free fingerprints, this is a rather new technological development, and very few universities and companies are involved in their development. Figure 2 presents the major players involved in contactless fingerprinting.

Fig. 2. Contactless Fingerprinting Systems (Valencia, 2009)

Among these leaders of contactless fingerprinting only the Advanced Electronic Technology Center of UMass (AETC) has designed and tested systems where 180o contactless fingerprinting is combined with Infra-Red (IR) testing of blood vessels. Currently, the IR mapping of blood vessels is used to check whether a finger is alive and not a decoy; however, blood vessels can potentially be an additional biometric identifier. At the Advanced Electronic Technology Center, we designed and tested a novel, mobile, and contactless fingerprinting system based on line-scan technology. The system is capable of imaging the human palm, four fingers together, and nail-to-nail rolled-fingerprint equivalents, in a contactless fashion.

In the following sections, we discuss contactless fingerprinting technology, its benefits, its capabilities, and some of its applications in law enforcement, financial transactions, and network security.

2. Mobile contactless fingerprinting

In this section we describe alternative designs based on aerial and line scan cameras, compare different imaging and optical systems and the trade-offs involved with various design choices. There are two main imaging technologies used within current optical-based fingerprinting units. The first, and most widely used due to its simplicity, uses a two-dimensional array of imaging elements, much like the standard CCD found in most consumer cameras. A typical imaging setup involves a positioning system used to hold the finger steady, and a lens system used to form an image of the finger on the surface of the sensor. This technique allows for full fingerprint images to be taken in a single exposure using inexpensive camera systems. Images can be taken quickly and successively, but the

sides of the fingerprint image will carry an unavoidable distortion due to the finger's edges being round.

Edges of the finger are not perpendicular to the image sensor, which causes their projection onto the flat surface of the sensing element to be a distorted representation of the actual finger. Figure 3 displays an image of a standardized test finger made of aluminium, taken with a traditional CCD. This test finger was made by US-National Institute of Standards and Technology. The dense grid pattern in the center of the image consists of circular dots regularly spaced 0.4mm apart. It can be seen that the pattern appears to be compressed towards the top and bottom edges of the image due the fact these edges on the test finger are "falling away" from the sensor as shown in the red boxes. Depending on the systems optics, this compression can be accompanied by image blur in these areas due to depth of field limitations within the specific lens system. These limitations tend to increase as the distance between the lens and the finger decreases, making it difficult to achieve desired image resolutions as fingerprinting machinery decreases in size. If higher quality images are desired, an alternative method of optical imaging must be used.

Fig. 3. Image of a test finger taken with a traditional CCD. Distortion in the form of compression and image blur in the top and bottom portions of the image can be seen

It is possible to record a high quality, two-dimensional image of a finger using an image sensor, as opposed to the standard two-dimensional pixel array, is one pixel wide by 512, 1024, 2048, or even 4096 pixels long. Such a sensor, known as a line-scan sensor, can capture an individual line of an image per exposure. By scanning an object across the sensing element, a two-dimensional picture can be built up from the individual image slices. When extended to fingerprinting, this technique can overcome the distortion issues encountered when using two-dimensional CCDs.

A two-dimensional image of a finger is built by orienting a line-scan sensor lengthwise to the finger and then rotating it around the main axis of the finger completing an 180° arc. The image captured represents an uncoiled view of the finger equivalent to a "rolled-ink" print. The line-scanner views each portion of the finger perpendicularly, removing the projection errors inherent in conventional two-dimensional scanning techniques. If this semicircular path is concentric with the finger, the path maintains a constant object-to-lens distance throughout the scan, eliminating any possible depth of field issues. Figure 4 displays an image of the same test finger shown in Figure 3 taken using the line-scanning technique. It can be seen that the irregularities present in Figure 3 are nonexistent, and the regular spacing of the grid pattern have been preserved.

Fig. 4. Image of standardized test finger taken with the rotational line scanning approach

In such a setup, high quality fingerprint images at greater than 1000 ppi resolution have been obtained by various prototype units. This high precision enables the detection of fingerprint ridges as well as pores, facilitating the extraction of additional biometric data, such as the locations of the pores on a fingers surface.

2.1 Optical design

The raw image quality of any optical fingerprinting system is directly related to the system's optical design and configuration. Due to optical limitations which are deeply rooted in optical physics, designers of these systems are faced with design trade-offs which greatly affect the images produced by the system. For instance, a machine designed for imaging close objects will suffer from a smaller depth of field than the same machine if it were designed for imaging objects that are farther away. When viewed from a fingerprinting perspective, these optical limitations translate to practical considerations for a user or perspective buyer of fingerprinting hardware. For instance, the machine designed for up-close imaging will have less tolerance in the positions that a fingerprintee is allowed to place their finger, as deviations from this allowed position will result in a higher degree of image degradation than the more zoomed out model. The trade-off is that for the increased resolution associated with the up-close unit, the price of an increased number of re-prints, longer fingerprinting times, and higher operator training costs will most likely be paid.

2.2 Light control systems

In addition to a fingerprint reader's optical subsystem, the reader's final image quality is also dependent on the unit's lighting control system. Lighting control systems can be either *Active* or *Passive,* the distinctions, benefits, and drawbacks of each technique will be discussed in this section.

Passive lighting control is the simplest form a lighting control system can take. A passive system has the ability to *only* turn the lights on to a preset intensity value at the start of a scan, and turn the lights off at the completion of a scan. This type of system can be created with relatively cheap, easily accessible parts, and with minimal development time. This translates to a cheaper cost of fingerprinting units that employ this technology. However, because there is no feedback on whether the lighting intensity is at an optimal value, the resulting image can become over or under exposed in some operating conditions. This is because differences in skin tone, cleanliness, and moisture affect the amount of light that it will reflect back to the image sensor, having a direct effect on the exposure of the resulting image.

If over exposed, the image sensor becomes saturated, and all usable data from over exposed region is lost. In the case of under exposure, image enhancement techniques applied in post processing of the image can usually extract fingerprint ridge information, but in severe cases data can be lost.   If optimal exposures are required across all fingerprintees and all operating conditions an active lighting control system should be used.

Active lighting control is a control system that utilizes feedback in real-time to adjust the light's output intensity. Active systems must use actual images from the camera to determine the correct lighting values for each use. This feedback requires the use of image processing, and relatively sophisticated control electronics. For this reason, active control systems can be more costly than passive control systems.

Active control systems can control one or many lighting zones over an image. Systems that have more than one lighting zone have the added benefit of being able to correctly control the exposure over different parts or "zones" of a finger, where systems with only one zone can only adjust the lighting intensity of the image as a whole. Multi-zoned lighting systems can be useful when fingers that are dirty over specific regions are trying to be imaged. For instance, a finger with grease only on its tip will have a high reflectivity on the dirty area, and a relatively low reflectivity on the areas not covered in grease. In a single lighting zone imaging system, the high dynamic range of the reflected lighting values can create exposure problems in the resulting image. In a multi-zoned lighting system, the lights illuminating the greasy portion of the finger can simply be turned down.

2.3 An example of a mobile contactless fingerprint system

The Advanced Electronic Technology Center designed and fabricated a line of contactless fingerprinting systems. The design of these machines is based on the general principles described above. Figure 5 depicts a front view of one of these fingerprinting systems. The finger is positioned in the center of the optical system illuminated by blue LEDs, and then a line by line image is taken by a system of three mirrors, where one of the mirrors rotates clockwise around the finger. The nail-to-nail fingerprint image is a collection of multiple one pixel thick line-scan images and is acquired in less than one second. In anti-clockwise the map of blood vessels is taken by using IR LEDs.

Fig. 5. Front view of the contactless line scans fingerprinting system

The machine is provided with a touch screen interface, which is used to interact with an operator. The operator has the options to perform a contactless scan alone or with blood vessel imaging. Once the image acquisition is complete, it is transmitted over WLAN to a remote server which performs image processing and stores it. From this server the image may be send to law enforcement agencies for further processing.
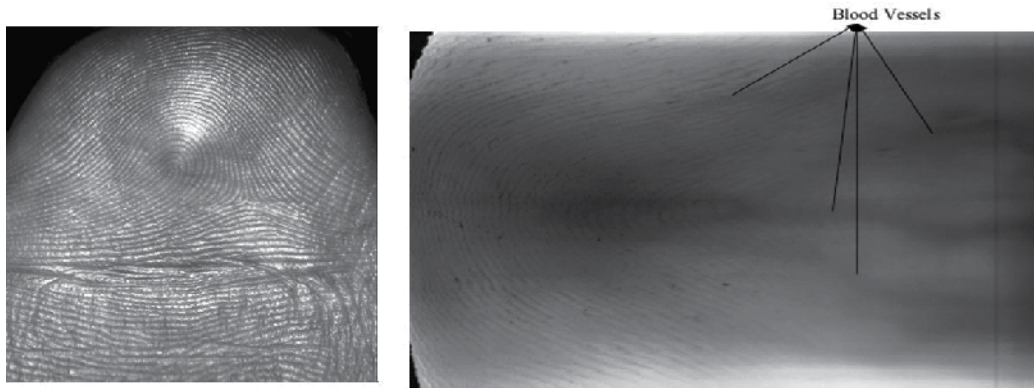

Fig. 6. a) presents an image of a fingerprint recorded in contact less fashion. b) presents an image of blood vessels taken by the contactless machine

Figure 6 has two sub-images, Figure 6a and Figure 6b. Figure 6a presents a fingerprint taken by line-scan camera where the mirror goes clockwise around the finger. The scan takes ¾ of a second. Figure 6b presents a map of blood vessels in the finger. The finger is seen in transmitted IR light where the anti-clockwise rotation of the mirror takes ¾ of a second.

3. Fingerprint processing algorithms

Different modes of fingerprint acquisition pose challenges in the form of format, size of images, non-linear distortions of fingerprint ridges, differences in orientation, and variation of gray scale values.
These challenges are mitigated by developing algorithms which pre-process raw images taken directly from acquisition devices, and facilitate reliable recognition of fingerprints.
We have developed a new binarization method that is used to eliminate variations in gray scale levels of each image, leaving the resulting images looking like a traditional wet-ink rolled fingerprint. In this study we tested 720 fingerprints generated by wet-ink, flat digital scanners, taken from FVC 2004 and by the novel contactless fingerprinting scanner described in (J. Palma et. all, 2006) and (S. Mil'shtein et. all, 2008). In following sections, we describe the binarization steps, and the fingerprint alignment process.

3.1 Binarization procedure
Most fingerprint recognition algorithms rely on the clarity and details of ridges. Hence, it is necessary to clearly differentiate the fingerprint ridges and valleys using only two distinct values; this process is called binarization. Regardless of the quality of any image recognition algorithm, a poorly binarized image can compromise its recognition statistics.

A good binarization algorithm would produce an image which would have very clear and uniform black ridges on a white background even if the image is overexposed to a certain degree. We used the following binarization techniques:
1.   Region-Based thresholding as described below.
2.   Filter-Based technique mentioned by (Meenen and Adhami, 2005).
The region based thresholding starts with division of the image into an N-by-N grid of smaller blocks. Identification of ridge regions within these smaller blocks is performed. This is implemented by taking the gradients in the x and y direction and then finding the co-variance data for the image gradients. Once this step is completed, the orientation of ridges is computed by finding the angle with respect to the coordinate axis. Then, estimation of ridge frequencies in these blocks is performed. This is done to find out which blocks have a higher and a lower density of ridges. The image block is then rotated to make the ridges vertical. and is cropped to remove invalid regions. A projection of the grayscale values, down the ridges, is obtained by summing along the columns. Peaks in projected grey values are found by performing dilation and finding where the dilation equals the original values. The spatial frequency of the ridges is determined by dividing the distance between the 1st and last peaks by the number of peaks. If no peaks are detected, or the frequency of ridge occurrence is outside the allowed bounds, the frequency is set to 0. The information about ridge regions, orientation and frequencies returns a mask of a fixed size which defines the actual area where the fingerprint exists. The ridges are then enhanced with the help of a median filter. The image obtained after this process is thresholded to obtain the binary fingerprint. The threshold for binarization depends on the resolution for the image. This process can also be called Adaptive Binarization.

This method works very well with the images that are obtained from the contactless fingerprinting system described in section 2.3. This binarization technique is not affected by varying brightness levels throughout the image, and results in a binary image that has consistent information throughout. The drawback of this process is that a relatively large number of calculations are needed, which adds to the time needed for the overall recognition algorithm to complete.
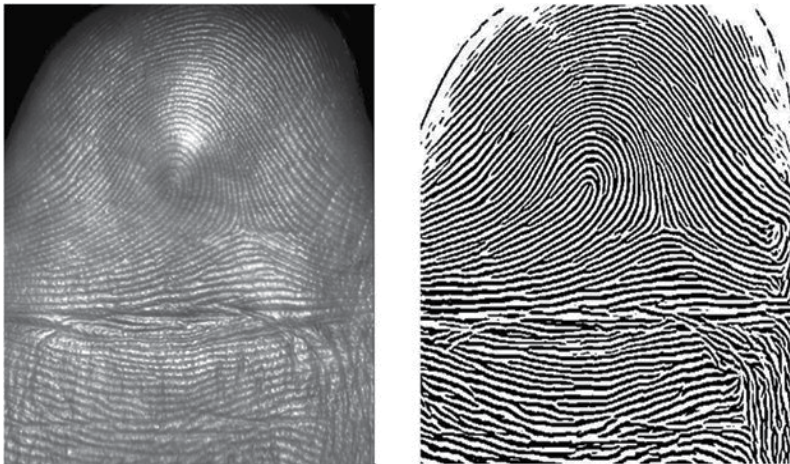

Fig. 7. Grayscale image of a selected fingerprint (Left) and the corresponding binarized image (Right)
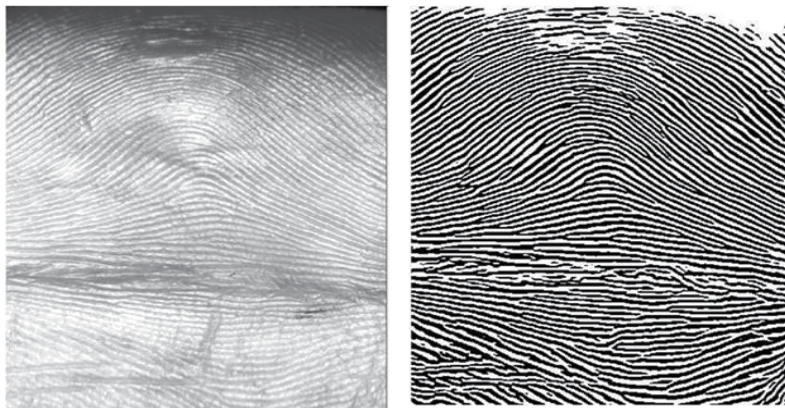
Fig. 8. Strength of binarization even if the image is seemingly overexposed

3.2 Fingerprint alignment

Fingerprint alignment is an important stage that is performed before fingerprint recognition. One must be sure that the regions being compared are the same. Fingerprint alignment using eight special types of ridges extracted from thinned fingerprint image is reported by (Hu et. all, 2008). Other alignment techniques based on phase correlation of minutiae points as described by (Chen and Gao, 2007), using line segments as pivots based on minutiae as mentioned by (Carvalho and Yehia, 2004) and using similarity histogram detailed by (Zhang et. all, 2004), have also been reported, creates a need for for a new novel alignment technique not based on minutiae. In this study, an alignment technique based on the Fourier Mellin Transform will be described.

The Fourier-Mellin Transform is a useful mathematical tool in image processing because its resulting spectrum is invariant in rotation, translation and scale. The Fourier Transform itself (FT) is translation invariant. By converting the resulting FT to log-polar coordinates, we can convert the scale and rotation differences to vertical and horizontal offsets that can be quantified. A second transform, called the Mellin Transform (MT), gives a transform-space image that is invariant to translation, rotation and scale. An application of the Fourier-Mellin Transform for image registration can be found in (Guo et. all, 2005).

The Mellin transform can be expressed as:

$$M(u,v) = \int_\infty^\infty \int_\infty^\infty f(x,y)\, x^{-ju-1} y^{-jv-1} dx dy;\; x,\, y > 0 \tag{1}$$

Convert to polar coordinates using:

$$r = \sqrt{x^2 + y^2} \tag{2}$$

We now have:

$$M\{f(r)\} = \int_\infty^\infty f(r)\, r^{-ju-1}\, dr \tag{3}$$

Making $r = e^{\dot{y}}$ and $dr = e^{\dot{y}} d\dot{y}$ we have :

$$M\{f(e^{\grave{y}})\} = \int\limits_{-\infty}^{\infty} f(e^{\grave{y}}) \, e^{-ju\grave{y}} \, d\grave{y} \tag{4}$$

By changing coordinate systems from the Cartesian system to a Log-Polar system, we can directly perform a DFT over the image to obtain the scale and rotation invariant representation. The figures below show some of the results of the alignment using The Fourier-Mellin Transform. Figures 9 and 10 are the base image and the image in need of alignment. Figure 11 shows the two images aligned using Fourier-Mellin Transform.

The inverse Fourier transform of the Mellin Transformed images helps to see how well the image is aligned with respect to the base image. While this step is necessary to see the alignment results, the Fourier transforms; however are stored in as separate database as from here they are now the templates that will be used for comparison. This will eliminate the need to take again the FFT of the aligned image and the base image when it comes to comparing the fingerprints.
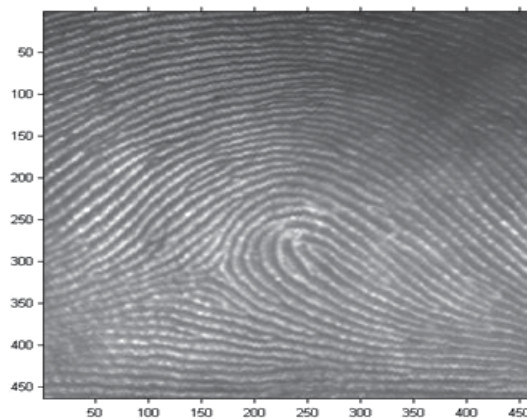

Fig. 9. Image of the 1st fingerprint

The Integrated Automated Fingerprint Identification System (IAFIS) of the FBI has fingerprints for more than 66 million people and more than 25 million civil prints (fbibiospecs.org). Most of these fingerprints have been taken with the ink and paper technique on FBI cards and then converted to a digital database using a high resolution scanner. But, it is well-known fact that most of these fingerprints are of poor quality due to smudging and smearing of the ink. In order to have improved quality of images and also improve the recognition rates, live-scan systems were used to obtain fingerprints. In these systems the image is acquired by sensing the tip of the finger using a sensor that is capable of digitizing the fingerprint on contact (Maltoni et. all, 2005). But recent studies (Mil'shtein and Doshi, 2004) have proven that fingerprints taken using the live-scan technique are also subject to pressure induced distortions. The distance between the ridges reduces depending on the pressure applied. Very recently, contactless fingerprinting techniques have been catching a lot interest due to their ability to produce almost distortion-free fingerprints. Also, because of their high resolution (Ross et. all, 2006), level 3 features can be used for identification. Figure 12 shows the fingerprint of the same finger taken using ink and paper, live-scan and touch-less technique.
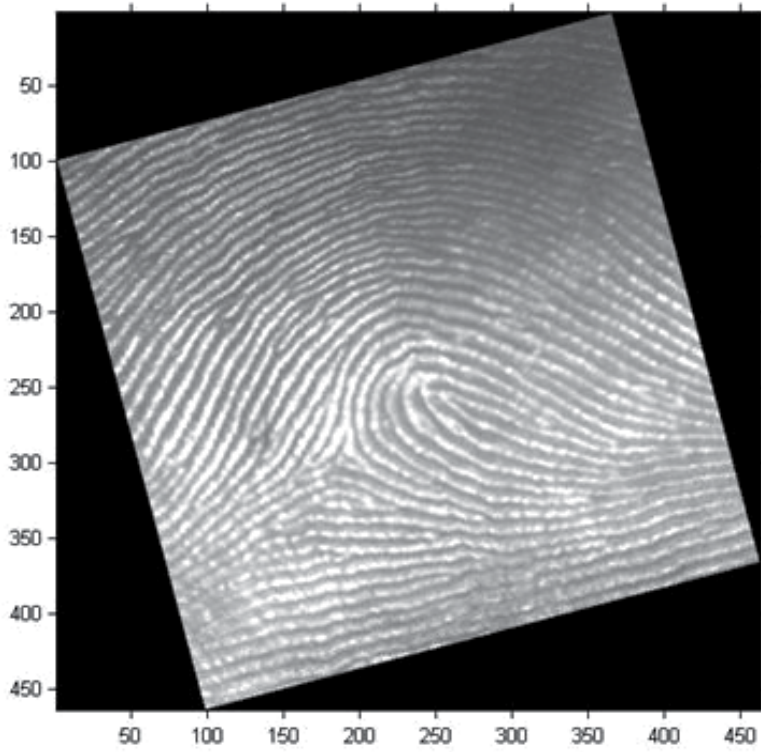
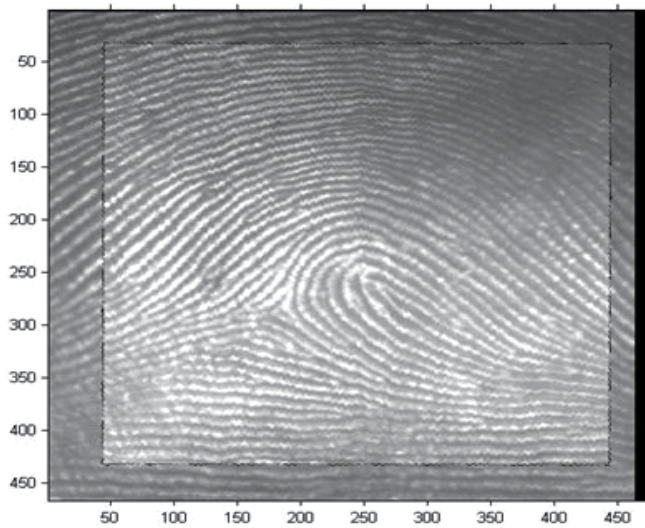Fig. 10. Image of 2nd fingerprint (In need of alignment)

Fig. 11. Aligned 1st and 2nd images (Image 2 superimposed upon image 1)

Fig. 12. Fingerprints taken via wet-ink, live scan and contactless techniques

But the issue of conversion of existing database to a contactless database is now looming and large. Another issue is how wet-ink and live-scan fingerprints compare to their touch-less counterparts. Can law enforcement agencies compare a fingerprint taken by wet-ink / live-scan methods to the same fingerprint taken using contactless technique? How will minutiae algorithm perform when it comes to comparing these two types of fingerprints? Dominant presence of wet-ink / live-scan fingerprints requires that these questions be answered immediately. In the following sections, we attempt to highlight these issues for contactless fingerprints that will come in the way of comparison of the databases.

3.3 Problems
3.3.1 Grayscale variations
Figure 13 shows an image taken from (S. Mil'shtein et. all, 2008). Even though the resolution of the fingerprint is well above FBI requirements (fbi.gov), one can clearly see that the intensity is not consistent throughout the image. As a result, digitization leads to loss of information. Hence, the resulting image is rendered useless.

3.3.2 Aspect ratio differences
Figure 12 showed comparison between a fingerprint taken by conventional wet-ink method, live-scan, and contact-less method. The aspect ratios are nearly the same in the first two images but for the contactless fingerprint, the image looks stretched out and as a result the aspect ratio varies along with the location of the minutiae points.

3.3.3 Lack of optimum illumination
Very often, the illumination circuitry in contactless fingerprinting technologies are preset to certain values. As a result, when fingers with different skin tones are fingerprinted, the fingerprint images lack the optimum brightness which result in a completely under exposed or oversaturated image.

3.3.4 Inverted background and foreground
In Figure 14 and 15, the background and foreground in a fingerprint taken via wet-ink technique and contactless technique are switched. This creates significant problems for minutiae algorithms in locating and identifying the minutiae points.

Fig. 13. Grayscale variations within the image
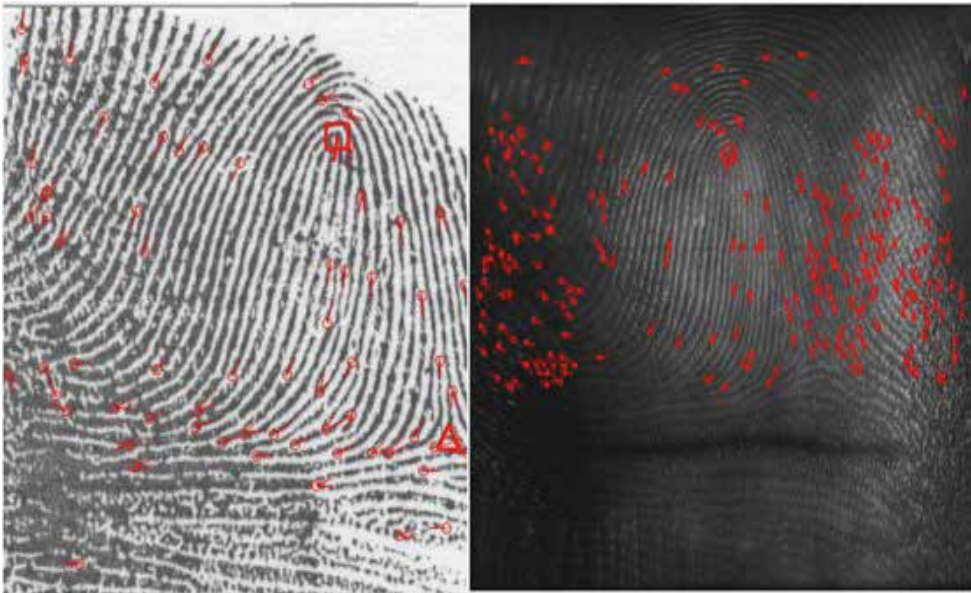
Fig. 14. Minutiae points on fingerprints taken via different techniques
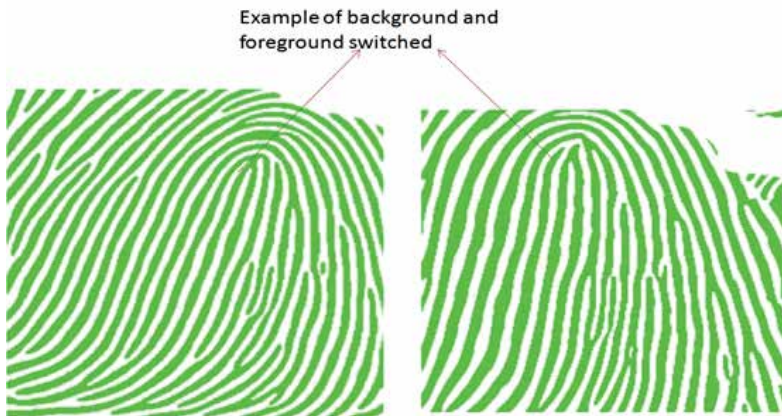
Fig. 15. Different foreground and background

3.4 Solutions
3.4.1 Adaptive binarization
One of the methods to solve the problem of grayscale variations is to binarize   the image adaptively.

3.4.2 Solution to the aspect ratio problem
Since the fingerprint images are of different aspect ratios, there needs to be sensing mechanism that senses the size of the finger and then adjusts the resolution at which the camera is taking the image. Using software packages such as MATLAB or Adobe Photoshop to resize the images to equal sizes is also possible but it does not lead to a totally accurate result, and is not scalable to high throughput operations.

3.4.3 Adaptive histogram equalization
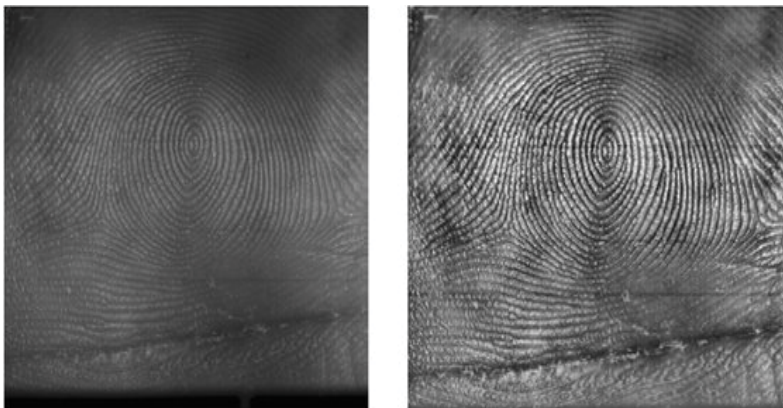
Fig. 16. Result of Histogram equalization

To make up for insufficient illumination, the contactless technology can be equipped with an adaptive histogram equalization algorithm (Pizer et. all, 1987). This ensures that fingerprint image has consistent in brightness and contrast which in turn results in good

quality binarization. It could also be implemented at a hardware level by using the procedure described in (Jusoh et. all, 2001). Fig. 16 shows the result of this algorithm

3.4.4 Inverted background and foreground

There are two solutions to the problem described in the section 3.3.4. One can take the complement of the image at the machine level and then transfer it to a database where it is binarized and then compared. Or the image can be binarized first and then complemented again using MATLAB or Adobe Photoshop. In our comparisons we found that the later method gave better results for comparisons.

4. Applications of contactless fingerprint technology

Increased security threats with respect to terrorism and cyber-crime have recently necessitated the development of biometric systems to be used at commercial facilities, border crossings, airports, and government building access points. Additionally, fraud with credit card accounts, hacking of retail store websites, and most importantly, the critical interruption of governmental agencies such as the Department of Defense and the Department of Homeland Security, requires the development of systems capable of identifying individuals accurately to mitigate such attacks.

When automated fingerprinting systems were introduced in the late 1960s, digital contact-based fingerprinting replaced the old method of ink rolling. This facilitated a new range of fingerprinting applications. The increased accuracy of contactless fingerprinting will create new applications in fingerprinting as well. Specifically, applications will be found in the fields of information security, access control, and law enforcement. Below are some examples of how contactless fingerprint systems can be used in both the private and the public sector.

4.1 Law enforcement agencies

Every organization has unique requirements for stored fingerprints, depending on how these prints are utilized. For example, agencies that deal with crime scene fingerprints prefer to have nail-to-nail images, because crime scene images may be partial; and the more information available at registration stage helps in matching them with partial prints lifted from a crime scene. This specific need of individual organizations has resulted in different agencies having their own unique and often incompatible databases.

Recently, creation of a unified and accurate database across all agencies has been recognized as a necessary step in the evolution of law enforcement's capabilities. A standardized method which captures a nail-to-nail image will help these agencies migrate towards a single large database, from which a specific portion of the image may be extracted depending on individual requirements.

To make this step, a new standard of fingerprints and fingerprinting hardware will need to be developed and followed across all agencies. Implementation details of such project will not be discussed here; however applications of such a database will be explored.

Currently, it is impossible for law enforcement to identify an individual in real-time based on fingerprints only. This is partially due to a lack of computer processing power to sort through databases containing millions of images. Most important, the current databases contain fingerprints taken using contact-based methods, and thus have varying degrees of distortion.

A database system containing high quality images and a repeatable method for fingerprint capture would facilitate applications such as real time recognition of individuals. For

example, police officers carrying mobile fingerprint capture units can successfully execute an arrest as soon as he verifies the identity of the individual.

4.2 Access control

Access control can also benefit from such devices. Current fingerprint based access control devices have a certain disadvantage in usability. Often a user may need to repeatedly scan their finger before they are granted access. The need to use a finger few times is caused by inconsistencies between the fingerprint data recorded by the capture device and the data stored within the system's database. This inconsistency increases the systems margin of error, translates to increased false rejections and a lower degree of confidence with every match.

In high security access control, an additional measure can be taken to further increase the degree of confidence with every match. Spoof detection is a technique that focuses on determining whether a finger is currently alive and attached to the body, and is in fact the person's real finger. Systems have been designed, such as the AETC's "Infrared Spoof Detection System", that satisfy such requirements (M. Baier, et. all, 2011).

4.3 Financial transaction

In the commercial sector, accurate biometric based authentication can be implemented in electronic commerce and confidential email exchange. Methods of authentication such as tokens, cards, badges, passwords and pins are widely being used today. These methods can be supplemented by accurate fingerprint based authentication to obtain a higher degree of user confidence, as well as decrease the presence of fraud in online spaces. At places of financial transactions, Automatic Teller Machines, and E-commerce are all areas that can potentially find solutions to long-standing security related problems through the use of commercialized contactless fingerprinting devices.

5. Network security

Using fingerprinting for computer and network safety is another example of applications of contactless fingerprinting. The importance of network security motivated us to present a separate section on this subject.

Security of network transactions is a major concern today, as we move towards society that is increasingly dependent on the comfort of performing day to day activities like bill payment, shopping at home, etc. Use of a public network service requires some form of authentication, because it is easily accessible to anyone connected to the network; and is prone to unauthorized and potentially malicious usage (ICC, 2009). Majority of network based authentication is performed using knowledge based methods wherein a password is used for authentication. Contrary to appearance, this type of authentication is inherently (White paper, M86 Security, 2010) flawed. A compromised password may be repeatedly used by a malicious user. There are limited means by which a second authentication system may be added in to the current infrastructure.

Few options such as RSA SecureID, VeriSign token and eToken from Aladdin knowledge systems require the user to carry additional devices which generate one-time passwords. Although these systems are marginally effective, they have the disadvantage of the user having to carry these devices. This may be inconvenient, and also if users forget to carry the device or if the device fails, they may not be able to use the system. There is also the possibility of the devices being stolen and used for authentication by malicious users. Similar problems exist in systems which send one time passwords to user's registered mobile phones.

5.1 Biometrics for authentication

Shortcomings of the above mentioned systems may be easily overcome using systems which use biometric modes of authentication, in addition to conventional static passwords. Biometrics is the use of characteristic features of face, fingerprints, iris or retina scans, voice, signature etc for authentication. The merit of biometric systems is that its uniqueness, and users do not have to carry additional hardware. However these systems have a major deficiency when it comes to usability. Face recognition, fingerprints, voice and signature are not invasive and are convenient for users. But these often pose problems with data acquisition, resulting on high false recognition. Iris and retina scans are invasive, and often pose problems for people using eye-wear such as spectacles or contact lenses. Fingerprints by far are the least invasive and most secure in terms of individuality when compared to all other biometric modalities.

5.2 Fingerprint based authentication

Fingerprint acquisition devices are usually contact based and thus pose problems during recognition stage. Fingerprints differ if they are taken by contact methods due to contact based distortion. We propose using contactless fingerprinting for network based authentication. Using contactless acquisition technologies provide high resolution, undistorted and consistent images which may be used to generate high entropy keys. Such high resolution images provide better minutiae points if minutiae (Afsar et. all. 2004) based extractor is used.

A major drawback of fingerprinting is that once a digitized fingerprint is compromised, the attacker merely needs to duplicate (Ratha et. all, 2001) it for authentication. The possibility of entire fingerprints being compromised is increased in case of transactions requiring transmission of full prints. This problem can be addressed by using partial fingerprints for identification, which is obtained from high resolution fingerprints using contactless technologies (Mil'shtein et. all, 2009).

5.3 Network security using randomized partial fingerprint

We describe a randomized partial contactless fingerprint based security protocol which uses a portion of the user's fingerprint. Effective use of partial prints is enabled by the use of distortion-less high resolution images obtained from a contactless fingerprint reader. A simplified implementation of such a system is described below:

1. Once a fingerprint is registered, it is stored in the server.
2. Upon receiving an authentication request from the authentication device, the server calculates a random Co-ordinate information key (CIK) using the image stored in it. It also stores the transaction identifier (TID) and the CIK for verification. Figure 17 describes these portions of the image. The cross-point shows the co-ordinate axis used, the circle shows the core, and the rectangle signifies the partial region used for matching.
3. Server sends back CIK along with the TID to the acquisition device.
4. The device perform acquisition and depending on the CIK, and transmits the valid portion of the image along with transaction ID back to the server.
5. The CIK is recomputed and handshaking is repeated till the server can conclusively accept or reject the user's request.

Figure 18 describes the above steps. For additional security a trusted third party maybe added in between the communication entities, as described in figure.
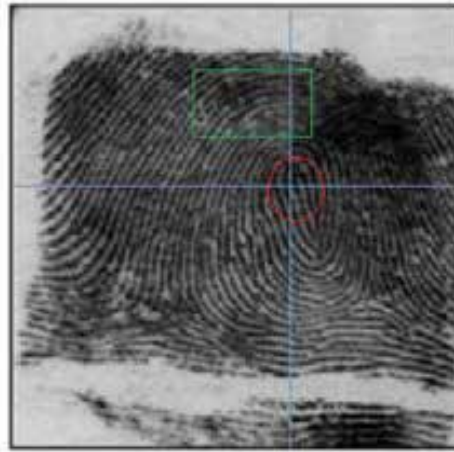
Fig. 17. A fingerprint showing CIK based region
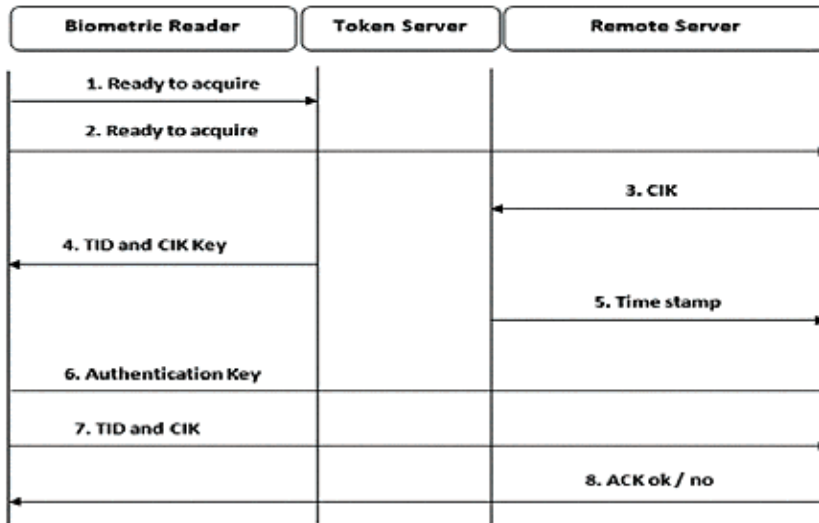
Fig. 18. A trusted third party based secure network transaction

The system described above has the following benefits:
1. The user does not have to remember multiple long passwords.
2. Once the acquisition devices are standardized, there is no need for additional hardware to be carried by the user.
3. An attacker who has access to multiple CIK will not be able to predict the next CIK, or infer which CIK will be used for a later transaction.
4. Since only a portion of the fingerprint is actually sent over the network, even if it were compromised, it would not be possible to bypass the system because of the inability to predict CIK, and thus the resulting image for any given transaction.
5. Since no complicated image processing is performed at the acquisition device, it keeps the device simple and thus economical, making it viable to be standardized.

6. Conclusions

In the last decade, both the hardware and the software of biometric technologies have been rapidly improving. The wet-ink procedures are being replaced by digitized fingerprinting, where a finger is pressed against a computer scanner. Simultaneously, novel contactless methods are being developed. Responding to the needs of forensic investigations and requirements of law enforcement new systems have been designed which made the examination of the entire palm, four -slap fingers, and nail to nail fingerprinting possible. Although, based on the old minutia algorithms current recognition software was modified to replace an operator by a computer to analyze and compare fingerprints. Fast network communication between police stations and database centers became a reality of the everyday operation of law enforcement. The Advanced Electronic Technology Center of UMass contributed to the recent modification of fingerprinting technology by combining contactless fingerprinting with blood vessel mapping in a line of newly designed hand-held systems which allow for examining of a palm and four fingers simultaneously, as well as each individual finger from nail-to-nail.

The best way to understand the future development of fingerprinting technology is to analyze the deficiencies of existing fingerprinting methods. In brief, these deficiencies could be classified into two groups.

1.   General deficiencies related to biological conditions of human body are:
a.   The shape of fingerprint changes with age due to the appearance of wrinkles on human skin. Periodic recertification of individual fingerprints is a potential answer to this problem
b.   Medical conditions of the individual might modify the reflectivity of the skin and change the IR light absorption by blood vessels. Adjustable light intensity and contrast in a fingerprinting system is one of the potential solutions of this problem.
2.   The deficiencies related to technical limitations of fingerprinting systems are:
a.   Low accuracy of recognition in some existing systems necessitates use of a second method of recognition of an individual. Often, computer scanners after taking fingerprint images need to scan an employee badge or a picture to confirm the identity of an individual. Non-distorted images generated by contactless fingerprinting systems and tight recognition algorithms are the answer to the problem.
b.   Fingerprinting based on a physical contact of a finger with a scanner generates pressure induced distortions of the fingerprint. Network security requires non-distorted, thus contactless images, for the computer (not an operator) to verify the identity of an individual.
c.   There are three libraries of fingerprints produced by wet-ink technology, digitized techniques and contactless methods. A compatibility study of these three libraries is urgently needed.

7. References

A. Ross, A. K. Jain, Multimodal Biometrics: An Overview, *Proc. of 12th European Signal Processing Conference*, pp:1221-1224, 2004.

C. Carvalho, H. Yehia, Fingerprint Alignment using Line Segments, *Biometric Authentication, Springer*, pp 1-10, 2004

C. Hu, J. Yin, E. Zhu, H. Chen, Y. Li, Fingerprint Alignment using Special Ridges, *Springer,* 2008

Davide Maltoni, Dario Maio, Anil K. Jain, Salil Prabhakar, Handbook of fingerprinting, *Springer* 2005

F.A. Afsar, M. Arif and M. Hussain, Fingerprint Identification and Verification System using Minutiae Matching, *National Conference on Emerging Technologies*, 2004

L. Nanni and A. Lumini, Descriptors for image-based fingerprint matchers, *Expert Systems With Applications*, vol.36, no.10, pp.12414-12422, December 2009

M. Baier, S. Foret, V. Oliyil Kunnil, M. Paradis, P. Bustos, S.Mil'shtein, Automatic Contactless Fingerprinting System, *To appear:Proc. of PECCS 2011*, Portugal.

Noor Ashedah Binti Jusoh et. all, Adaptive lightning system and method for machine vision apparatus, *patent # US 6,207,946 B1, 2001.*

N.K. Ratha, J.H. Connel, and R.M. Bolle, Enhancing security and privacy in biometrics-based authentication system, *IBM Systems Journal*, 2001

P. Meenen, R. Adhami, Approaches to image binarization in current automated fingerprint identification systems, *Proceedings of the Thirty-Seventh Southeastern Symposium on System Theory*, 2005, ISBN:0-7803-8808-9

Palma J, Liessner C, Mil'shtein S, Contactless Optical Scanning of Fingerprints with 180° View, *Scanning*, 28, 6, pp 301-304, 2006

Ross, Arun A., Nandakumar, Karthik, Jain, Anil K., Handbook of Multi-biometrics, *Springer,* 2006

S. Mil'shtein, M. Baier, C. Granz and P.Bustos, Mobile System for Fingerprinting and Mapping of Blood -Vessels across a Finger, *IEEE Intern. Conf. on Technologies for Homeland Security*, ISBN 978-1-4244-4179-2, p.30-34, 2009

S. Mil'shtein, J. Palma, C. Liessner, M. Baier, A. Pillai, and A.Shendye, Line Scanner for Biometric Applications, *IEEE Intern. Conf. on Technologies for Homeland Security*, ISBN 4244-1978-4 P 205-208, 2008.

S. Mil'shtein, U. Doshi, Scanning of the Pressure-Induced Distortion of Fingerprints, *Scanning,* 26, 4, 323-327, 2004.

S. M. Pizer, E. P. Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, B. T. H. Romeny, John B. Zimmerman and Karel Zuiderveld, Adaptive histogram equalization and its variations, *Computer Vision, Graphics, and Image Processing,* Volume 39, Issue 3, September 1987, pp 355-368.

T. Zhang, J. Tian, Y. He, J. Cheng, X. Yang, Fingerprint alignment using similarity histogram, *International conference on audio and video-based biometric person authentication*, pp 854-861, 2003

V. Valencia, Intuitive Touchless Fingerprinting The Pressure's Off, *Biometric Consortium Conference*, Tampa Bay, Florida, 2009.

W. Chen, Y. Gao, Minutiae-based Fingerprint Alignment Using Phase Correlation, *Mechatronics and Machine Vision in Practice*, Springer Link, pp 193-198, 2007

Xiaoxin Guo, Zhiwen Xu, Yinan Lu, Yunjie Pang, An Application of Fourier-Mellin Transform in Image Registration, *The Fifth International Conference on Computer and Information Technology*, Sep 2005, pp 619-623.

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

http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/iafis/iafis

https://www.fbibiospecs.org/biometric_specs.html

http://techbiometric.com/biometric-market/global-market-overview/

http://www.onin.com/fp/fphistory.html

Internet Crime complaint Center, statistics for 2009
        http://www.ic3.gov/media/annualreport/2009_IC3Report.pdf

Whitepaper M86 Security - August 2010, Cybercriminals Target Online Banking Customers - Use Trojan and Exploit Kits to Steal Funds from Major UK Financial Institution.

Cancelable Biometric Identification by Combining Biological Data with Artifacts

Nobuyuki Nishiuchi and Hiroka Soya
Tokyo Metropolitan University
Japan

1. Introduction

In present day information-oriented society, the ability to accurately and rapidly identify an individual in various situations, such as identity verification at an ATM, login authentication, and permitting access to secured rooms, has taken on considerable importance. Personal identification systems that rely on knowledge, for example, a password and ID number, or possession, for example, an ID card or keys, are subject to loss, counterfeiting, and theft. In addition, such systems suffer from the inability to identify the genuine user if the information is borrowed on permission of the user. Due to these limitations, the development of an identification system based on biometrics has attracted a great deal of interest as it obviates the requirement for physical possession or memorization of a security code and has the potential to differentiate individuals with high accuracy (Ashbourn, 2000; Prabhakar et al., 2003; Jain et al., 2004a, 2004b). To date, fingerprints, veins, iris, retina patterns, facial and other features have been used for biometric identification. The ideal biological data for biometrics has the following five characteristics (Wayman, 2000):

i.   Distinctive: the biological data differs from one person to another.
ii. Repeatable: the biological data remains constant over a long period.
iii. Accessible: it is easy to view the biological data.
iv. Acceptable: it is not objectionable to show the biological data.
v.   Universal: all people possess the biological data.

From different viewpoints, the five characteristics are associated with the potential problems and limitations of biometric identification.

1.1 Problems of biometric identification

Current biometric identification systems have a number of problems that are related with the five characteristics of biological data described in the above section. The three main problems are as follows:

Problem 1: The biological data cannot be replaced.

> For instance, if a user's fingers are lost, or if fingerprint information is stolen, the user cannot use a fingerprint identification system. This problem is related with characteristics (ii) and (v).

Problem 2: Users are specified only from the biological data.

As biological data is information linked directly with individuals, if biological data is leaked, the user can be specified using only the leaked biological data. This problem is related with characteristic (i).

Problem 3: The biological data can be collected without consent of the user.

In general, because biological features are exposed on the surface of the body, such as the face, fingerprints, and iris, it is difficult to keep these features concealed from others. This problem is related with characteristics (iii) and (iv).

Due to these problems, current biometric identification systems have a major vulnerability: spoofing. Yamada et al. (2000), Stén et al. (2003), Hirabayashi et al. (2004), and Matsumoto (2006) described this vulnerability of biometric identification and demonstrated that it is possible with existing technology to obtain fingerprint information from adhered surface residue and replicate the fingerprint on an artificial finger. The theft and counterfeit of exposed biological information can be accomplished by first capturing an individual's targeted information as a two-dimensional image, and then using the data to reproduce a counterfeit model.

As a result of this vulnerability to spoofing, and despite progress with various types of biometric systems, users are often hesitant to submit their unique biological data during the initial enrollment process (Gunn, 2010). It is easy to envision that users of restricted facilities, such as buildings, commercial establishments, accommodations, and amusement parks, may not willingly submit the necessary biological information for a biometric identification system.

To overcome these limitations, novel approaches for the development of practical biometric identification systems that do not retain or require potentially sensitive user information are needed.

1.2 Proposed method of cancelable biometric identification

In this chapter, we introduce a novel method of cancelable biometric identification that combines biological data with the use of artifacts and is resistant to spoofing. In this identification system, the user first attaches an artifact (a sticker with two dots) to the fingernail of the thumb or forefinger during the enrollment step, and subsequently presents the finger (biological data) with the attached artifact to the system for imaging. The position and direction of the artifact are uniquely detected based on the individual's biological data (outline of finger) using image processing. In the identification step, the user presents the finger with the attached artifact, and identification is accomplished by comparison of the probe and reference data. As the randomness of the position and direction of the artifact on the fingernail is quite high, the user can be uniquely identified. Notably, this system represents cancelable biometric identification, because once the artifact is removed from the fingernail, re-enrollment is required. From the viewpoint of ease of use, our proposed method is more acceptable than other identification methods using artifacts, such as RFID implants (Rotter et al., 2008).

This chapter is organized as follows. In Sections 2 and 3, the details of the proposed method of cancelable biometric identification are described. In Sections 4 and 5, the results of experiments and simulations using this method are presented and discussed. In Section 6, the features of the proposed method are summarized and applications of the method are proposed. Finally, conclusions and future directions are offered in Section 7.

2. Experimental setup

The artifact and hardware prototypes used in the experimental biometric identification system are shown in Figures 1 and 2. At the actual application stage, the artifact will be designed to have a less intrusive appearance and provide a higher level of security, and the imaging hardware will be smaller and more compact.

2.1 Artifact

We evaluated two artifact prototypes, having either a circular- or square-shaped design (Figure 1). For both types, a white base sticker was marked with one red and one blue dot. The circular artifact was 6 mm in diameter, and the square artifact was 5 × 1.5 mm in size. In the enrollment step, the user first attaches the artifact to the fingernail of the thumb or forefinger, and image processing is used to extract the dots on the artifact. As our initial evaluation did not detect any differences between the two types of artifacts during the data extraction, we selected the square type for use in further experiments because of its ease of fabrication. For practical use, the base sticker should be circular, transparent, and have dots printed with dye that can only be detected only under specific lighting, to maximize the difficulty of re-attaching the artifact at the same position and angle. To facilitate user acceptance of the system, ideally, the fingernail would not appear different from its usual appearance, and the attached artifact would not interfere with daily life. Moreover, it may also be possible to mark the fingernail directly with dye to serve as the artifact.
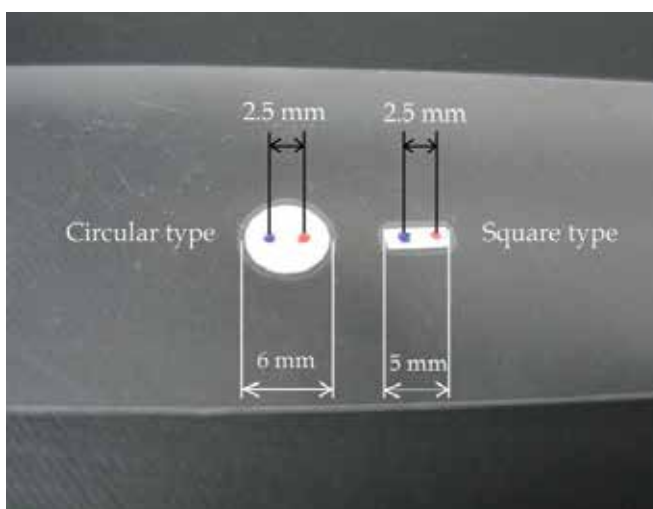
Fig. 1. Artifact used in the present system (left: circular type, right: square type)

2.2 Experimental device configuration

The device configuration of the experimental biometric identification system is illustrated in Figure 2. After placing the thumb or forefinger (in this experiment, the thumb was used) with the attached artifact on the stage, an image of the user's thumb was obtained with a single CCD camera (XC-505; Sony, Japan) under illumination by a LED light (NSF-150W; Shimatec, Japan). A black cloth was used as a backdrop to facilitate image processing. All images were analysed using Visual C++ 2008 software (Microsoft) and a 640 × 480 pixel

capture board (FDM-PCI4; Photron Ltd., Japan). A representative input image obtained using this system is shown in Figure 3.
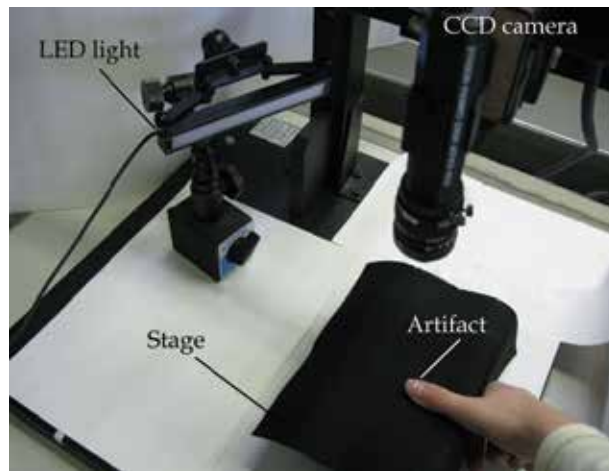
Fig. 2. Configuration of the experimental biometric identification system during image capture
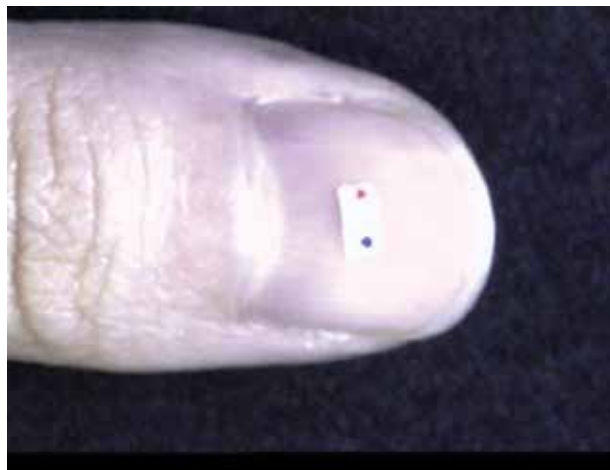
Fig. 3. A representative input image showing the artifact on a fingernail

3. Algorithm for cancelable biometric identification

The algorithm flow of the proposed cancelable biometric identification system is outlined in Figure 4. The enrollment step proceeds until feature extraction (edge pursuit and distance calculation) is performed, and the obtained reference data is then stored in the database. The algorithm flow of the identification step can be divided into four parts and begins with processing of the first input image for the artifact (binarization and center extraction; Figure 4). The second step involves image processing for the finger (binarization and edge extraction), while the third and fourth steps consist of feature extraction (edge pursuit and

distance calculation) and comparison, respectively. The details of the algorithm are described in the following subsections.
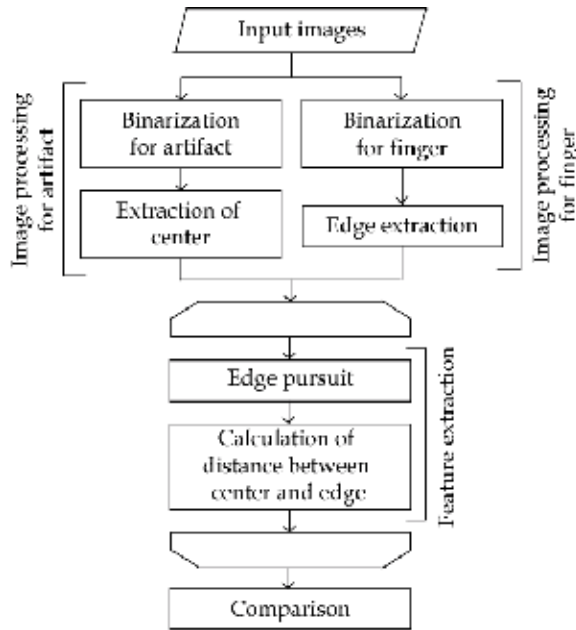
Fig. 4. Flow chart of the algorithm used for the identification step

3.1 Image processing for artifacts

In this step, the center of each dot on the artifact in the input image is determined. First, the input image is binarized by the color of each dot (blue and red), and the area of each dot is extracted (Figure 5(b)). To determine the center of the blue area, horizontal maximum X_{bmax} and minimum X_{bmin} and vertical maximum Y_{bmax} and minimum Y_{bmin} are searched by horizontal and vertical scanning, respectively. The intersection point of line segment $X_{bmax}X_{bmin}$ and $Y_{bmax}Y_{bmin}$ is determined to represent the center of blue point area (B_c). Using the identical process, the center of the red area (R_c) is also detected (Figure 5(c)).

(a)   (b)   (c)

Fig. 5. Image processing for the artifact, showing a (a) representative input image, (b) extraction of the two colored dots, and (c) detection of the center of each dot area (zoomed image)

In this study, we used colored dots on the artifact and the above algorithm to detect the center of each colored dot. However, as only the position of two points (or a vector) is needed, it is possible to introduce variations to the shape and color of the artifact.

3.2 Image processing for fingers

In the next step of image processing, the finger outline is determined from the input image. The input image is first processed by binarization to separate the background and finger area into a binary image (Figure 6(b)). The finger outline is then obtained by edge extraction using a Laplacian filter (Figure 6(c)).

(a)   (b)   (c)

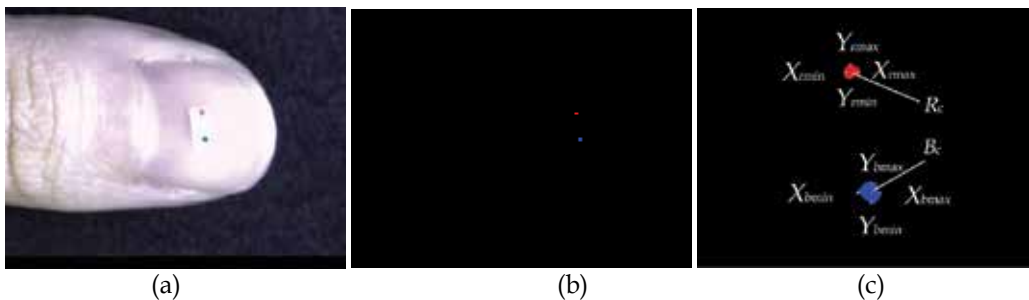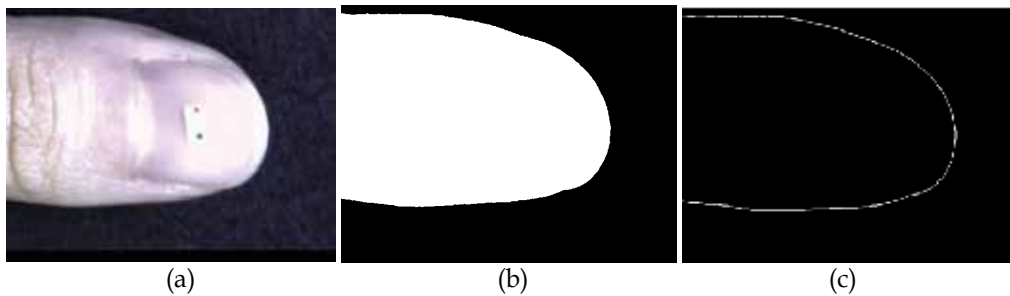Fig. 6. Image processing for the finger, showing a (a) representative input image, (b) extraction of the finger area (binary image), and (c) extraction of the finger outline (edge image)

3.3 Feature extraction

As a preprocessing step for feature extraction to equalize the volume of finger outline data, the finger outline is excised at a set distance (450 pixels) from the edge of the fingertip (indicated by a vertical line in Figure 7(a)), and the edges opposite the fingertip (towards the first finger joint) are connected with a line (Figure 7(a)). The fingertip location is decided based on the horizontal maximum point of the finger outline by horizontal scanning.

(a)   (b)
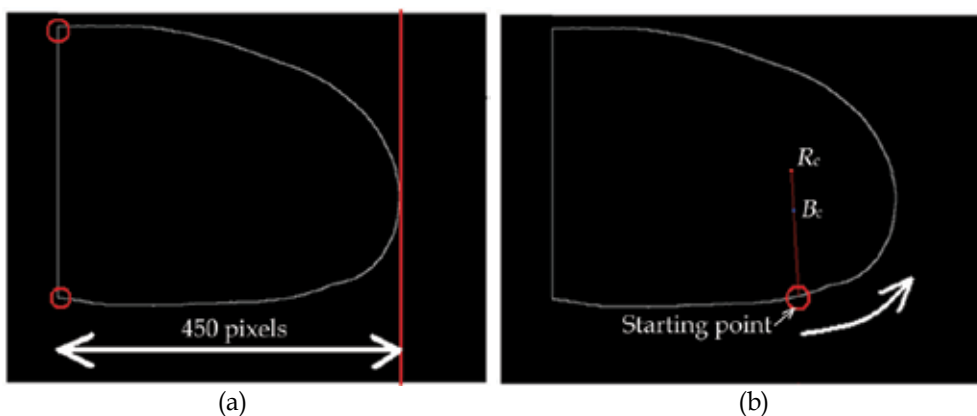
Fig. 7. Feature extraction. (a) Extracted outline of the finger and connection of the edges (red circles indicate the edges), (b) Detection of the starting point for pursuing the finger outline based on the points on the artifact

For the feature extraction processing, the finger outline pixels are pursued in an anti-clockwise direction from the starting point until returning to that point (Nishiuchi, 2010), and the distance between pixels on the finger outline and the middle of the two dots (between B_c and R_c) on the artifact is measured continuously. The starting point for pursuing the finger outline is detected based on the intersection between the finger outline and the extended line connecting points R_c and B_c on the artifact (Figure 7(b)).

A representative graph based on the feature extraction processing procedure is presented in Figure 8, where the horizontal and vertical axes represent the position of the pursued pixels and the measured distance, respectively. The data shown in Figure 8 is used for the reference and probe data during identification to determine whether a presented finger and artifact are genuine or an imposter. The red area in Figure 8 corresponds to the line connecting the two edges (red circles in Figure 7(a)) of the outline of the finger. The data within this area is not used during the comparison step.
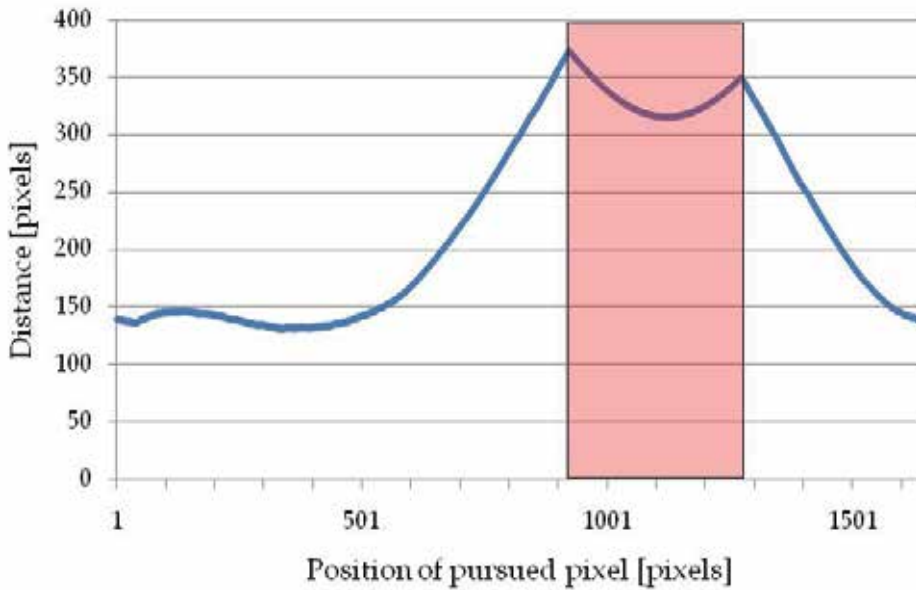
Fig. 8. Distance between the pixels on the outline of finger and the middle of the two dots on the artifact

3.4 Comparison

In the final comparison step, the correlation coefficient (R) is used for the comparison between the reference and probe data. Correlation coefficient R is calculated using Equation (1):

$$R = \frac{\sum_{i=1}^{n}(x_i - x_{aa})(y_i - y_{aa})}{\sqrt{\sum_{i=1}^{n}(x_i - x_{aa})^2}\sqrt{\sum_{i=1}^{n}(y_i - y_{aa})^2}} \tag{1}$$

In Equation (1), x_i (i=1, 2, 3, …,n) represents reference data, y_i (i=1, 2, 3, …, n) represents probe data, and x_{aa} and y_{aa} represent the arithmetic average of x_i and y_i, respectively.

4. Experimental evaluation

To evaluate the proposed biometric identification method, the following three experiments were conducted.
EXP. 1 Genuine trial: validation of repeatability
EXP. 2 Imposter trial: validation of anti-spoofing
EXP. 3 Genuine trail- artifact is removed and re-attached: validation of anti-spoofing
The set of EXP. 1 and EXP. 2 was performed as a general evaluation of the proposed biometric identification method to allow comparison with previous biometric systems, and EXP. 3 was conducted as a validation of the anti-spoofing property of our system using a genuine user who had removed and re-attached the artifact. The details and outcomes of each experiment are described in the following subsections.

4.1 Genuine trial: validation of repeatability

To validate the repeatability of the proposed biometric identification method, five images of a finger with an attached artifact were each captured for five subjects (A-E) with the finger resting on the stage of the imaging system. A representative set of captured images for subject A is shown in Figure 9. The reference data (Data A1) was then compared with the probe data (Data A2 to Data A4) of the genuine subject.

The result of the comparison for subject A is shown in Figure 10, where the horizontal and vertical axes represent the position of the pursued pixels and the measured distance, respectively, and the five lines represent each feature extracted from the five images of the genuine finger with the attached artifact. On comparison of the plotted reference and probe data, it is clear that they are quite similar. This determination was confirmed by examining the correlation coefficients resulting from the comparison for all five subjects (Table 1). As the minimum values of the correlation coefficients are all 0.996 or greater, the repeatability for the identification was considered to be high. In addition, the repeatability could be increased by using a guide for fixing the finger in place during the capture of the input image (data not shown).

(a) A1   (b) A2   (c) A3

(d) A4   (e) A5
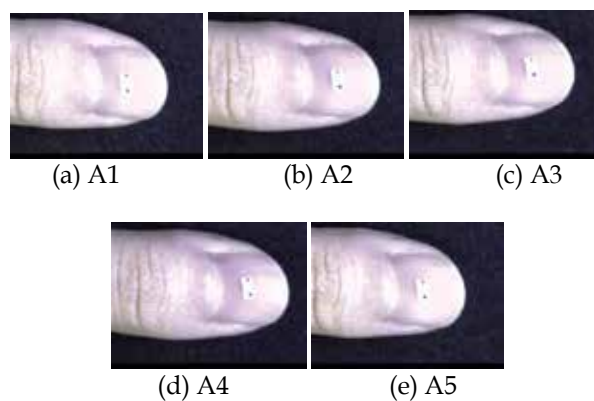
Fig. 9. Five images of the identical genuine finger of subject A with an attached artifact
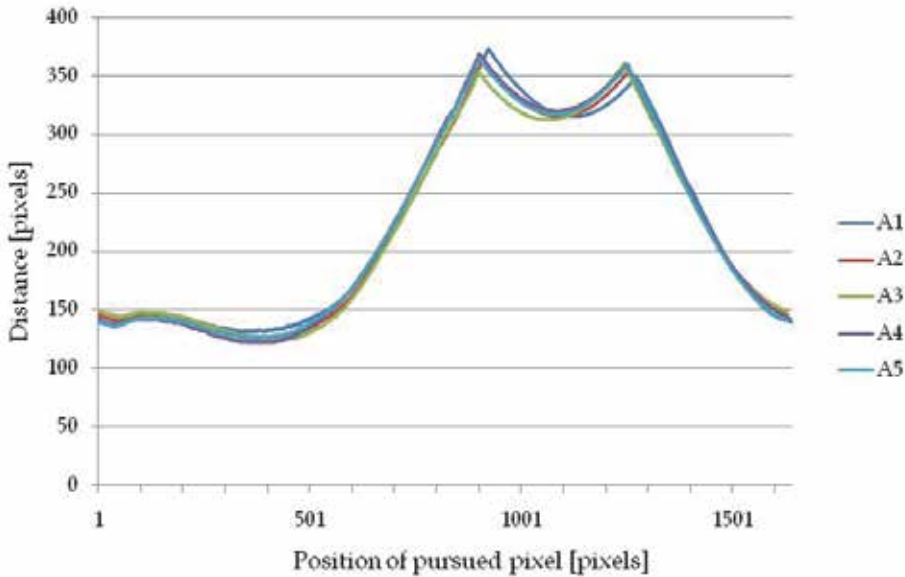
Fig. 10. Distance between the pursued pixels on the outline of finger and the artifact for the genuine trial of subject A

Subject	A	B	C	D	E
Average	0.9972	0.9976	0.9989	0.9993	0.9996
Maximum	0.9984	0.9986	0.9998	0.9996	0.9998
Minimum	0.9962	0.9964	0.9971	0.9991	0.9995

Table 1. Correlation coefficients for the comparison of the reference and probe data obtained during the genuine trial for subject A-E

4.2 Imposter trail: validation of anti-spoofing

After validating the repeatability of the proposed method, its resistance to spoofing was next evaluated by capturing five images of fingers with an attached artifact from five subjects (A-E) (Figure 11). The reference data (Data A) was then compared with the probe data (Data B to E) of the four imposter subjects. As an added element to evaluate the spoofing resistance, the imposter subjects (B-E) attempted to mimic the position and angle of the artifact of the genuine user (A) by referring to an image of the genuine user's finger with the attached artifact.

The result of the comparison between the data of the imposters and genuine user is shown in Figure 12, where the horizontal and vertical axes represent the position of the pursued pixels and the measured distance, respectively, and the five lines represent the features of each subject (A-E). It can be seen that lines of subjects D and E are quite similar to subject A. Table 2 lists the correlation coefficients resulting from the comparison between genuine user A with each of the imposters. As can be seen in Table 2, the correlation coefficients of A-D

and A-E tended to be high. However, for the genuine trial, the correlation coefficient values were 0.996 or higher, whereas the imposter trial resulted in values ranging from 0.680 to 0.983. In addition, the distributions from the genuine and imposter trials did not interfere. When the threshold value for identification was set at 0.995, both the false rejection rate (FRR) and false acceptance rate (FAR) were 0%. Even though the imposter subjects attempted to mimic the artifact position of the genuine user, it was difficult to set the artifact at the identical position and angle as that of the genuine user, demonstrating the resistance of our proposed system to spoofing.

(a) A   (b) B   (c) C

(d) D   (e) E
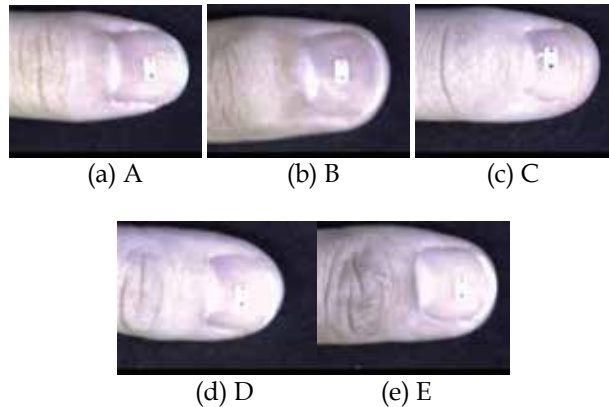
Fig. 11. Images of fingers with attached artifacts for five subjects (A, genuine user; B-E, imposter subjects)
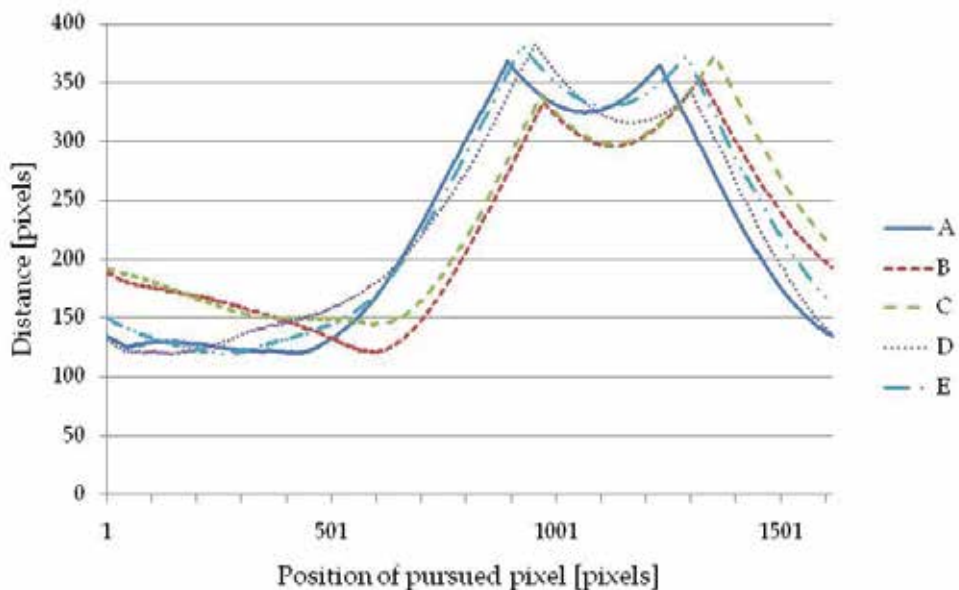
Fig. 12. Distance between the pixels on the outline of the finger and the artifact for the five subjects of the imposter trial

A-B	A-C	A-D	A-E
0.6844	0.7313	0.9705	0.9826

Table 2. Correlation coefficients for the comparison of the reference (A) and probe data (B-E) obtained for the imposter trial

4.3 Genuine trial: artifact is removed and re-attached

In this experiment, we validated the ability of the proposed biometric identification system to reject a genuine user who had removed and re-attached the artifact. Two captured images of the identical finger of subject A with an attached artifact that was removed once and attached again in a random position are shown in Figure 13. The reference data (Data A) was then compared the probe data (Data A′; re-attached artifact) of subject A.

(a) A   (b) A′
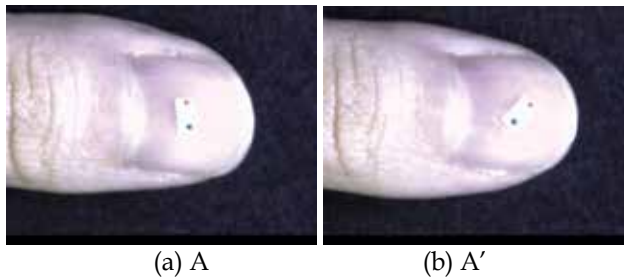
Fig. 13. Images of a genuine finger with an attached artifact (left) that was removed once and re-attached in a random position (right)
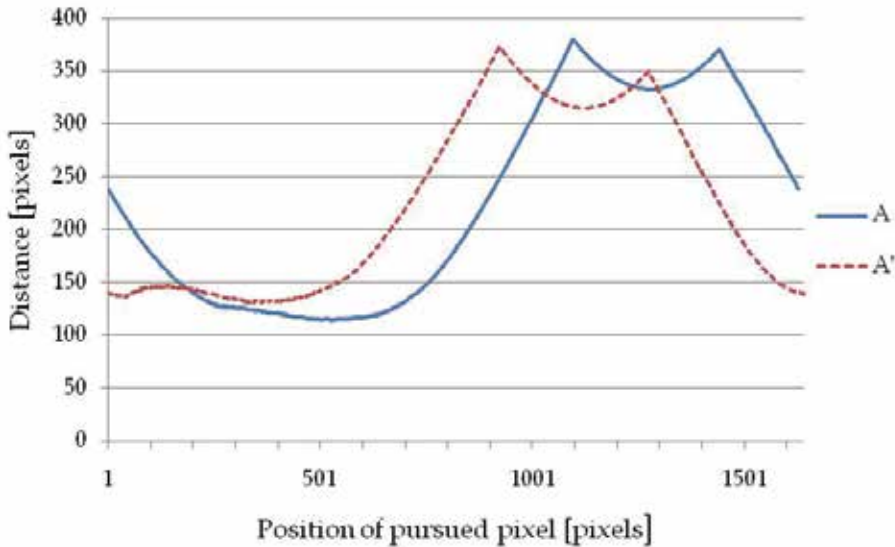
Fig. 14. Distance between the pixels on the finger outline and the reference artifact (A) and the artifact that was removed once and re-attached randomly (A′)

The result of the comparison between Data A and A′ is shown in Figure 14. From the plotted data, it is clear that the two lines are markedly different with respect to the shift on the horizontal axis, which was also reflected in the low correlation coefficient between Data A and A′ of 0.660. Thus, even if the genuine user attempts to access the system after removal of the artifact, re-enrollment is necessary. In Section 4.2, the finger shapes of a few imposters were quite similar to the genuine user. However, even when an imposter attempted to spoof using the genuine finger outline and an imitation finger, spoofing is prevented by the randomness of the position and angle of the artifact. In Section 5.1, we confirmed the security level of the proposed biometric identification method depending on the randomness of the position and angle of the artifact.

5. Validation of security level

To validate the security level of the proposed biometric identification method, the following two simulations were conducted.
SIM. 1: Security level depending on the position and angle of the artifact
SIM. 2: Security level depending on the amount of biological data
The level of security, as verified by each simulation, is an important factor for demonstrating the practical use of the proposed system. The details of each simulation are described in the following two subsections.

5.1 Security level depending on the artifact position and angle

In this simulation, we verified the allowable range of the position and angle of the artifact for identification when the artifact is removed and re-attached. Specifically, we attempted to determine the degree of change in the artifact position or angle that prevents the imposter from being verified by the system, as determined by the correlation coefficient. The position and angle of the artifact of Figure 5(a) were changed in the simulation program based on the following two conditions:

Condition 1: The artifact is moved in the direction of x (horizontal direction) and the direction of y (vertical direction) by one pixel (approximately 0.05 mm).
Condition 2: The artifact is rotated by one degree.

The results of the simulation under conditions 1 and 2 are shown in Figures 15 and 16, respectively. If the threshold value for identification is set at 0.995 based on the results presented in Section 4.2, and the artifact is moved 11 pixels (0.55 mm) or more in the x direction, or 10 pixels (0.50 mm) or more in the y direction, the correlation coefficient falls below the threshold level and the genuine user is not accepted into the system (Figure 15). If the acceptable range for placement of the artifact on the fingernail is assumed to be 5.0 × 5.0 mm, the randomness of the artifact position is calculated as follows:

$$5.0 / 0.55 × 5.0 / 0.50 = 90 \text{ patterns}$$

If the threshold value is set at 0.995 and the artifact is rotated 5.0 degrees or more, the genuine user is not accepted into the system (Figure 16). Under this condition, the randomness of the angle of the artifact is calculated as follows:

$$360 / 5 = 72 \text{ patterns}$$

Considering the combination between the position and angle of the artifact, and assuming that the position and angle are independent parameters, the randomness of the method is calculated as follows:

90 × 72 = 6480 patterns

The relative scale of the range of positions (0.55 × 0.50 mm) and angles (5.0 degrees) with respect to the finger outline are shown in Figure 17. From these simulations, it was clearly demonstrated that the randomness of the artifact is quite high. Therefore, if someone attempted to mimic the position and angle of the genuine artifact, the inherent randomess of the proposed identification system would effectively prevent such spoofing attempts.
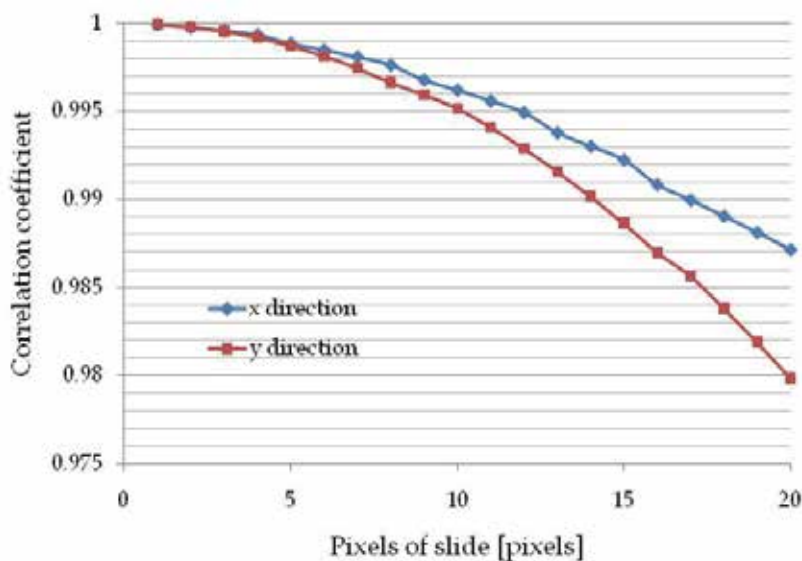

Fig. 15. Effect on the correlation coefficient by moving the artifact in the x (blue line) or y (red line) direction
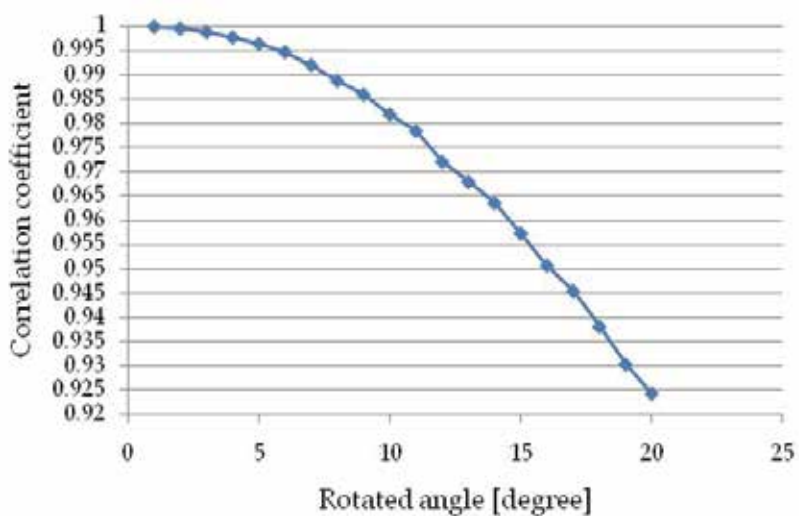

Fig. 16. Effect of rotating the artifact position on the correlation coefficient

Notably, this estimation is based on the placement range of 5.0 × 5.0 mm for the artifact; however, the acceptable range for placement of the artifact on the fingernail is thought to be even wider.

From the viewpoint of fingernail growth, which relates to Problem 1 described in the Introduction, the proposed method possesses the advantage of cancelable identification. It is estimated that the fingernails of adults grow approximately 0.1 mm per day. Thus, based on the simulation results and the constant growth rate of fingernails, a genuine user would need to re-enroll in the system within six days.

| (a) | (b) |
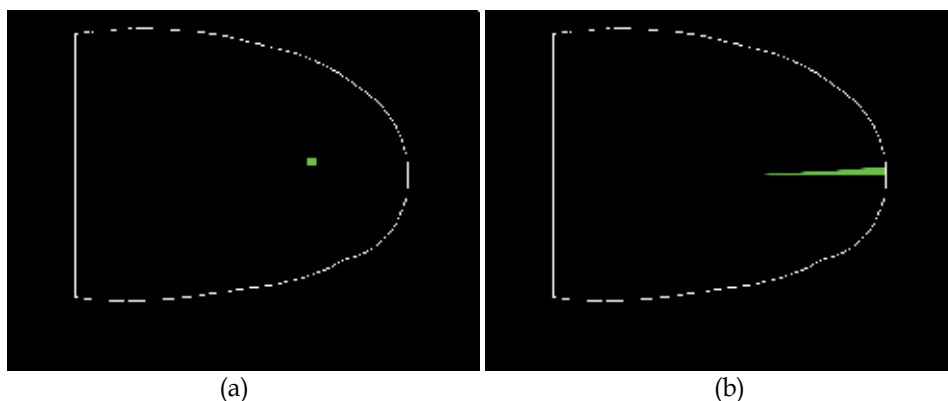
Fig. 17. Relative scale of the allowed range for identification with respect to the finger outline; (a) positional range (0.55 × 0.50 mm) and (b) angular range (5.0 degrees)

5.2 Security level depending on the amount of biological data

In a second simulation, we verified the relationship between the amount of biological data (finger outline data) and the security level of the proposed biometric identification system. The amount of finger outline data corresponds to the number of pursuit pixels counted from the starting point during the image processing step. All finger outline data shown in Figure 8, with the exception of the red area, were used for the comparison in the experiments in Section 4. Although the uniqueness of the finger outline is relatively low, it represents biological data, similar to that provided by fingerprints, veins, or the iris. Therefore, depending on the situation, it is conceivable that users may hesitate to enroll finger outline data in the identification system, which relates to Problem 2 described in the Introduction. Thus, we have proposed identification using biological data that is not specific to the user, but is specific only with respect the artifact position. In the second simulation, the amount of required biological data was examined by decreasing the amount of finger outline data that allowed distinguishing between the genuine user and an imposter.

All data of subject A and the imposters (B-E) were used for the simulation. Figure 18 is a graphical result of the simulation, where the vertical axis represents the correlation coefficient and the horizontal axis represents the decrease ratio of finger outline data. When the decrease ratio in the horizontal axis in Figure 18 is replaced with the number of data (the distance in pixels between the artifact and finger outline), 100% corresponds to 1283 and 0.26% corresponds to 4. From the simulation results, even if the decrease ratio is decreased by as much as 0.56% (the number of data is 8), the genuine user can be distinguished from imposters. Based on this simulation experiment, it is clear that the collected biological data

cannot be used to identify the user, but can be used to specify the position of the artifact and identify the genuine user. Moreover, the meaningfulness of the collected data can be canceled by simply removing the artifact. Thus, for identification using the proposed method, the users are not required to enroll their highly unique biological data, which is unavoidable using current biometric systems.
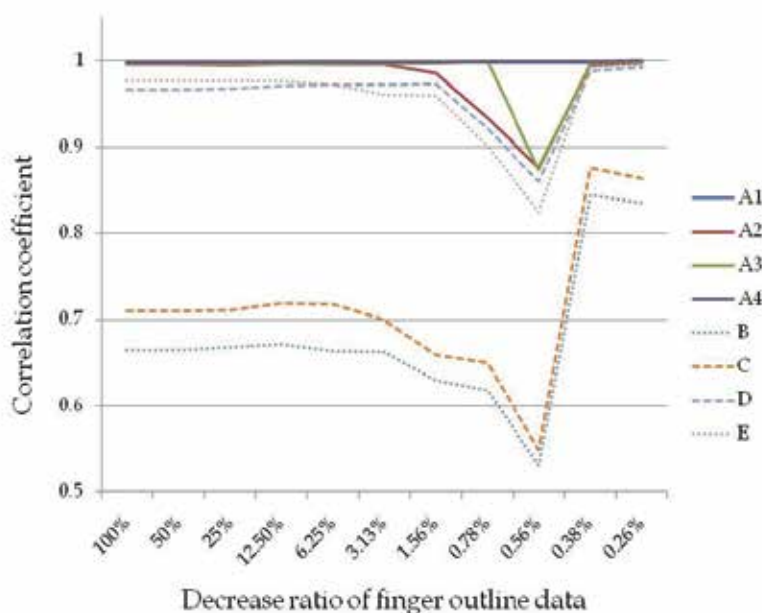
Fig. 18. Effects of decreasing the amount of finger outline data used in the identification on the correlation coefficient

6. Summary and proposed application system

6.1 Summary

The proposed cancelable biometric identification system has a number of advantages over current systems. The features of the proposed method are summarized as follows:

1. Cancelable biometric identification: Registered information can be canceled by simply removing the artifact. Even if the genuine user attempts to access the system, once the artifact is removed, re-enrollment is necessary (refer to Section 4.3). Moreover, due to the constant growth of fingernails, identification is not possible after a certain period of time (approximately one week).

2. Controllable security level: The security level of the system can be adjusted by controlling the amount of permissible biological and artifactual data.
    I. Artifacts:
        I-a. Using sufficient information to allow identification of the artifact (random pattern, code, or artifact-metrics proposed by Matumoto et al. (2000): the secondary biological data is artificially appended to the fingernail).
        I-b. Using only information that allows the position and direction of the artifact to be detected (a few dots).

    II.   Biological data:

        II-a. Using all information that allows identification of the user (all data shown in Figure 8, except the red area).

        II-b. Using only information that allows the position and direction of the artifact to be detected. In other words, the user cannot be identified only by biological data (refer to Section 5.2).

When (I-a) and (II-a) are combined, three identifications are performed. The first is the identification of the biological data, the second is identification of the artifact, and the third is the relation between the biological data and the artifact. Yamagishi et al. (2008) described this method using fingerprints and a random pattern as an artifact. When (I-b) and (II-b) are combined, only the third identification is conducted.

3.   Non-necessity for the registration of unique biological information: Although the outline of the finger constitutes biological data, the information it provides in itself is not sufficient for individual identification. Moreover, when the amount of biological data collected for identification is decreased, it is impossible to identify an individual (refer to Section 5.2). Therefore, the enrollment of biological data that can uniquely identify the user is unnecessary in this proposed biometric verification system.

4.   Strength against spoofing: If the biological data and artifact are stolen, spoofing can be prevented due to difficulty of reproducing the biological data in relation to the artifact that is required for identification (refer to Sections 4.1, 4.2, 4.3, and 5.1).

6.2 Proposed application of the system

The conditions for the application of the proposed method are as follows: (a) only a limited number of enrollments can be accommodated by the system, and (b) the use period is approximately five days. The flow of use of the proposed system at a hotel is illustrated in Figure 19. At the time of check-in, the user attaches the artifact to the fingernail and enrolls
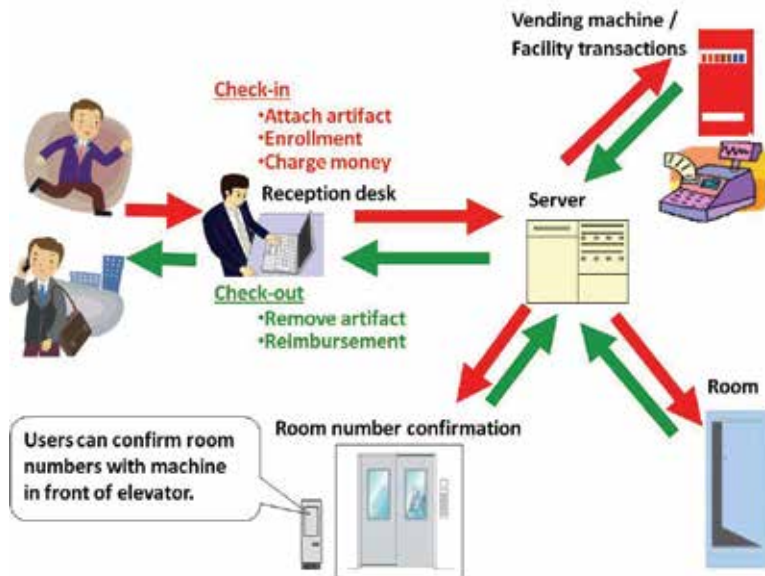
Fig. 19. Potential applications of the proposed system in a hotel setting

the artifact and biological data with the system. After the enrollment, the user can enter and exit his/her room without possession of a room key. Moreover, the possession of a wallet or purse becomes unnecessary while the user is within the hotel facilities, thus improving safety and convenience. For practical use, a transparent sticker containing the two dots marked with a dye that can only be detected using specific lighting (Takayanagi, 2009) would serve as the artifact (Figure 20).
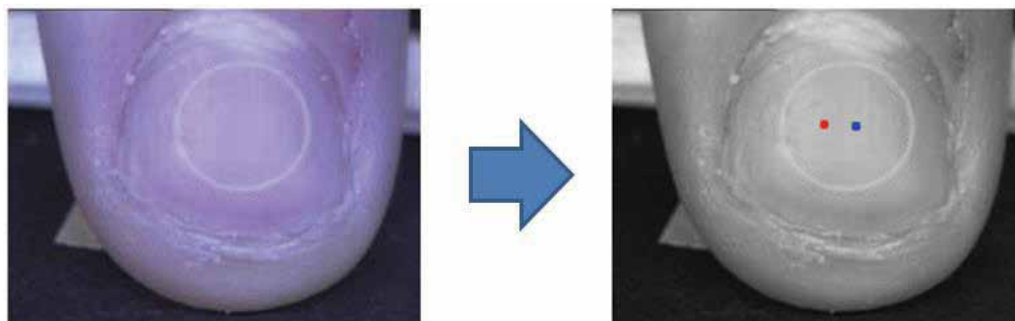
Fig. 20. Image of a fingernail with a transparent artifact (left: appearance under natural light, right: appearance under specific lighting)

Under these conditions, the proposed biometric identification system would also be suitable for use as a one-day pass for office and factory buildings, and amusement and medical facilities, among numerous other potential applications.

7. Conclusions

We have described a novel method of cancelable biometric identification that combines biological data with the use of an artifact. The algorithm of the identification step can be divided into four parts: processing of the input image for the artifact; image processing for the finger; feature extraction, which involved determining the distance between the artifact and finger outline; and comparison of the reference and probe data. Based on the results of the three evaluative experiments and two simulations described here, several strengths of the proposed method can be recognized. First, the proposed method is a type of cancelable biometric identification, as registered information can be canceled by simply removing the artifact. Second, the proposed method allows control of the security level by adjusting the amount of biological and artifactual data. Third, the registration of unique biological information is not necessary for the identification system. Finally, the proposed method is resistant to spoofing.

Despite these apparent strengths, a few limitations of the proposed method warrant mention. First, the application of the proposed method is limited by two conditions: (a) only a limited number of enrollments can be accommodated by the system, and (b) the use period is approximately five days. Although the potential field of applications is limited by these two conditions, the proposed method is characterized by user friendliness and relative simplicity that do not exist in current identification methods. Second, the usability of the identification system should be improved. Specifically, it is necessary to develop a material for use as the artifact that remains firmly in place and a mechanism that permits the user to easily detach the artifact on exiting the system.

8. Acknowledgments

9. References

Ashbourn, J. (2000). Biometrics: Advanced Identity Verification, The Complete Guide, Springer-Verlag.

Gunn, L. (2010). VIEWPOINT: Would you use a biometric solution to authorise a transaction at an ATM?, European ATM Security Team (EAST), https://www.european-atm-security.eu/Welcome%20to%20EAST/&action=fullnews&id=62. Accessed in April 2011.

Hirabayashi, M., Tanabe, T., and Matsumoto, T. (2004). Can We Make Artificial Fingers That Fool Fingerprint Systems? (Part VI), Technical Report of Institute of Electronics, Information, and Communication Engineers, ISEC2003-139, pp. 151-154.

Jain, A. K.; Ross, A. and Prabhakar, S. (2004a).An Introduction to Biometric Recognition, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, pp. 4-20.

Jain, A. K., Pankanti, S., Prabhakar, S., Hong, L., Ross, A., and Wayman, J. L. (2004b). Biometrics: A Grand Challenge, Proceedings of International Conference on Pattern Recognition, pp. 935 – 942.

Matsumoto, H., Matsumoto, T. (2000). Artifact-metric Systems, Technical Report of the Institute of Electronics, Information and Communication Engineers, ISEC2000-59, pp. 7-14.

Matsumoto, T. (2006). Biometric Authentication Systems: Vulnerability of Biometric Authentication - On the Issue of Physiological Spoofing –, IPSJ (Information Processing Society of Japan) Magazine, Vol. 47, No. 6, pp. 589-594.

Nishiuchi, N., Komatsu, S., Yamanaka, K. (2010). Biometric verification using the motion of fingers: a combination of physical and behavioural biometrics, International Journal of Biometrics, Vol. 2, No. 3, pp. 222-235 .

Prabhakar, S., Pankanti, S., Jain, A. K. (2003). Biometric Recognition: Security & Privacy Concerns", IEEE Security and Privacy Magazine, Vol. 1, No. 2, pp. 33-42.

Rotter, P., Daskala, B., Compañó, R. (2008). RFID implants: opportunities and challenges for identifying people. IEEE Technology and Society Magazine, Vol. 27, Issue 2, pp. 24–32.

Stén, A., Kaseva, A., Virtanen, T. (2003). Fooling Fingerprint Scanners - Biometric Vulnerabilities of the Precise Biometrics 100 SC Scanner, Proceedings of 4th Australian Information Warfare and IT Security Conference 2003, pp. 333-340.

Takayanagi, Y., Kamijo, K., Katto, J. (2009). Invisible Barcode Extraction using Color Channel Estimation, Technical report of IEICE. Multimedia and virtual environment 109(149), pp. 31-36.

Wayman, J. (2000). National Biometric Test Center Collected Works 1997-2000, pp. 1–3.

Yamada, K., Matsumoto, H., and Matsumoto, T. (2000). Can We Make Artificial Fingers That Fool Fingerprint Systems?, Technical Report of Institute of Electronics, Information, and Communication Engineers, ISEC2000-45, pp. 159-166.

Yamagishi, M., Nishiuchi, N., Yamanaka, K. (2008). Hybrid Fingerprint Authentication Using Artifact-Metrics, International Journal of Biometrics, Vol. 1, No. 2, pp. 160-172.

Part 3

Application of Encryption

Biometric Keys for the Encryption of Multimodal Signatures

A. Drosou[1], D.Ioannidis[2], G.Stavropoulos[2], K. Moustakas[2] and D. Tzovaras[2]

[1]*Imperial College London*
[2]*Ce.R.T.H. - Informatics and Telematics Institute*
[1]*UK*
[2]*Greece*

1. Introduction

Biometrics have long been used as means to recognize people, mainly in terms of their physiological characteristics, for various commercial applications ranging from surveillance and access control against potential impostors to smart interfaces (Qazi (2004)) (Xiao (2005)). These systems require reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. The biometric methods, that are usually incorporated in such systems, can be categorized to physiological and behavioral (Jain et al. (2004)), depending on the type of used features.

The most popular physiological biometric traits are the fingerprint (Maltoni et al. (2009)) that is widely used in law enforcement for identifying criminals, the face (Chang et al. (2005)) and the iris (Sun & Tan (2009)). However, despite their high recognition performance, static biometrics have been recently overwhelmed by the new generation of biometrics, which tend to cast light on more natural ways for recognizing people by analyzing behavioural traits.

Specifically, behavioural biometrics are related to specific actions and the way that each person executes them. In other words, they aim at recognizing livingness, as it is expressed by dynamic traits. The most indicative cases of behavioural biometric recognition is gait (Goffredo et al. (2009b)), facial expressions (Liu & Chen (2003)) or other activity related, habitual traits (Drosou, Ioannidis, Moustakas & Tzovaras (2010)). As a result behavioural biometrics have become much more attractive to researchers due to their significant recognition potential and their unobtrusive nature. They can potentially allow the continuous (on-the-move) authentication or even identification unobtrusively to the subject and become part of an Ambient Intelligence (AmI) environment.

The inferior performance of behavioural biometrics, when compared to the classic physiological ones, can be compensated when they are combined in a multimodal biometric system. In general, multimodal systems are considered to provide an excellent solution to a series of recognition problems. Unimodal systems are more vulnerable to theft attempts, since an attacker can easily gain access by stealing or bypassing a single biometric feature. In the same concept, they have to contend with a variety of problems, such as noisy data, intraclass variations, restricted degrees of freedom, non-universality, spoof attacks, and unacceptable error rates, i.e., it is estimated that approximately 3% of the population does not have legible

fingerprints (Fairhurst et al. (2003)). Such biometric system may not always meet performance requirements, may exclude large numbers of people, and may be vulnerable to everyday changes and lesions of the biometric feature.

In this context, the development of systems that integrate more than one biometrics is an emerging trend, since it has been seen that true multimodal biometric systems, that capture a number of unrelated biometrics indicators, have significant advantages over unimodal ones. Specifically, most of the aforementioned limitations can be addressed by deploying multimodal biometric systems that integrate the evidence presented by multiple sources of information. A multimodal biometric system uses multiple applications to capture different types of biometrics. This allows the integration of two or more types of biometric recognition systems, in order to meet stringent performance requirements. Moreover, such systems are much more invulnerable to fraudulent technologies, since multiple biometric characteristics are more likely to resist against spoof attacks than a single one.

Last but not least, a major issue of biometric systems is the protection of the sensitive biometric data that are stored in the database, so as to prevent unauthorized and malicious use. Given the widespread deployment of biometric systems and the wide exposition of personal data, public awareness has been raised about security and privacy of the latter. Seemingly, the voting of several laws concerning the ethical and privacy issues of private data provide a universal solution unless it is accompanied by the appropriate technological tools.

Unfortunately, simple password-based systems, that provide regular cryptographic solutions (Uludag et al. (2004)) can not be easily applied, since the representation of behavioural biometric traits is not fixed over time. Thus, the current issue has been confronted with modern, sophisticated encryption methods that do not require the exact match of the prompted and the original signatures in order to grant access.

1.1 Related work

With respect to behavioural biometrics, previous work on human identification can be mainly divided in two main categories. a) sensor-based recognition (Junker et al. (2004)) and b)vision-based recognition. Recently, research trends have been moving towards the second category, due to the obtrusiveness imposed by the sensor-based recognition approaches (Kale et al. (n.d.)). Additionally, recent work and efforts on human recognition have shown that the human behavior (e.g. extraction of facial dynamics features (Hadid et al. (2007)). However, the most known example of behavioural biometrics is the human body shape dynamics (Ioannidis et al. (2007) or joints tracking analysis (Goffredo et al. (2009a)) for gait recognition. In the same respect, the analysis of dynamic activity-related trajectories (Drosou, Moustakas & Tzovaras (2010)) provide the potential of continuous authentication for discriminating people, when considering behavioural signals.

Although there have been already proposed a series of multimodal biometric systems concerning static physiological biometric traits (Kumar et al. (2010)) (Sim et al. (2007)) there are only a few dealing solely with behavioural traits (Drosou, Ioannidis, Moustakas & Tzovaras (2010)). In any case, the main issue in a multimodal biometric system is the optimization of its fusion mechanism. In a multimodal biometric system, integration can be done at (i) feature level, (ii) matching score level, or (iii) decision level. However, matching score level fusion is commonly preferred because matching scores are easily available and contain sufficient information to distinguish between a genuine and an impostor case. In this respect, a thorough analysis of such score-level fusion methods regarding biometric traits has been presented in (Jain et al. (2005)).

Since all biometric systems deal with the issue of storing biometric data, different approaches regarding their security have been suggested. In the current work, an extension of the security template scheme, presented in (Argyropoulos et al. (2009)), is proposed, that bases on Error Correcting Codes (*ECC*) and the modeling of channel statistics. Channel codes have been previously used for the development of authentication schemes. Earlier, in (Wadayama (2005)), a generic authentication scheme based on channel codes was proposed to improve security and prevent unauthorized access in secure environments. Also, in (Davida et al. (1998)), a channel coding scheme was presented for secure biometric storage. Error correcting codes were employed to tackle the perturbations in the representation of biometric signals and classification was based on the Hamming distance between two biometric representations. Based on this concept, the fuzzy commitment scheme was introduced to tolerate more variation in the biometric characteristics and provide stronger security (Juels & Sudan (2006)). In this scheme, the user selects at the enrolment a secret message c. Then, the template consists of the difference between the user's biometric data x and c along c with an encrypted version of c. At the authentication, the stored difference d is added to the new biometric representation y and $y + d$ is decoded to the nearest codeword c' using error correcting codes.

In this respect, a series of encryption methods have been developed to account for the inherent variability of biometric signals. Apart from (Davida et al. (1998)), a methodology based on the Slepian-Wolf theorem (Slepian & Wolf (1973)) for secure storage biometric via Low-Density Parity Check (*LDPC*) codes was presented in (Martinian et al. (2005)). The multimedia authentication problem in the presence of noise was investigated, the theoretical limits of the system were identified, and the tradeoff among fidelity, robustness, and security was discussed. This approach provides intuition for the proposed method in this paper; the biometric recognition problem is considered as the analogous of data transmission over a communication channel, which determines the efficiency of the system. Interestingly, the problem of coding distributed correlated sources has also attracted much interest in the field of video coding recently. The same framework was also employed in (Draper et al. (2007)) in order to secure fingerprint biometrics, image authentication (Yao-Chung et al. (2007)) and biometric authentication as a wire-tap problem (Cohen & Zemor (2004)).

In the seminal work of (Pradhan & Ramchandran (2003)), the distributed source coding using syndromes scheme was proposed. Based on this work, the field of distributed video coding (Girod et al. (2005)) has emerged as a new trend in video coding. Finally, an interesting approach of applying the *LDPC* methodology in multimodal biometric systems has been proposed in (Argyropoulos et al. (2009)).

Similarly to above, one of the major concerns in applications that grant access based on a password, a pin or a token, is the protection of the original data to prevent malicious use from those who try to access them by fraudulent means. Although this problem in such systems has been investigated in depth and sophisticated encryption methods have been developed (Stallings (2006)), a significant issue remains the possibility of having the password stolen or forgotten. Thus, methods which enable a biometric-related key have been proposed (Álvarez et al. (2009)). Thus, the required pin is always carried by the user, since it is encoded on himself.

1.2 Contribution

In the current chapter, a novel framework for activity related authentication in secure environments based on distributed source coding principles and automatically extracted biometric keys is proposed. The main novelty is the development of an integrated framework

that utilizes biometric key based encryption, in order to assist the channel decoding process and to boost the system's recognition performance. It is shown that the proposed system increases the security of the stored templates and ensures privacy of personal data, while indirectly provides "hybrid" fusion between static and dynamic biometric traits towards improved recognition results. Moreover, unobtrusive, multimodal, on-the-move biometric authentication is presented and evaluated in a bimodal scenario, which utilizes two behavioural traits of the user. Namely, the gait and the activity-related motion trajectories of the head and the hands during specific movements which are seen to provide a powerful auxiliary biometric trait are inspected in terms of biometric means for user authentication.

2. Proposed methodology

The architecture of the proposed biometric recognition framework is illustrated in Figure 1. Initially, from the captured gait image sequence, the moving silhouettes are extracted, the shadows are removed and the gait cycle is estimated using state-of-art (SoA) algorithms (Ioannidis et al. (2007)), (Cucchiara et al. (2001)). Using a stereoscopic camera, we detect those frames in the sequence, whereby the user is standing and we discard them from those where the user is walking. Then the visual hull of the moving silhouette is extracted using disparity estimation. Once a view normalization is applied by rotating the silhouette, the $3D$ reconstructed silhouettes are denoised via spatiotemporal filtering, in order to improve their quality. Finally, two SoA geometric descriptors are extracted based on the sequence Gait Energy Image (GEI).
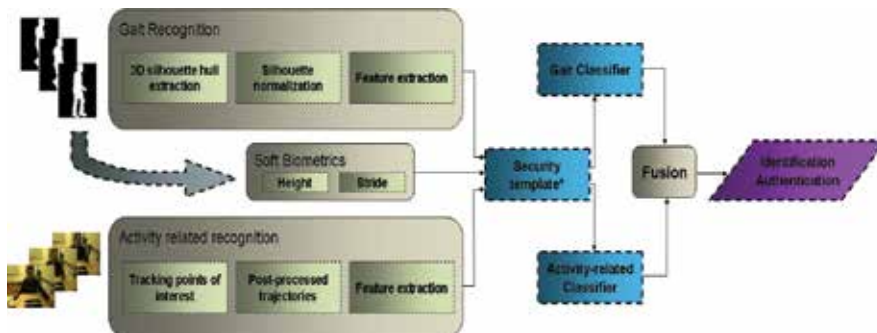
Fig. 1. System Architecture.

The gait recognition follows the principle of a model-free, feature-based analysis of the extracted human silhouettes, whereby geometric methods implement a robust classifier. In the following, the activity-related recognition is performed on the users' movements while they interact with a security panel installed at the end of the corridor. The extracted motion trajectories that are used as the user's biometric traits are classified by a Dynamic Time Warping classifier and its result is finally fused with the corresponding gait results at the score level towards an overall recognition outcome.

3. Behavioural biometrics

As it has already been mentioned, the development a novel biometric recognition method or the improvement of current State of Art (SoA) methodologies in this area, are not within the scope of the current work. In this context, a set of simple but robust activity-related

recognition modules have been utilized in the context of the proposed security framework in order to build a behavioural multimodal recognition system, where the proposed enhanced security template framework (see Section 4) could be tested and evaluated.

In particular, the first biometric modality consists of *SoA* gait recognition methodology (Ioannidis et al. (2007)) that bases on features extracted from spatiotemporal gait patterns. Similarly, the second modality that has been utilized refers to a novel activity-related concept that has been initially proposed in (Drosou, Moustakas, Ioannidis & Tzovaras (2010)) and deals with the motion related traits left by the user during the performance of some simple activities that are performed on a regular basis. Both aforementioned modalities are not only directly related to the users' physiology, but they are also highly governed by the users' habitual response to external stimuli. Thus, they have been seen to provide significant recognition capacity, both as stand-alones, as well as in multimodal recognition systems (Drosou, Ioannidis, Moustakas & Tzovaras (2010)).

For the convenience of the reader, a short functional description of the aforementioned modalities is included hereafter. Before presenting the security framework, which is the main contribution of the current work, a short description of the utilized biometric modalities is included.

3.1 Gait recognition

Let the term *"gallery"* refer to the set of reference sequences, whereas the term *"probe"* stands for the test sequences to be verified or identified, in both presented modalities.

Initially, the walking human binary silhouette is extracted as described in (Ioannidis et al. (2007)). The feature extraction process of the gait sequences is based on the Radial Integration Transformation (RIT) and the Circular Integration Transform (CIT) (Ioannidis et al. (2007)), but instead of applying those transforms on the binary silhouette sequences themselves, the Gait Energy Images (GEI) are utilized, which have been proven from one hand to achieve remarkable recognition performance and on the other hand to speed up the gait recognition (Han et al. (2006)) (Yu et al. (2010)).

Given the extracted binary gait silhouette images I' and each gait cycles k, the gray level (GEI) (Figure 2) is defined over a gait cycle as:

$$GEI_k = \frac{1}{C_L} \cdot \sum_{k=CycleStart}^{CycleEnd} I'(k) \tag{1}$$

where C_L is the length of the gait cycle and k refer to the gait cycles extracted in the current gait image sequence.
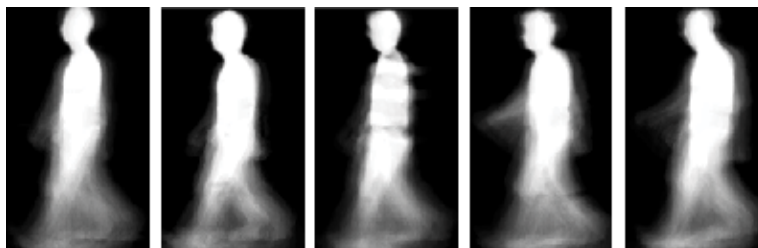
Fig. 2. Gait Energy Images from several users.

The RIT and CIT transforms are applied on the GEI, in order to construct the gait template for each user, as shown in Figure 3 in according to the following equations:

$$RIT_{f(\theta)} = \int f(x_0 + u\cos\theta, y_0 + u\sin\theta)du \qquad (2)$$

where u is the distance from the starting point (x_0, y_0).

$$RIT(t\Delta\theta) = \frac{1}{J}\sum_{j=1}^{J} GEI(r_0 + j\Delta u \cdot \cos(t\Delta\theta), y_0 + j\Delta u \cdot \sin(t\Delta\theta)) \qquad (3)$$
$$\text{for } t = 1,...,T \text{ with } T = 360^0/\Delta\theta$$

for $t = 1,...,T$ with $T = 360^0/\Delta\theta$, where $\Delta\theta$ and Δu are the constant step sizes of the distance u and angle θ and J is the number of the pixels that coincide with the line that has orientation R and are positioned between the center of gravity of the silhouette and the end of the image in that direction.
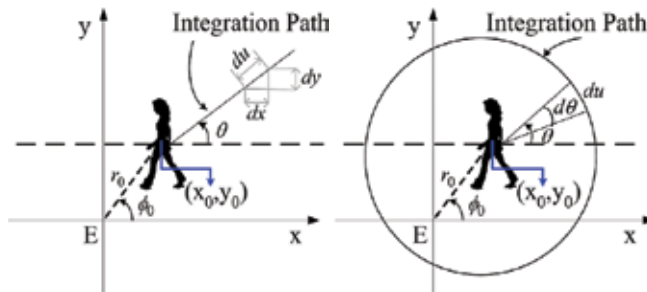
Fig. 3. Applying the RIT (left) and CIT (right) transforms on a Gait Energy Image using the Center of Gravity as its origin.

Similarly, CIT is defined as the integral of a function $f(x,y)$ along a circle curve $h(\rho)$ with center (x_0, y_0) and radius ρ. The CIT is computed using the following equation:

$$CIT_{f(\rho)} = \oint_{h(\rho)} f(x_0 + \rho\cos\theta + \rho\sin\theta)du \qquad (4)$$

where du is the arc length over the path of integration and θ is the corresponding angle.
The center of the silhouette is again used as the origin for the CIT. The discrete form of the CIT transform is used, as depicted graphically in Figure 3/right.

$$CIT(k\Delta\rho) = \frac{1}{T}\sum_{t=1}^{T} GEI(x_0 + k\Delta\rho \cdot \cos(t\Delta\theta), y_0 + k\Delta\rho \cdot \sin(t\Delta\theta)) \qquad (5)$$

for $k = 1,...,K$ with $T = 360^0/\Delta\theta$, where $\Delta\rho$, and $\Delta\theta$ are the constant step sizes of the radius and angle variables and finally $K\Delta\rho$ is the radius of the smallest circle that encloses the grayscaled GEI (Figure 2).
The extracted RIT and CIT feature vectors are then concatenated, in order to form a single 1D biometric trait.

3.1.1 Matching

The comparison between the number of gallery G_{GEI} and probe P_{GEI} gait cycles for a specific feature $E \in \{RIT, KRM\}$ is performed through the dissimilarity score d_E.

$$d_E = \min_{i,j} \left(||\mathbf{s}_i^G - \mathbf{s}_i^P|| \right) \ \forall i,j; i \in [1, G_{GEI}) \text{ and } j \in [1, P_{GEI}) \tag{6}$$

whereby $|| \cdot ||$ is the L_2-norm between the $\mathbf{s}^G$ and $\mathbf{s}^P$ values of the corresponding extracted feature (i.e. RIT & CIT) for the gallery and the probe collections, respectively.

3.2 Activity-related recognition

The proposed framework extends the applicability of activity-related biometric traits (Drosou, Moustakas, Ioannidis & Tzovaras (2010)), and investigates their feasibility in user authentication applications.

In (Kale et al. (2002)) and (Drosou, Moustakas & Tzovaras (2010)), it is claimed that the traits of a subject's movements during an activity that involves reaching and interacting with an environmental object can be very characteristic for recognition of his/her identity. Indeed, given the major or minor physiological differences between users' bodies in combination with their individual inherent behavioural or habitual way of moving and acting it has been reported that there is increased authentication potential in common everyday activities such as answering a phone call, etc.

In the following, an improved activity-related recognition framework is proposed, that employs a novel method for the normalization of the trajectories of the user's tracked points of interest. The proposed algorithm also introduces a warping method that compensates for small displacements of the environmental objects and has no effect on the behavioural information of the movement at all.

As of today, activity related biometrics, where the activity is associated with reaching and interacting with objects, have always assumed a fixed environment (Drosou, Moustakas & Tzovaras (2010)), which is not always the case in real life scenarios. However, significant performance degradations can be observed due to the small variances in the interaction setting, which are introduced by the arbitrary position of the environmental objects in respect to the user at each trial. Thus, a post-processing algorithm towards the improvement of the overall authentication performance that can be employed into biometric systems which utilize the reaching and interacting concept, is presented in the following.

3.2.1 Motion trajectory extraction

The core of the proposed authentication system used on dynamic motion tracking (4f) is extensively described in (Drosou, Moustakas, Ioannidis & Tzovaras (2010)) and is briefly described in the following so as to make the paper self-contained. The userŠs movements are recorded by a stereo camera and the raw captured images are processed, in order to track the users head and hands via the successive application of filtering masks on the captured image.

Specifically, a skin-colour mask (Gomez & Morales (2002)) (4a) combined with a motion-mask (Bobick & Davis (2001)) (Figure 4d) can provide the location of the palms, while the head can be accurately tracked via a combination of a head detection algorithm (Viola & Jones (2004)) enhanced by a mean-shift object tracking algorithm (Ramesh & Meer (2000)) (4b). Given the pre-calibrated set of CCD sensors mounted on the stereo camera, the real 3D information can be easily calculated first by performing disparity estimation (4c) from the
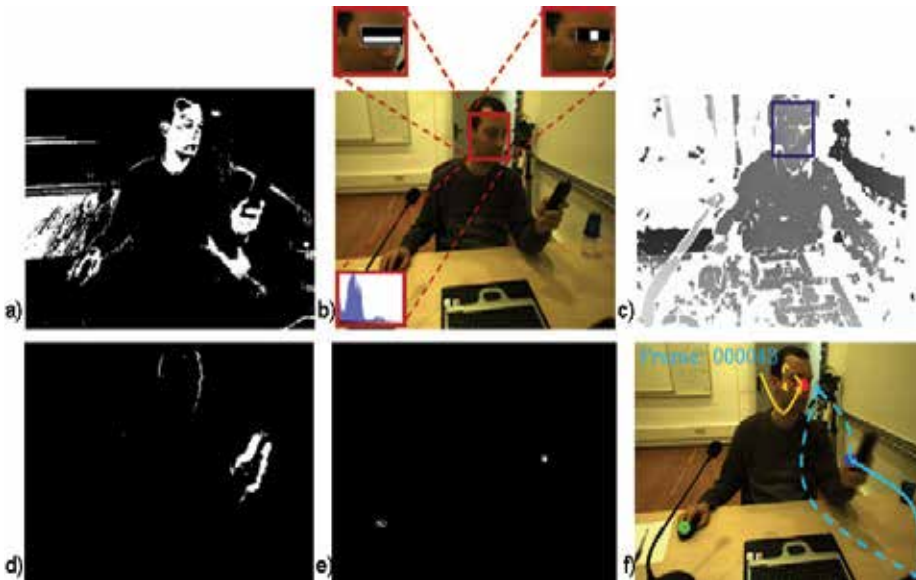
Fig. 4. Tracking methodology: a) Skin filtering - b) Head Tracking - c) Disparity image - d) Motion Detection - e) Possible hands' locations - f) Motion trajectories.

input stereoscopic image sequence and then by mapping the 2.5D information onto the $3D$ space. After post-processing (Drosou, Moustakas, Ioannidis & Tzovaras (2010)) that is applied on the raw tracked points, based on moving average window and Kalman filtering, equally sized and smooth $3D$ motion trajectories are extracted (Figure 5), which are then used as activity related biometric traits for proposed modality.

A motion trajectory for a certain limb l (head or palms) is considered as a $3D$ N-tuple vector $\mathbf{s_l(t)} = (x_l(t), y_l(t), z_l(t))$ that corresponds to the x,y,z-axes location of limbs center of gravity at each time instance t of an $N - frame$ sequence. The x,y and z data of the trajectories $\mathbf{s_l}$, are concatenated into a single vector and all vectors, produced by the limbs that take part in a specific activity c form the trajectory matrix S_c. Each repetition of the same activity by a user creates a new matrix. Both gallery and probe user-specific set of matrices are subsequently used as input to the Dynamic Time Warping (DTW) algorithm 3.2.2 that has been utilized as classifier for the current biometric modality, in order to provide an authentication score with respect to the claimed ID (gallery).

3.2.2 Matching via DTW

DTW is used for calculating a metric about the dissimilarity between two (feature) vectors. It is based on the difference cost that is associated with the matching path computed via dynamic programming, namely the Dynamic Time Warping (DTW) algorithm. The DTW algorithm can provide either a valuable tool for stretching, compressing or aligning time shifted signals (Sakoe & Chiba (1990)) or a metric for the similarity between two vectors (Miguel-Hurtado et al. (2008)). Specifically, it has been widely used in a series of matching problems, varying from speech processing (Sakoe & Chiba (1990)) to biometric recognition applications (Boulgouris et al. (2004)). The matching between the two vectors is done and a path is found using a rectangular grid (Figure 6).
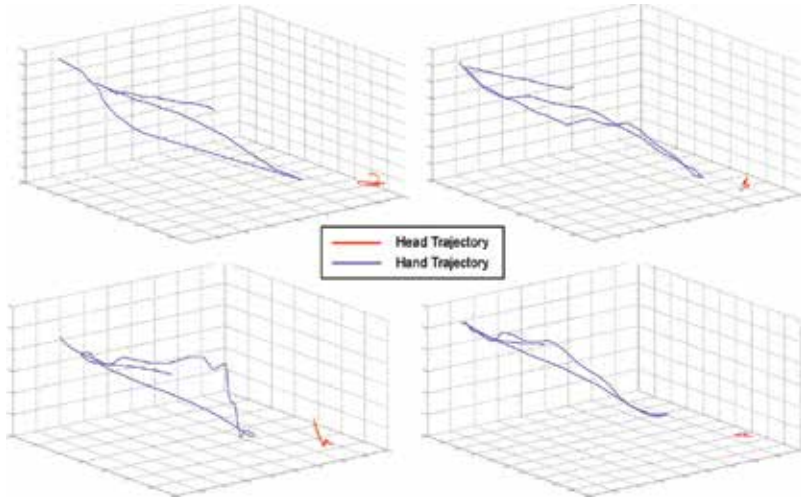
Fig. 5. 3*D* Motion Trajectories extracted during a "Phone Conversation" activity.

A short description of the functionality of *DTW* algorithm for comparing two one-dimensional vectors (probe & gallery signal) is presented below:

The probe vector **p** of length L is aligned the X-axis while the gallery vector **g** of length L' is aligned along the Y-axis of a rectangular grid respectively. In our case $L \equiv L'$ as a result of the preprocessing steps (Section 3.2.1). Each node (i,j) on the grid represents a match of the i_{th} element of **p** with the j_{th} element of **g**. The matching values of each $\mathbf{p}(i),\mathbf{g}(j)$ pair are stored in a cost matrix C_M associated with the grid. $c(1,1) = 0$ by definition and all warping paths are a concatenation of nodes starting from node $(1,1)$ to node (L,L).

The main task is to find the path for which the least cost is associated. Thus the difference cost between the two feature vectors is provided. In this respect, let $(y_1(k), y_2(k))$ represent a node on a warping path at the instance t of matching. The full cost $D(y_1, y_2)$ associated to a path starting from node $(1,1)$ and ending at node $(y_1(K), y_2(K))$ can be calculated as:

$$D(y_1, y_2) = D(y_1(k-1), y_2(k-1)) + c(y_1, y_2) = \sum_{m=1}^{k} c(y_1(m), y_2(m)) \tag{7}$$

Accordingly, the problem of finding the optimal path can be reduced to finding this sequence of nodes $(y_1(k), y_2(k))$, which minimizes $D(y_1(k), y_2(k))$ along the complete path.
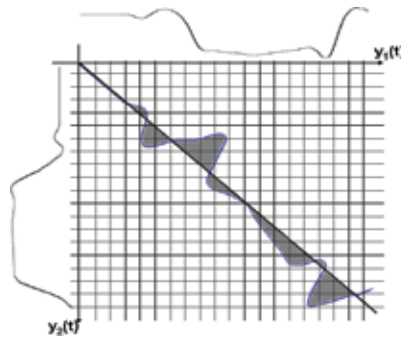
Fig. 6. Dynamic Time Warping Grid.

As stated by Sakoe and Chiba in (Sakoe & Chiba (1990)), a good path is unlikely to wander very far from the diagonal. Thus, the path with minimum difference cost, would be the one that draws the thinnest surface around the diagonal as shown by the dashed lines in Figure 6. In the ideal case of perfect matching between two identical vectors, the area of the drawn surface would be eliminated. The size of the closed area S_A around the diagonal can be calculated by counting the nodes $V(p_i, q_j)$ between the path and the diagonal at every row (Jayadevan et al. (2009)) as indicated by the following equation.

$$V(p_i, q_j) = \begin{cases} 1 \text{ , if } (i > j) \text{ of } N(p_i, q_j) \\ \qquad \text{for } j = j, j+1, ..., j+d, \text{ where } d = i - j \\ 1 \text{ , if } (i < j) \text{ of } N(p_i, q_j) \\ \qquad \text{for } i = i, i+1, ..., i+d, \text{, where } d = i - j \\ 1 \text{ , if } (i = j) \text{ of } N(p_i, q_j) \\ 0 \text{ , otherwise} \end{cases} \tag{8}$$

Thus, the value $V(p_i, q_j) = 1$ to these nodes. On the contrary, all other nodes lying outside the closed area will be assigned the value $V(p_i, q_j) = 0$. Then, the total area S_A created by the path is mathematically stated as following:

$$S_A = \sum_{i=1}^{L} \sum_{j=1}^{L} V(p_i, q_j) \tag{9}$$

whereby
Finally the total dissimilarity measure D_M between vector **p** and **g** (Equation 9) can be computed as the product of area size S_c and the minimum full cost $D(L, L)$ (Equation 7):

$$D_M = S_A \cdot D_{min}(L, L) \tag{10}$$

4. Biometric template security architecture

As far as the security of the biometric data is regarded, multiple feature descriptors from the gait modality and the activity-related modality are initially encoded via a Low Density Parity Check ($LDPC$) encoder. In the following, the parity bits of the activity-related modality are encrypted via a biometric-dependent key, so that double secured, non-sensitive biometric data is stored in the database or in smart cards, which are useless to any potential attackers of the system.

The proposed method, which resembles a channel coding problem with noisy side information at the decoder, is shown to improve the authentication rates as they are provided from the modality-specific classifiers. Additionally to the already known key-encryption methodologies, the encryption of the parity bits of the second modality takes place before their storage to the database. The novelty lies in the fact the personal encryption/decryption key used, is inherent in the biometric trait of the first modality and thus, it remains unknown even to each user. Specifically, in the implemented scenario the biometric key is selected according to the height and the stride length of the user.

The architecture (Figure 7) of the proposed security is thoroughly described in the next two Sections, whereby a functional analysis of the utilized distinct components is provided.
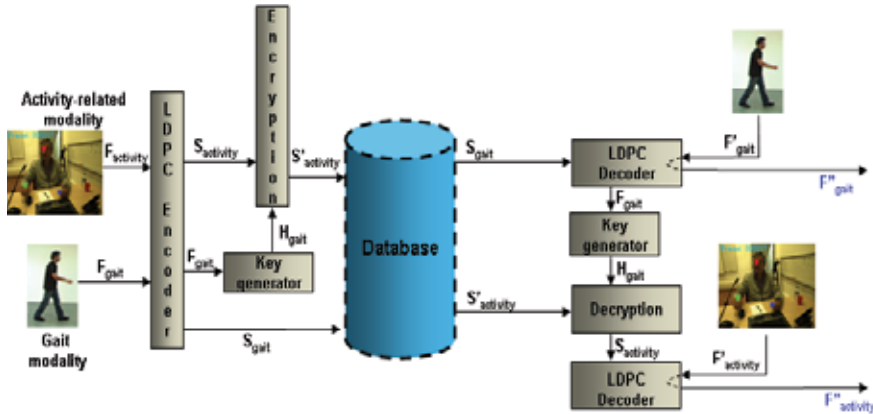
Fig. 7. Security subsystem Architecture.

4.1 Encoding scheme

The first step towards biometric template protection in the current multimodal biometric cryptosystem is based on distributed source coding principles and formulates biometric authentication as a channel coding problem with noisy side information at the decoder, as presented in (Argyropoulos et al. (2009)). The main idea lies on the fact that perturbations in the representation of the biometric features at different times can be modelled as a noisy channel, which corrupts the original signal. Thus, the enrolment and authentication procedures of a biometric system are considered as the encoding and decoding stages of a communication system, respectively. The proposed formulation enables the exploitation of the Slepian-Wolf theorem to identify the theoretical limits of the system and minimize the size of the templates. Moreover, casting the problem of biometric authentication as a communication problem allows the use of well known techniques in communication systems such as the exploitation of correlation (or noise) channel statistics by integrating them in the soft decoding process of the channel decoder.

The architecture of the multimodal biometric authentication system is included in Figure 7. At the enrolment stage, the feature vectors F_{gait} and $F_{activity}$ from the *Gait* and the *Activity - related* modality are initially extracted as described in the previous section. Then, the extracted feature vectors are quantized and encoded using an (n, k) *LDPC* channel encoder. It must be stressed that the rate of the *LDPC* encoders in Figure 7 is different for each modality and is calculated according to the Slepian-Wolf theorem

$$R_X \geq H(X|Y) \; R_Y \geq H(Y|X) \; R_X + R_Y \geq H(X,Y) \tag{11}$$

where R_X and R_Y the achievable rates, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies and $H(X,Y)$ is the joint entropy of X and Y gallery and probe feature vectors, respectively.

The resulting codewords S_{gait} and $S_{activity}$ comprise the biometric templates of the suggested modalities and are stored to the database of the system. Thus, if the database of the biometric system is attacked, the attacker can not access the original raw biometric data or their corresponding features but only S_{gait} and $S_{activity}$, which are not capable of revealing sensitive information about the users.

Similarly the gait and activity-related feature vectors F'_{gait} and $F'_{activity}$ are extracted and quantized at the authentication stage. Subsequently, the syndromes S_{gait} and $S_{activity}$ which correspond to the claimed ID are retrieved from the database and are fed to the *LDPC*

decoders. A similar multimodal approach is thoroughly described in (Argyropoulos et al. (2009)). Thereby, two biometric traits, i.e. face characteristics and face, have been combined via concatenation of their feature vectors. Specifically, once the first modality was successfully decoded, the decoded feature vector was concatenated to probe feature vector of the second modality. The full feature vector was fed to a second decoder. Thus, enhanced security was offered, since the second decoder requires that both feature vector resembles the gallery input. In the proposed approach, the system deals with two behavioural biometric traits separately, as far as the *LDPC* algorithm is regarded. However, it must be noted that the two biometric templates in the proposed scheme are not secured independently from each other.

The basic guidelines of the *LDPC* encoding/decoding scheme will be presented below in short, in order to provide a self-consistent paper.

Given the unimodal protection scheme had been used for every biometric modality independently the rate required to code each feature vector. This in turn would affect the size of the templates and the performance of the system.

Even if liveness detection is out of the scope of the paper, the multimodal framework provides tools to guarantee that even if the user is wearing a mask, in order to fake the system, he/she should also mimic the gait modality. Thus, we are not proposing a solution that will support liveness detection at the sensor level, however, we can support security at the signal level due to the multimodal nature of the proposed framework.

Initially, at the enrolment stage, the biometric signatures of an individual for *gait* and *activityrelated* modalities are obtained. The extracted features form the vector $\mathbf{F_i} = [f_1, \ldots, f_k]$, whereby $i \in gait, activity\ related$ and $f_i \in \mathbb{R}^k$. The feature vector $\mathbf{F_i}$ is then uniformly transformed from the continuous to the discrete domain of 2^L levels through the function $u : \mathbb{R}^k \to \mathbb{Q}^k$ whereby $\mathbb{Q} = 0, 1 \ldots, l-1$. Each one of the resulting vectors $\mathbf{q} = u(\mathbf{F_i})$ is fed to the Slepian.Wolf encoder, which performs the mapping $e : \mathbb{Q}^k \to \mathbb{C}^n$ where $C = \{0, 1\}$ outputs the codeword $\mathbf{c} = e(\mathbf{q})$, $\mathbf{c} \in \mathbb{C}^n$.

As already mentioned, herein the Slepian-Wolf algorithm has been implemented by a systematic *LDPC* encoder (Gallager (1963)) (see Figure 8). *LDPC* codes were selected due to their excellent error detecting and correcting capabilities. They also provide near-capacity performance over a large range of channels while simultaneously admitting implementable decoders. An *LDPC* code (n, k) is a linear block code of codeword length n and information block length k which is defined by a sparse $(n-k) \times n$ parity matrix H, where $n-k$ denotes the parity bits produced by the encoder. The code rate is defined as $r = k/n$. A code is a systematic code if every codeword consists of the original $k - bit$ information vector followed by $(n-k)$ parity-bits. In the proposed system, the joint bit-plane encoding scheme of (Girod et al. (2005)) was employed to avoid encoding and storing the L bit-planes of the vector $\mathbf{q}$ separately.
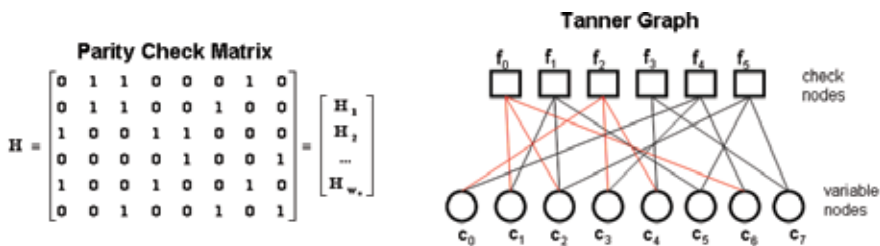

Fig. 8. Encoding via a Parity Check Matrix.

Subsequently, the k systematic bits of the codeword ci are discarded and only the syndrome $\mathbf{s}$, that is the $n - k$ parity bits of the codeword $\mathbf{c}$, is stored to the biometric database. Thus, the biometric templates of an enrolled user consist of the syndromes $\mathbf{s} = [c_{k+1}...c_n]$, $\mathbf{s} \in \mathbb{C}^{(n-k)}$, and their size is $n - k$. It must be stressed that the rate of the two *LDPC* encoders is different because the statistical properties of the two modalities are different.

Similarly to the enrollment procedure the biometric feature vector $\mathbf{F}'_i$ is obtained quantized at the authentication stage. This, together with encoded syndrome $\mathbf{s}_i^{encoded}$ are fed to the *LDPC* decoder. The decoding function $d : \mathbb{C}^{(n-k)} \times \mathbb{R}^k \rightarrow \mathbb{Q}^k$ combines $\mathbf{F}_i$ with the corresponding syndromes which are retrieved from the biometric database and correspond to the claimed identity I. The decoder employs belief-propagation (Ryan (n.d.)) to decode the received codewords.

If the errors introduced in the side information with respect to the originally encoded signal are within the error correcting capabilities of the channel decoder then the correct codeword is output after an experimentally set (N_c=30) number of iterations and the transaction is considered as a client transaction. To detect whether a codeword is correctly decoded we add 16 Cyclic Redundancy Check (*CRC*) bits at the beginning of the feature vector $\mathbf{F}_i$. By examining these bits the integrity of the original data is detected. If the codeword is correctly decoded, then the transaction is considered as genuine. Otherwise, if the decoder can not decode the codeword ($N_{iter} \geq N_c$) a special symbol $\varnothing$ is output and the transaction is considered as an impostor transaction.

From the above, it is obvious that the level of security and the performance of the system significantly bases on the number of the parity bits in syndrome $\mathbf{s}_i$, apart from the error correcting performance of the channel code.

On the one hand, a channel code with low code rate exhibits high error correcting capabilities, which results in the decoding of very noisy signals. This means, that the channel decoder will be able to decode the codeword even if the noise in the biometric signal has been induced by impostors. Additionally, will consist of many bits and will be more difficult to forge. On the other hand, channel codes of high code rate exhibit limited error-correcting capabilities and reduce the security of the system since the parity bits produced by the channel encoder consist of a few bits. Thus, the design of an effective biometric system based on the channel codes involves the careful selection of the channel code rate to achieve the optimal trade-off between performance and security. In this respect, a method for further securing the syndrome $\mathbf{s}_i$ is proposed in the following section (4.2). Thus, both the security of a long syndrome is preserved, while improved performance is provided.

4.2 Encryption scheme

The second phase of the security template algorithm, that is implemented via an encryption algorithm ("*Keygenerator*" box in Figure 7) has a dual mission. On the one hand, it further ensures the security of the stored biometric syndromes $\mathbf{S}_{gait}$ and $\mathbf{S}_{activity}$ (see Section 4.1) and on the other hand, it provides a novel method for fusing static physiological information with dynamic behavioural traits. An interesting novelty introduced by the specific methodology is that the user is no longer obliged to memorize a pin, in order to secure his data. On the contrary, the personal password is automatically extracted from a series of N_b soft biometric features. Thus, the password can neither be stolen nor copied. The utilized methodology is presented below.

In the current implementation of the proposed framework $N_b = 2$ soft biometric characteristics have been included. However, the framework can be easily extended to

any arbitrary number of soft biometric features, depending on the utilized modalities. In particular, the *Height* and the *StrideLength* (see Figure 9) of the user have been utilized hereby according to the following extendable methodology.
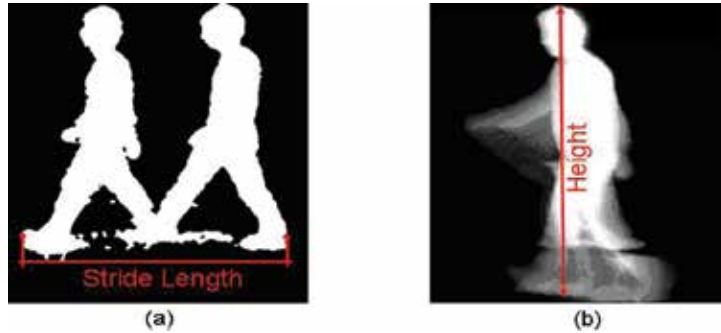
Fig. 9. Stride length (left) and Height (Right) of a user drawn on his/her Gait Energy Image (GEI).

It has been experimentally noticed that the measurement regarding the user's *Stride* are much more noisy than the ones for his/her *Height*. Thus, the ratio of $\frac{Height}{Stride}$ has been preferred over pure *Stride* scores, in order to provide a more uniform score distribution for the second soft-biometric trait.

First, a two dimensional hash table is formed, whereby its dimension is limited by the minimum and maximum expected value of each soft biometric trait, as illustrated in Figure 10/left. The resolution f_s^{height} and f_s^{stride} of the grid in each dimension respectively is scalable, in order to be optimally adjusted to the needs of each implementation of the framework (see Section 6). Thereafter, a unique biometric key is estimated for each cell on the grid (or cube or hypercube in the case of $N_b \geq 2$), according to the corresponding Soft Biometric values. Thus, we can write for the general case of N_b available soft biometric traits

$$Key(n_1, n_2, \ldots, n_{N_b}) = \frac{\sum_{i=1}^{N_b} n_i}{N_b} \tag{12}$$

whereby n_i stands for the index of the hash table (see Figure 10/left) and is calculated as $n_i = int(\frac{v_i}{f_s^i})$. v_i stands for the extracted value of the ith Soft Biometric trait.

In this context, it is expected that the same user will always obtain the same biometric key, since his soft biometric properties will always assign his identity with the same hypercube in the grid.

In the following, the syndromes S_i of the ith modality are encrypted using the Rijndael implementation (Daemen & Rijmen (1999)) of the Advanced Encryption Standard (*AES*). Specifically, the $128 - bit$ extracted key is used to shuffle the syndrome bits. Simple zero-padding technique is performed on the syndrome bits vector, in the cases where their number is not a whole multiplier of 2^7 bits. Similarly, a $256 - bit$ key could have been extracted, however it has been experimentally seen that it offers a bad trade-off between computational resources and security improvement.

In this respect, the biometric key is used to shuffle/deshuffle the syndromes for the claimed ID in the enrollment/authentication phase of the biometric system, respectively. It is easy to understand that most probably an impostor would be assigned to a different cell on the grid, given his different soft biometric characteristics with respect to the claimed ID. Thus, the
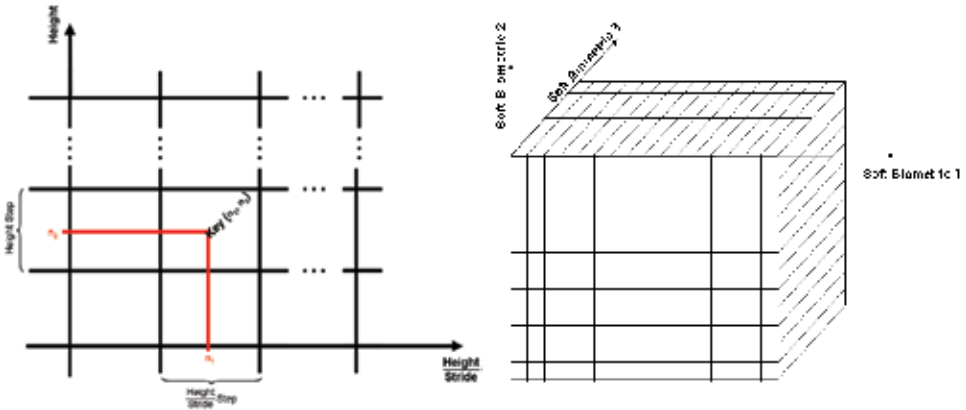
Fig. 10. 2D Soft Biometric Grid (Left) - 3D Soft Biometric Grid (Right)

requested syndrome bits will be wrongly decrypted and the applied dynamic biometric trait will never match the decoded one (see Section 4.1).

5. Score fusion

In order to provide an overall distance score between the user requesting access and the corresponding claimed ID, a fusion of the partial matching distances for each modality has to be performed. The fusion approach that has been utilized for the current biometric system is based on score level fusion. Thus, the optimal fusion score, that would combine unequally amounts of information from each RP is defined as follows

$$S_{tot} = \mathbf{W} \cdot \mathbf{S} = \sum_{j=1}^{N} w_j s_j = w_1 s_1 + w_2 s_2 + \dots + w_N s_N \tag{13}$$

whereby $\mathbf{W}$ is the weight coefficient matrix with N w_j elements and $\mathbf{S_j}$ the score matrix having as elements the corresponding partial matching distances s_j.

In this respect, the most common problem that has to be solved in a score-level fusion approach is the optimal estimation of matrix $\mathbf{W}$. Given the general structure of a multimodal biometric system, it is expected that the authentication capacity would be higher for some modalities than for some others. Thus, a rational way for defining the partial weight coefficients w_j for each modality is to assign a value proportional to their overall authentication performance, as follows:

$$w_j = 1 - \frac{EER_j}{\sum_{j=1}^{N} EER_j} \tag{14}$$

where EER_j stands for the Equal Error Rate score for the jth modality.
For the current bi-modal ($N = 2$) biometric system, the values for each w_j are defined as:

$$w_1 = 1 - \frac{EER_1}{EER_1 + EER_2} \; ; w_2 = 1 - \frac{EER_2}{EER_1 + EER_2} \tag{15}$$

In order to provide normalized scores in the range of values for each modality, all scores have been normalized to a common basis according to the following equation:

$$s^{norm} = \left(\frac{0.5}{T_L}\right)e^{\left(-\frac{s}{s^{max}}\right)} \tag{16}$$

where s_k^{norm} is the normalized score value, s_k the non-normalized score, s_k^{max} the maximum possible score value and T_k an experimentally set threshold for the kth modality; $k \in \{RIT, CIT, DTW\}$.

6. Results

The current Section starts with a short description of the database on which the experiments have been carried out. In the following, the identification and authentication results of the presented framework implementation are exhibited and qualitatively evaluated. A short discussion about the proposed framework is also included.

6.1 Database

The evaluation of the proposed secure multimodal biometric framework has been performed on the proprietary ACTIBIO-dataset (*ACTIBIO ICT STREP* (2008)). The current annotated database was captured in an ambient intelligence indoor environment and consists of 29 subjects, performing a series of everyday workplace & office activities. The workplace recordings include 29 subjects walking in various paths of *6m*, while being recorded by a stereoscopic camera that was placed *2.5m* away from the walking path and lateral to the walking direction. In order to test the permanence of the biometric features, the recordings have been repeated in a second session, few months after the first one.

Regarding the office recordings, the same 29 subjects have been recorded in an ambient intelligence (*AmI*) indoor environment, while they have been performing a series of everyday office activities with no special protocol, such as a phone conversation, typing, talking to a microphone panel and drinking water. Each subject repeated the same sequence of activities 8 times in total, split in two sessions while a manual annotation of the database has followed. Among the five cameras that have been recording the users from different view-angles, only the recordings from a frontal stereoscopic camera have been used for the current work.

Within the current work, the traits of the aforementioned modalities have been combined, in order to create 29 user-specific multimodal signatures. In this respect, each subject has been registered to the system (*gallery signatures*) by using his gait biometric signature together with his behavioural response during a phone conversation. Despite the fact that the current dataset does also include complicated gait scenarios, whereby the subject is carrying a bag or a coat, the simplest version has been utilized within the presented work. Similarly, only the "Phone Conversation" activity has been used from the office environment. Similarly, the recordings from each modality for a different repetition have been combined in order to form the *probe signatures* for the system.

6.2 Authentication & verification results

As it has already been stressed out the major contribution of the current framework is that it allows higher level of security of the biometric templates stored in the database, while higher recognition performance is simultaneously provide via the encoding of soft biometrics. The improved level of security can be easily noticed, when considering that the information stored in the database is encrypted. In this respect, not data can be retrieved from the

database without providing the correct key. Further, even if this encryption step is somehow bypassed, the obtained data remain still of no use to the attacker, since they reveal no biometric information as explained in Section 4.1.

Moreover, in order to illustrate the advances performed in the recognition performance via indirect fusion/encoding of the soft biometric traits with the dynamic ones, the initial recognition capabilities of the utilized traits is shown in Figure 11. In the same Figures the reader can notice a slight degradation in the recognition performance of the activity related modality, when the templates that are stored in the database are secured via the *LDPC* encoding algorithm (Section 4.1). Contrary to the $1D$ feature vector of the gait modality, activity related feature vectors are much more complex. Thus, a degradation in the authentication performance is more likely due to the noisy errors at the decoding. Moreover, a degradation is expected, since this is the trade off for adding enhanced security in the biometric system. Specifically, such a deterioration have been mainly caused by the unintended reconstruction of an impostor's feature vector, so that it resembles a genuine user.
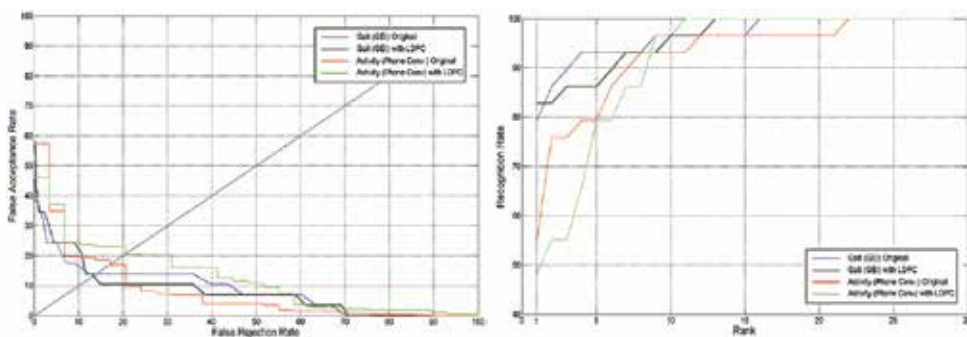
Fig. 11. EER Scores (left) and Identification performance (right) of the proposed modalities prior to encryption.

As it has been mentioned in Section 4.2 the current framework allows a scalable resolution of the hash table that is used for encryption, so that optimal performance of the system is achieved, given different soft biometrics. In this respect, the *Optimal Functional Point (OFP)* of the current system has been set according to the results illustrated in Figures 12. Specifically, one can notice that an intense degradation of the system's recognition performance for high resolution values ($\equiv$ large number of available keys in Hash Matrix of Figure 10left). This is caused by the noisy measurements of the soft biometric trait in different repetitions. For instance, let us assume a user that has been registered to the system with a $v_{Height} = 1.79$ and $v_{Stride} = 1.62$. He/she would be assigned the key $K(17,14)$. A noisy measurement of his soft biometrics at the authentication stage might result that his stored syndrome **s** was attempted to be decoded by a different key $Key(n_1^{probe}, n_2^{probe}) \neq K(17,14)$. Thus, the decrypting would never be successful and the recognition would fail. The reason for which the *EER* scores of the activity related modality exhibits more fluctuations than the one of the gait bases on the following fact: The soft biometric measurements of some impostors in the authentication stage did not only lie within the same hash bin as the client, but also their activity related traits managed to be decoded via *LDPC* using the syndrome of the claimed user's ID.

In this concept, it can be concluded that high resolution values, which refer to a big number of bins in the hash table, are intolerant to noisy measurements. On the other hand, small resolution values may result to the fact that all subjects are assigned to the same key K and
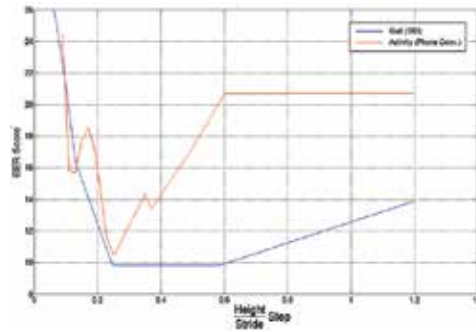
Fig. 12. *EER* scores of the proposed modalities as a function of the $\frac{Height}{Stride}$ step.

thus, the encryption scheme would have no meaning. On the contrary, there is always an functional point, whereby the recognition performance of the system is optimal.

Moreover, the reader can notice in Figure 12 that for the $\frac{Height}{Stride} step = 0.25$ both modalities achieve their authentication performance. Thus, this value can be considered the system's optimal functional point for a given *Height* resolution in the hash table, experimentally set at f_s^{Height}.
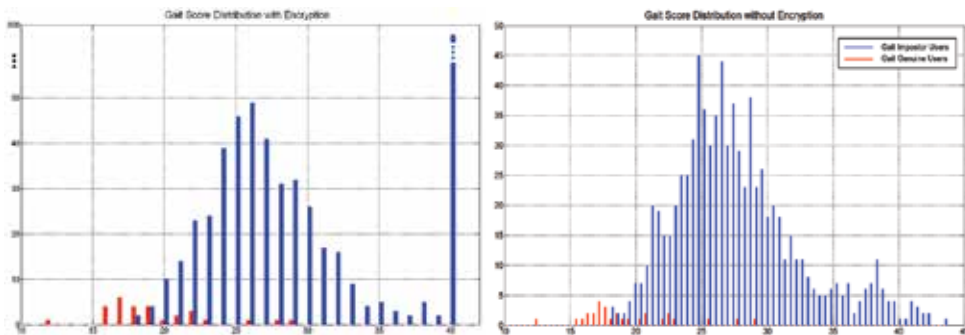
Fig. 13. Client/Impostor Distributions for the gait modality at the optimal functional point(left) and without Encryption(right) via the Biometric key.

Although there seems to be only small changes in the *EER* scores for small resolution values, the distribution of the genuine/impostor scores significantly changes (see Figure 13).
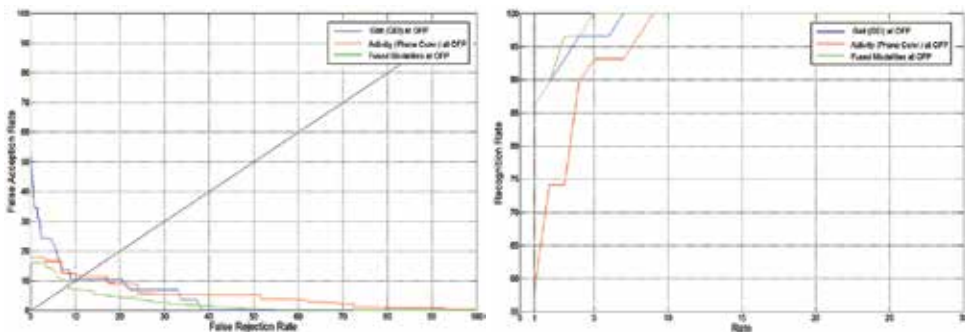
Fig. 14. EER Scores (left) and Identification performance (right) of the proposed system.

Given the optimal functional point of the encryption system, optimal fusion of the soft biometrics with the dynamic traits has been also achieved. In this respect, the current test case of the proposed framework has been evaluated in terms of its overall recognition performance, by performing fusion between the two utilized behavioral biometrics (Sections 3.1 & 3.2) as described in Section 5. The derived optimal recognition performance of the bimodal biometric system system is illustrated in Figures 14, in terms of both authentication and identification capacity.

Concluding, it must be noted that the potential of the proposed framework in terms of recognition performance is significantly high. Given a larger number of soft biometrics, an almost $1 - 1$ proportion for keys-users can be achieved, which would lead to further decreasing of the recognition error.

7. Conclusions

Summarizing, the advantages of the proposed method in terms of security and impact on matching accuracy for recognition purposes have been thoroughly analyzed and discussed. The performance of the proposed method is assessed in the context of ACTIBIO, an EU Specific Targeted Research Project, where activity-related and gait biometrics are employed in an unobtrusive application scenario for human recognition. The experimental evaluation on a multimodal biometric database demonstrates the validity of the proposed framework. Most important, the dual scope of the current framework has been illustrated. Specifically, the utilization of the encryption algorithm does not only provide enhanced template security; it does also provide indirect fusion with soft biometric characteristics and thus it improves the recognition potential. Finally, the proposed user-specific biometric key, which exclusively depends on the user's biometry, increases the level of unobtrusiveness of the system, since the user is not obliged anymore to memorize pins or to carry ID cards.

8. Acknowledgments

9. References

ACTIBIO ICT STREP (2008).

Álvarez, F. H., Encinas, L. H. & Ávila, C. S. (2009). Biometric Fuzzy Extractor Scheme for Iris Templates, *World Congress in Computer Science, Computer Engineering and Applied Computing*.

Argyropoulos, S., Tzovaras, D., Ioannidis, D. & Strintzis, M. G. (2009). A Channel Coding Approach for Human Authentication From Gait Sequences, *IEEE Trans.Inf. Forensics Security.* 24(3): 428–440.

Bobick, A. & Davis, J. (2001). The recognition of human movement using temporal templates, *IEEE Trans. Pattern Anal. Mach. Intell.* 23(3): 257–267.

Boulgouris, N., Plataniotis, K. & Hatzinakos, D. (2004). Gait recognition using dynamic time warping, *IEEE 6th Workshop on Multimedia Signal Processing*, Siena, pp. 263–266.

Chang, K. I., Bowyer, K. W. & Flynn, P. J. (2005). An evaluation of multimodal 2D+3D face biometrics., *IEEE transactions on pattern analysis and machine intelligence* 27(4): 619–24.

Cohen, G. & Zemor, G. (2004). Generalized coset schemes for the wire-tap channel: application to biometrics, *IEEE International Carnahan Conference on Security Technology (ICCST)Symposium on Information Theory*, Chicago, IL, p. 46.

Cucchiara, R., Grana, C., Piccardi, M., Prati, A. & Sirotti, S. (2001). Improving shadow suppression in moving object detection with HSV color information, *Intelligent Transportation Systems* pp. 334–339.

Daemen, J. & Rijmen, V. (1999). The Rijndael Block Cipher.

Davida, G. I., Frankel, Y. & Matt, B. J. (1998). On enabling secure applications through off-line biometric identification, *IEEE Symp. Security and Privacy*, Oakland, CA, pp. 148–157.

Draper, S. C., Khisti, A., Martinian, E., Vetro, A. & Yedidia, J. (2007). Using distributed source coding to secure fingerprint biometrics, *Int. Conf. Acoustics, Speech and Signal Processing*, Honolulu, HI, pp. 129–132.

Drosou, A., Ioannidis, D., Moustakas, K. & Tzovaras, D. (2010). Unobtrusive Behavioural and Activity Related Multimodal Biometrics: The ACTIBIO Authentication Concept, *The Scientific World - Special Issue on: Biometrics Applications: Technology, Ethics and Health Hazards* p. accepted for publication.

Drosou, A., Moustakas, K., Ioannidis, D. & Tzovaras, D. (2010). On the potential of activity-related recognition, *The International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2010)*.

Drosou, A., Moustakas, K. & Tzovaras, D. (2010). Event-based unobtrusive authentication using multi-view image sequences, *Proc. of ACM Multimedia/Artemis Workshop ARTEMISŠ10*, Florence, pp. 69 – 74.

Fairhurst, M., Deravi, F., Mavity, N., George, J. & Sirlantzis, K. (2003). *Intelligent Management of Multimodal Biometric Transactions*, Springer Berlin-Heidelberg.

Gallager, R. (1963). *Low-Density Parity-Check Codes*, MIT Press, Cambridge, MA.

Girod, B., Aaron, A. M., Rane, S. & Rebollo-Monedero, D. (2005). Distributed video coding, *Proc. IEEE* 93: 71–89.

Goffredo, M., Bouchrika, I., Carter, J. N. & Nixon, M. S. (2009a). Performance analysis for automated gait extraction and recognition in multi-camera surveillance, *Multimedia Tools and Applications* .

Goffredo, M., Bouchrika, I., Carter, J. N. & Nixon, M. S. (2009b). Self-Calibrating View-Invariant Gait Biometrics., *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society* .

Gomez, G. & Morales, E. F. (2002). Automatic feature construction and a simple rule induction algorithm for skin detection, *Proc. of the ICML Workshop on Machine Learning in Computer Vision (MLCV)*, pp. 31–38.

Hadid, A., Pietikäinen, M. & Li2, S. Z. (2007). *Learning Personal Specific Facial Dynamics for Face Recognition from Videos*, Springer Berlin / Heidelberg, pp. 1–15.

Han, X., Liu, J., Li, L. & Wang, Z. (2006). Gait recognition considering directions of walking, *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, pp. 1–5.

Ioannidis, D., Tzovaras, D., Damousis, I. G., Argyropoulos, S. & Moustakas, K. (2007). Gait Recognition Using Compact Feature Extraction Transforms and Depth Information, *IEEE Trans.Inf. Forensics Security.* 2(3): 623–630.

Jain, A. K., Ross, A. & Prabhakar, S. (2004). An Introduction to Biometric Recognition, *IEEE Trans. Circuits Syst. Video Technol.* 14(1): 4–20.

Jain, A., Nandakumar, K. & Ross, A. (2005). Score normalization in multimodal biometric systems, *Pattern recognition* 38(12): 2270–2285.

Jayadevan, R., Kolhe, S. R. & Patil, P. M. (2009). Dynamic Time Warping based Static Hand Printed Signature Verification, *ournal of Pattern Recognition Research* 4(1): 52–65.

Juels, A. & Sudan, M. (2006). A fuzzy vault scheme, *Designs Codes Cryptography* 38(2): 237–257.

Junker, H., Ward, J., Lukowicz, P. & Tröster, G. (2004). User Activity Related Data Sets for Context Recognition, *Proc. Workshop on 'Benchmarks and a Database for Context Recognition'*.

Kale, A., Cuntoor, N. & Chellappa, R. (2002). A framework for activity-specific human identification, *IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 4, pp. 3660–3663.

Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N. P., Roy-Chowdhury, A. K., Kruger, V. & Chellappa, R. (n.d.). Identification of Humans Using Gait, *IEEE Trans. Image Processing* 13: 1163–1173.

Kumar, A., Kanhangad, V. & Zhang, D. (2010). A New Framework for Adaptive Multimodal Biometrics Management, *IEEE Trans.Inf. Forensics Security.* 5(1): 92 – 102.

Liu, X. & Chen, T. (2003). Video-based face recognition using adaptive hidden markov models, *IEEE Proc. Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* 1: 340–345.

Maltoni, D., Jain, A. & Prabhakar, S. (2009). *Handbook of fingerprint recognition*, 2nd edn, Springer Professional Computing.

Martinian, E., Yekhanin, S. & Yedidia, J. (2005). Secure biometrics via syndromes, *43rd Annual Allerton Conf. on Communications, Control, and Computing*, Monticelo, IL.

Miguel-Hurtado, O., Mengibar-Pozo, L. & Pacut, A. (2008). A new algorithm for signature verification system based on DTW and GMM, *IEEE International Carnahan Conference on Security Technology ICCST* 42: 206–213.

Pradhan, S. S. & Ramchandran, K. (2003). Distributed source coding using syndromes (discus): Design and construction, *IEEE Trans. Inf. Theory* 49(3): 626–643.

Qazi, F. A. (2004). A survey of biometric authentication systems, *Security and Management* pp. 61–67.

Ramesh, D. C. V. & Meer, P. (2000). Real-Time Tracking of Non-Rigid Objects Using Mean Shift, *IEEE Proc. Computer Vision and Pattern Recognition 2007 (CVPR)*, Vol. 2, pp. 142–149.

Ryan, M. G. (n.d.). Visual Target Tracking.

Sakoe, H. & Chiba, S. (1990). Dynamic programming algorithm optimization for spoken word recognition, *Readings in speech recognition* .

Sim, T., Zhang, S., Janakiraman, R. & Kumar, S. (2007). Continuous Verification Using Multimodal Biometrics, *IEEE Trans. Pattern Anal. Mach. Intell.* 29(4): 687 – 700.

Slepian, J. D. & Wolf, J. K. (1973). Noiseless coding of correlated information sources, *IEEE Trans. Inf. Theory* 19: 471–480.

Stallings, W. (2006). *Cryptography and Network Security: Principles and Practices.*, Prentice-Hall: Upper Saddle River, NJ.

Sun, Z. & Tan, T. (2009). Ordinal Measures for Iris Recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 31(12): 2211.

Uludag, U., Pankanti, S., Prabhakar, S. & Jain, A. K. (2004). Biometric Cryptosystems: Issues and Challenges, *Proceedings of the IEEE* 92(6): 948–960.

Viola, P. & Jones, M. J. (2004). Robust Real-Time Face Detection, *International Journal of Computer Vision* 57(2): 137–154.

Wadayama, T. (2005). An authentication scheme based on a low-density parity check matrix, *Int. Symp. Information Theory*, pp. 2266–2269.

Xiao, Q. (2005). Security issues in biometric authentication, *Information Assurance Workshop, IAW 2005* pp. 8–13.

Yao-Chung, L., Varodayan, D. & Girod, B. (2007). Image authentication based on distributed source coding, *Int. Conf. on Image Processing*, San Antonio, TX, pp. 5–8.

Yu, C., Cheng, H., Cheng, C. & Fan, K.-C. (2010). Efficient Human Action and Gait Analysis Using Multiresolution Motion Energy Histogram, *Journal on Advances in Signal Processing (EURASIP)* p. 13.

Biometric Encryption Using Co-Z Divisor Addition Formulae in Weighted Representation of Jacobean Genus 2 Hyperelliptic Curves over Prime Fields

Robert Brumnik[1], Vladislav Kovtun[2], Sergii Kavun[3] and Iztok Podbregar[4]

[1]University of Maribor, Faculty of Criminal Justice and Security,
[2]National Aviation University, Kiev,
[3]National University of Economic, Kharkov,
[4]University of Maribor, Faculty of Criminal Justice and Security,
[1,4]Slovenia
[2,3]Ukraine

1. Introduction

Security in biometry is a prime concern of modern society. Identity theft is a growing problem in today's interconnected world. To ensure a safe and secure environment biometrics is used today in many commercial, government and forensic applications. To ensure a high level of security of a biometric system we use Cryptographic algorithms. Though a number of bio-crypto algorithms have been proposed, they have limited practical applicability due to the trade-off between recognition performance and security of the template. Overall, these are very secure, however, they do have a weak point in terms of the procedure and storage of the crypto keys.

Biometric authentication systems should have many exploitable crypto secure points that can be used to compromise the identification system within the optimization process. Biometric encryption with Jacobean Genus 2 Hyperelliptic curves is a security scheme that combines strong cryptographic algorithms with biometric authentication to provide enhanced security. This paper discusses the simple implementation Co-Z divisor addition formulae in a weighted representation of encryption systems for biometric software application.

In this article the authors show a newly developed Co-Z approach to divisor scalar multiplication in Jacobean of Genus 2 Hyperelliptic curves over fields with odd characteristics in weighted coordinates for application in biometric-based authentication systems. We assess the performance of these biometric generation algorithms using Co-Z divisor. This approach is based upon improved additional formulae of weight 2 divisors in weighted divisor representation which, in the most frequent cases are well suited to exponentiation algorithms based on Euclidean addition chains.

2. Cryptographic applications

The progress of civilization and the constantly increasing role of various technologies in human day-to-day activities has lead to the permanent development of access control

security systems for information and physical objects. A number of researchers have studied the interaction between biometrics and cryptogra-phy, two potentially complementary security technologies (Hao et al., 2005). For such systems, the importance of authentication and identification subsystems is undeniable. One of the approaches to the implementation of authentication and identification subsystems is biometric systems.

The process of creating biometric systems of authentication and identification has solved a large variety of problems among which the assurance of confidentiality and integrity of biometric information, which is by no means unimportant. As a rule, cryptographic transformations are used – encryption and electronic digital signature (EDS). The analysis, carried out by the authors, has shown the prospectivity of algebraic curves theory for the implementation of cryptographic transformations.

2.1 Elliptic curves

The usage of elliptic curves (EC) for cryptographic purposes was first suggested by (Koblitz, 1987) and (Miller, 1985).

Basiri et al. (2004) present two algorithms for the arithmetic of cubic curves with a totally ramified prime at infinity. The first algorithm, inspired by Cantor's reduction for hyperelliptic curves, is easily implemented with a few lines of code, making use of a polynomial arithmetic package.

In his research, Koblitz (1989) has proved the possibility of usage of hyperelliptic curves (HEC) in cryptographic applications. Their difference from EC is that for HEC the group (Jacobian) of more complex structures should be considered – divisors instead of curve points. It is known that HEC have a variety of advantages over EC: being a richer source of the Abelian groups (Koblitz, 1989; Menezes and Wu, 1998) they also use a base field of a smaller size (from the Abelian group, the size of which is defined by product of the base field size by a curve genus).

2.2 Hyperelliptic curves

Hyperelliptic curve cryptosystems (HCC for short) is a generalization of ECC. It has been drawing the attention of more and more researchers in recent years. The problem of how to decrease the amount of addition and scalar multiplication on the Jacobians of hyperelliptic curves so that the implementation speed can be improved is very important for the practical use of HCC (Zhang et al., 2001).

During the time in which HEC cryptosystems were restricted to academic interest only, they had no practical application due to the high complexity of software and hardware implementation, low performance, absence of in-depth studies in the field of cryptoanalysis of such cryptosystems and the absence of comprehensible algorithms of cryptosystem parameters generation. Active review research of papers (Koblitz, 1989; Menezes et al., 1998; Lange 2002c; Matsuo et al., 2001; Miyamoto et al., 2002; Takahashi et al., 2002; Sugizaki et al. 2002, etc) has allowed us to overcome the majority of the described difficulties. The authors of publications, offer a variety of approaches which increase the performance of HEC cryptosystems essentially, and bring them close to the EC cryptosystems.

2.3 Genus 2 HEC cryptosystems over prime fields

The given research is devoted to the development of the approach (Cohen et al., 1998) and to the improved efficiency of genus 2 HEC cryptosystems over prime fields.

Scalar multiplication operation is used in encryption, decryption and electronic digital signature based on HEC. These computations are relatively expensive when implemented on low-power devices. A widely used standard method is the left-to-right binary method. In accordance with (Koblitz, 1989; Menezes et al., 1998; Lange 2002c; Matsuo et al., 2001; Miyamoto et al., 2002; Takahashi et al., 2002; Sugizaki et al. 2002; Lange, 2002; Kovtun and Zbitnev, 2004) the power consumption traces of divisor addition and doubling are not the same, they can easily be distinguished between these operations and derive the bit of scalar. The first method proposed, with resistance to the side channel attacks (SCA), is Coron's dummy addition (CDA) (Coron, 1999).

Several SCA-resistant scalar multiplication algorithms have been proposed that are faster than the CDA method. There are three basic approaches with SCA resistance:

- The first is to use indistinguishable additions and doubling algorithms in scalar multiplication (Clavier and Joye, 2001). For example, Jacobi form and Hesse form of EC. However, this requires specially chosen EC and HEC curves and does not work for the standardized curves.
- The second is the double-and-always-add approach. The CDA method is the simplest algorithm of this type. In paper (Okeya and Sakuri, 2000), the authors proposed to use the Montgomery form of EC and extended it to general curves (Brier and Joye, 2002).
- The third approach is to use a special addition chain with a sequence of additions and doublings that does not depend on the bit information of the scalar (Izu and Takagi, 2002).

In this paper, we are interested in scalar multiplication algorithms that do not require specially chosen curves and based on approach (Meloni, 2007) for genus 2 HEC over prime fields.

3. Biometric cryptosystems

In a generic cryptographic system the user authentication is possession based. That is, possession of the decrypting key is a sufficient evidence to establish user authenticity. Because cryptographic keys are long and random, (e.g., 128 bits for the advanced encryption standard (AES) (NIST, 2008; Stallings, 2003), they are difficult to memorize. As a result, the cryptographic keys are stored somewhere (for example, on a computer or a smart card) and released based on some alternative authentication (e.g., password) mechanism, that is, upon assuring that they are being released to the authorized users only. Most passwords are so simple that they can be easily guessed (especially based on social engineering methods) or broken by simple dictionary attacks (Klein, 1990).

It is not surprising that the most commonly used password is the word "password"! Thus, the multimedia protected by the cryptographic algorithm is only as secure as the passwords (weakest link) used for user authentication that release the correct decrypting key(s). Simple passwords are easy to crack and, thus, compromise security; complex passwords are difficult to remember and, thus, are expensive to maintain. Users also have the tendency to write down complex passwords in easily accessible locations. Further, most people use the same password across different applications and, thus, if a single password is compromised, it may open many doors. Finally, passwords are unable to provide nonrepudiation; that is, when a password is shared with a friend, there is no way to know who the actual user is. This may eliminate the feasibility of countermeasures such as holding conniving legitimate users accountable in a court of law. Many of these limitations of the traditional passwords can be ameliorated by incorporation of better methods of user authentication. Biometric authentication (Jain et al., 1999; Maltoni et al., 2003) refers to verifying individuals based on

their physiological and behavioural characteristics such as face, fingerprint, hand geometry, iris, keystroke, signature, voice, etc. It is inherently more reliable than password-based authentication, as biometric characteristics cannot be lost or forgotten (cf. passwords being lost or forgotten); they are extremely difficult to copy, share, and distribute (cf. passwords being announced in hacker websites) and require the person being authenticated to be present.

A brief comparison of some of the biometric identifiers based on seven factors is provided (Wayman, 2001):

- universality (do all people have it?),
- distinctiveness (can people be distinguished based on an identifier?),
- permanence (how permanent is the identifier?), and
- collectability (how well can the identifier be captured and quantified?) are properties of biometric identifiers.
- performance (speed and accuracy), acceptability (willingness of people to use), and circumvention (foolproof) are attributes of biometric systems.

4. Background

Let's observe basic concepts of cryptosystems on HEC. More detailed information can be obtained from (Koblitz, 1989; Menezes and Wu, 1998).

Let K be a field and $\bar{K}$ be the algebraic closure of K. Hyperelliptic curve C of genus $g \geq 1$ over K is a set of points (u, v) that satisfy the equation:

$$C : v^2 + h(u)v = f(u), \ \ k[u, v] \tag{1}$$

and there are no solutions $(u, v) \in \bar{K} \times \bar{K}$ which simultaneously satisfy the equation (1) and the partial derivative equations with corresponding variables.

$$2u + h(u) = 0, \ h'(u)v - f'(u) = 0 \tag{2}$$

In case of genus 2 HEC, polynomials $f(u)$ and $h(u)$ will be represented as:

$$f(u) = u^5 + f_4 u^4 + f_3 u^3 + f_2 u^2 + f_1 u + f_0 \tag{3}$$

and

$$h(u) = h_2 u^2 + h_1 u + h_0, \ h_i, f_j \in K. \tag{4}$$

Divisor D is a formal sum of points in C:

$$D = \sum_{P \epsilon C} m_p P .. m_p \in Z \tag{5}$$

where only a finite number of the m_P are non-zero.

Divisor $D \in \mathbf{D}^0$ is a principal divisor, if $D = \mathrm{div}(R)$ for some rational function $R \in \bar{K}(C)^*$. The set of all principal divisors, denotes $P_C(\bar{K}) = \{\mathrm{div}(F) : F \in \bar{K}(C)\}$, in curve C over $\bar{K}$, moreover $P_C(\bar{K})$ is a subgroup of $\mathbf{D}^0$. Generally $P(C) = P_C(\bar{K})$ is called a group of principal divisors of curve C. The quotient group $J_C(\bar{K}) = \mathrm{div}_C^0(\bar{K}) / P_C(\bar{K})$ is called the Jacobian of the

curve C over $\bar{K}$. The quotient group $J(C) = \mathrm{div}^0(C)/P(C)$ is called Jacobian of the curve C.

Furthermore, we will operate with divisors in the Mumford representation (Menezes, 1998):

$$D = \left(x^2 + u_1 x + u_0, v_1 x + v_0\right), \ \deg v < \deg u \leq 2, \ u \mid f(u) - h(u)v - v^2$$

$$\text{where } \forall D_i \in J(C), \ weight(D_i) = 2, \ i = \overline{1,2} \tag{6}$$

The result $D_3 = D_1 + D_2$ will have a $weight(D_3) = 2$, which helps to avoid consideration of alternative addition methods for divisors of different weight and with intersecting support (containing intersecting sets of points) (Lange, 2003; Wollinger, 2004).

HECC uses a divisor scalar multiplication operation:

$$\underbrace{D + D + \cdots + D}_{k} = k \cdot D \tag{7}$$

At the intermediate computation phase of scalar divisor multiplication (scalar multiplier in binary notation) the binary algorithm performs the divisor addition and doubling operation. The addition and doubling algorithms use field $\mathbf{GF}(p)$ multiplicative inversion, which is the most computationally intensive and space critical operation. Projective divisor representation (Miyamoto et al., 2002; Takahashi, 2002; Lange, 2002c) is one of the most popular approaches which allows saving of a field inversion.

In her work Lange (2002c), suggested a weighted divisor representation, being the development of a projective approach.

In weighted representation, the divisor D we can present as Lange (2002c):

$$D = \left(x^2 + u_1 x + u_0, v_1 x + v_0\right) \tag{8}$$

which is of the form of:

$$D = \left[U_1, U_0, V_1, V_0, Z_1, Z_1^2, Z_2, Z_2^2\right] \tag{9}$$

while

$$D = \left(x^2 + U_1/Z_1^2 \, x + U_0/Z_1^2, V_1/Z_1^3 Z_2 x + V_0/Z_1^3 Z_2\right) \tag{10}$$

Note, that arithmetic in Jacobian genus 2 HEC in weighted representation (Lange, 2002c) is the most efficient (Kovtun and Zbitnev, 2004).

5. Co-Z approach

The Co-Z approach was first suggested by Meloni (2007), in order to increase the efficiency of scalar multiplication in EC over $\mathbf{GF}(p)$ with double-free addition chain for the resistance side channel attacks (SCA) (Goundar, 2008). It results in a fixed sequence of operations; hence attackers could not detect any information through SCA. Its principles lie in transformation of EC points in SCA resistant scalar point multiplication in projective and

modified Jacobi representation with the same denominator and further operation with points of identical Z-coordinates. Note that the Co-Z approach is applicable for algorithms, based on the Euclidian addition chains approach by Meloni (2007) and scalar in the Zeckendorf representation, in order to replace doublings by Fibonacci numbers computations refer to Algorithm A.1. Indeed the Fibonacci sequence is an optimal chain (Meloni, 2007).

The Zeckendorf number representation needs 44% more digits in comparison with the binary representation. For example a 80-bit integer will require around 115 Fibonacci digits. However, the density of 1's in this representation in lower (near 0.2764). This means that representing a 80-bits integer requires, an average 40 powers of 2 but only 32 Fibonacci numbers (near 115x0,2764).

More generally, for a n-bit integer, the classical double-and-add algorithm requires on average $1.5 \times n$ operations ($\frac{n}{2}\mathbf{A} + n\mathbf{D}$, where $\mathbf{A}$-addition operation and $\mathbf{D}$-doubling operation) and the Fibonacci-and-add requires $1.83 \times n$ operations ($1.44 \times n\mathbf{F} + 0.398 \times n\mathbf{A}$, where $\mathbf{A}$-addition step and $\mathbf{F}$-Fibonacci step). In other words, the Fibonacci-and-add algorithms A.1 require about 23% more operations (Table 1.). Note, that in paper (Lange, 2002c) the simplified version of HEC is used, such that $h(x) = 0$, $f_4 = 0$ (since replacement $y \mapsto y - h/2$ is admissible for the odd field characteristic and the replacement $x \mapsto x - f_4/5$ is admissible if $p \neq 5$), which allowed T. Lange to save 1 multiplication at the step A.2.7. However in this paper we will consider a more general case of curve with $\deg(h) = 2$, $h_i \in \mathbf{F}_2$, $i = \overline{0,2}$ $\deg(f) = 5$, $f_5 = 1$, $f_i \in \mathbf{GF}(p)$, $i = \overline{0,4}$.

Input: $D \in J_C$, $k = (d_l, ..., d_2)_Z$

Output: $[k]D \in J_C$

begin
   $(U, V) \leftarrow (D, D)$
   for $i = l - 1$ downto 2
      if $d_i = 1$ then $U \leftarrow U + D$ (add step)
      $(U, V) \leftarrow (U + V, U)$ (Fibonacci step)
   end
   return U
end

Table 1. Algorithm A.1 Fibonacci-and-add(k, P)

We can see, in Algorithm A.1 using the Addition and Fibonacci step for the adding divisors. In common case, addition of two reduced divisors in weighted coordinates can be represented by the Algorithm A.2 (Lange, 2002c) (Table 2).

Assuming that $Z_{11} = Z_{21} = Z_1$ and $Z_{12} = Z_{22} = Z_2$ for D_1 and D_2, which allows the transformation of algorithm A.2 into algorithm A.3. Apply the approach described by Meloni (2007) to the divisor addition algorithm suggested by Lange (2002c) and weighted divisor representation. Further circumscribe the derivation of expressions in different steps of A.3 (Table 3).

Input:	$\left[U_{11}, U_{10}, V_{11}, V_{10}, Z_1, Z_1^2, Z_2, Z_2^2\right]$,
	$\left[U_{21}, U_{20}, V_{21}, V_{20}, Z_1, Z_1^2, Z_2, Z_2^2\right]$
Output:	$\left[U_1', U_0', V_1', V_2', Z_1', Z_1'^2, Z_2', Z_2'^2\right] =$
	$\left[U_{11}, U_{10}, V_{11}, V_{10}, Z_1, Z_1^2, Z_2, Z_2^2\right] +$
	$+\left[U_{21}, U_{20}, V_{21}, V_{20}, Z_1, Z_1^2, Z_2, Z_2^2\right]$,
	$weight\left(D_1\right) = weight\left(D_2\right) = 2$

#	Expression	Cost
1	Precomputations: $z_{13} = Z_{11} \cdot Z_{12}$, $z_{23} = Z_{21} \cdot Z_{22}$, $z_{12} = z_{11} \cdot z_{13}$, $z_{22} = z_{21} \cdot z_{23}$, $\tilde{U}_{21} = z_{11} \cdot U_{21}$, $\tilde{U}_{20} = z_{11} \cdot U_{20}$, $\tilde{V}_{21} = z_{12} \cdot V_{21}$, $\tilde{V}_{20} = z_{12} \cdot V_{20}$	8M
2	Compute resultant r for u_1 and u_2: $y_1 = U_{11} \cdot z_{21} - \tilde{U}_{21}$, $y_2 = \tilde{U}_{20} - U_{10} \cdot z_{21}$, $y_3 = U_{11} \cdot y_1 + y_2 \cdot z_{11}$, $r = y_2 \cdot y_3 + y_1^2 \cdot U_{10}$, $Z_2' = Z_{11} \cdot Z_{21}$, $\tilde{Z}_2 = Z_{12} \cdot Z_{22}$, $Z_1 = Z_2'^2$, $\tilde{Z}_2 = \tilde{Z}_2 \cdot Z_1$, $\tilde{Z}_2 = \tilde{Z}_2 \cdot r$, $Z_2' = Z_2' \cdot \tilde{Z}_2$, $\tilde{Z}_2 = \tilde{Z}_2^2$, $z_2' = Z_2'^2$	4S, 11M
3	Compute almost inverse $inv = r/u_2 \bmod u_1$, $inv = inv_1 x + inv_0$: $inv_1 = y_1$, $inv_0 = y_3$	
4	Compute $s = \left(v_1 - v_2\right) inv \bmod u_1$, $s = s_1 x + s_0$: $w_0 = V_{10} \cdot z_{22} - \tilde{V}_{20}$, $w_1 = V_{11} \cdot z_{22} - \tilde{V}_{21}$, $w_2 = inv_0 \cdot w_0$, $w_3 = inv_1 \cdot w_1$, $s_0 = w_2 - U_{10} \cdot w_3$, $s_1 = \left(inv_0 + z_{11} \cdot inv_1\right) \cdot \left(w_0 + w_1\right) - w_2 - w_3 \cdot \left(z_{11} + U_{11}\right)$ If $s_1 = 0$ then consider special case	8M
5	Precomputations: $S_1 = s_1^2$, $S_0 = s_0 \cdot Z_1$, $Z_1' = s_1 \cdot Z_1$, $S_1 = Z_1' \cdot S_0$, $S_0 = S_0^2$, $R = r \cdot Z_1'$, $s_0 = s_0 \cdot Z_1'$, $s_1 = s_1 \cdot Z_1'$, $z_1' = Z_1'^2$	3S, 6M
6	Compute $l = s u_2$, $l = x^3 + l_2 x^2 + l_1 x + l_0$: $l_0 = s_0 \cdot \tilde{U}_{20}$, $l_2 = s_1 \cdot \tilde{U}_{21}$, $l_1 = \left(s_1 + s_0\right) \cdot \left(\tilde{U}_{21} + \tilde{U}_{20}\right) - l_0 - l_2$, $l_2 = l_2 + S$	3M
7	Compute $u' = \left(s\left(l + h + 2v_1\right) - k\right) u_1^{-1}$, $k = \left(f - v_1 h - v_1^2\right)/u_1$, $u' = x^2 + u_1' x + u_0'$: $V_1' = R \cdot \tilde{V}_{21}$, $U_1' = 2S - s_1 \cdot y_1 - z_2'$, $U_0' = S_0 + y_1 \cdot \left(S_1 \cdot \left(y_1 + \tilde{U}_{21}\right) - 2s_0\right) + y_2 \cdot s_1 + 2\tilde{V}_1' + \tilde{Z}_2 \cdot \left(y_1 + 2\tilde{U}_{21}\right)$	6M
8	Precomputations: $l_2 = l_2 - U_1'$, $w_0 = l_2 \cdot U_0'$, $w_1 = l_2 \cdot U_1'$	2M
9	Compute $v' \equiv -\left(h + s_1 l + v_2\right) \bmod u'$, $v' = v_1' x + v_0'$: $V_1' = w_1 - z_1' \cdot \left(l_1 + V_1' - U_0'\right)$,	3M

Table 2. Algorithm A.2 Addition reduced divisors

1. Compute resultant r of u_1, u_2:

$$y_1 = \frac{U_{11}}{Z_1^2} - \frac{U_{21}}{Z_1^2} , \quad y_2 = \frac{U_{20}}{Z_1^2} - \frac{U_{10}}{Z_1^2} , \quad y_1 = \frac{U_{11}}{Z_1^2} \cdot \frac{y_1}{Z_1^2} + \frac{y_2}{Z_1^2} = \frac{U_{11}y_1 + y_2 Z_1^2}{Z_1^4} , \quad r = \frac{y_2}{Z_1^2} \cdot \frac{y_3}{Z_1^4} + \frac{y_1^2}{Z_1^4} \cdot \frac{U_{10}}{Z_1^2} = \frac{y_2 y_3 + y_1^2 U_{10}}{Z_1^6} .$$

2. Compute almost inverse $inv = r/u_2 \bmod u_1$, $inv = inv_1 x + inv_0$:

$$inv_1 = \frac{y_1}{Z_1^2} , \quad inv_0 = \frac{y_3}{Z_1^4} .$$

3. Compute $s' = r \cdot s = (v_1 - v_2) inv \bmod u_1$, $s' = s_1' x + s_0'$:

$$w_0 = \frac{V_{10} - V_{20}}{Z_1^3 Z_2} , \quad w_1 = \frac{V_{11} - V_{21}}{Z_1^3 Z_2} , \quad w_2 = \frac{inv_0}{Z_1^4} \cdot \frac{w_0}{Z_1^3 Z_2} = \frac{inv_0 w_0}{Z_1^7 Z_2} , \quad w_2 = \frac{inv_1}{Z_1^2} \cdot \frac{w_1}{Z_1^3 Z_2} = \frac{inv_1 w_1}{Z_1^5 Z_2} ,$$

$$s_1' = \frac{(inv_0 + inv_1 Z_1^2)(w_0 + w_1)}{Z_1^2 Z_2} - \frac{w_2}{Z_1^7 Z_2} - \frac{w_3}{Z_1^5 Z_2}\left(1 + \frac{U_{11}}{Z_1^2}\right) , \quad s_0' = \frac{w_2}{Z_1^7 Z_2} - \frac{U_{10}}{Z_1^2} \cdot \frac{w_3}{Z_1^5 Z_2} = \frac{w_2 - U_{10}w_3}{Z_1^7 Z_2} .$$

4. Compute $s'' = x + s_0'/s_1'$:

$$w_1 = \frac{1}{r \cdot s_1} = \frac{Z_1^6 \cdot Z_1^7 \cdot Z_2}{r \cdot s_1} , \quad w_2 = r \cdot w_1 = \frac{r}{Z_1^6} \cdot \frac{Z_1^6 Z_1^7 Z_2}{s_1 \cdot r} = \frac{Z_1^7 Z_2}{s_1'} \quad \text{where } s_1' = s_1 \cdot r ,$$

$$w_3 = (s_1')^2 \cdot w_1 = \left(\frac{s_1'}{Z_1^7 Z_2}\right)^2 \cdot \frac{Z_1^{13} Z_2}{r \cdot s_1} = \frac{s_1'}{r \cdot Z_1 \cdot Z_2} , \quad w_4 = r \cdot w_2 = \frac{r}{Z_1^6} \cdot \frac{Z_1^7 Z_2}{s_1} = \frac{r \cdot Z_1 \cdot Z_2}{s_1} ,$$

$$s_0'' = s_0' \cdot w_2 = \frac{s_0'}{Z_1^7 Z_2} \frac{Z_1^7 Z_2}{s_1'} = \frac{s_0'}{s_1'} = \frac{s_0}{s_1} , \quad w_5 = w_4^2 .$$

5. Compute $l = su_2$, $l = x^3 + l_2 x^2 + l_1 x + l_0$:

$$l_2 = \frac{U_{21}}{Z_1^2} + \frac{s_0'}{s_1'} = \frac{U_{21}s_1' + s_0' Z_1^2}{Z_1^2 s_1} , \quad l_1 = \frac{U_{21}}{Z_1^2} \cdot \frac{s_0'}{s_1'} + \frac{U_{20}}{Z_1^2} = \frac{U_{21}s_0' + U_{20}s_1'}{Z_1^2 s_1'} , \quad l_0 = \frac{U_{20}}{Z_1^2} \cdot \frac{s_0'}{s_1'} = \frac{U_{20}s_0'}{Z_1^2 s_1'} .$$

6. Compute $u' = (s(l + h + 2v_1) - k) u_1^{-1}$, $k = (f - v_1 h - v_1^2)/u_1$, $u' = x^2 + u_1' x + u_0'$:

$$u_0' = (s_0'' - U_{11})(s_0'' - y_1 + h_2 w_4) - U_{10} + l_1 + (h_1 + 2V_{21})w_4 +$$

$$+ (2U_{21} + y_1 - f_4)w_5 = \left(\frac{s_0'}{s_1'} - \frac{U_{11}}{Z_1^2}\right)\left(\frac{s_0'}{s_1'} - \frac{y_1}{Z_1^2} + h_2 \frac{r \cdot Z_1 Z_2}{s_1}\right) - \frac{U_{10}}{Z_1^2} + \frac{l_1}{Z_1^2 s_1} +$$

$$+ \left(h_1 + 2\frac{V_{21}}{Z_1^3 Z_2}\right) \cdot \frac{r \cdot Z_1 \cdot Z_2}{s_1} + \left(s\frac{U_{21}}{Z_1^2} + \frac{y_1}{Z_1^2} - f_4\right) \cdot \frac{(r \cdot Z_1 \cdot Z_2)^2}{s_1^2} = \left| s_2 = s_0 Z_1^2, R = rZ_1 Z_2, s_3 = s_1 Z_1^2, \overline{R} = rZ_1^3 Z_2 \right|$$

$$= \frac{(s_2 - U_{11}s_1)(s_2 - y_1 s_1 + h_2 \overline{R})}{s_3^2} - \frac{U_{10}s_1 s_3 - l_1 s_3}{s_3^2} + \frac{(h_1 Z_1^3 Z_2 + 2V_{21})rs_3}{s_3^2} + \frac{(2U_{21} + y_1 - f_4 Z_1^2)R\overline{R}}{s_3^2}$$

$$= \frac{(s_2 - U_{11}s_1)(s_2 - y_1 s_1 + h_2 \overline{R})}{s_3^2} - \frac{U_{10}s_1 s_3 - U_{21}s_0 s_3 - U_{20}s_1 s_3}{s_3^2} + \frac{(h_1 Z_1^3 Z_2 + 2V_{21})rs_3}{s_3^2} + \frac{(2U_{21} + y_1 - f_4 Z_1^2)R\overline{R}}{s_3^2} .$$

$$u_1' = 2s_0'' - y_1 + h_2 w_4 - w_5 = 2\frac{s_0}{s_1} - \frac{y_1}{Z_1^2} + h_2 \frac{rZ_1 Z_2}{s_1} - \left(\frac{rZ_1 Z_2}{s_1}\right)^2 = \frac{2s_0 Z_1^2 - y_1 s_1}{s_1 Z_1^2} + h_2 \frac{RZ_1^2}{s_1 Z_1^2} - \frac{R^2}{s_1^2} =$$

$$\frac{2s_2 - y_1 s_1 + h_2 \overline{R}}{s_3} - \frac{\overline{R}^2}{s_3^2} = \frac{(2s_2 - y_1 s_1 + h_2 \overline{R})s_3 - \overline{R}^2}{s_3^2} .$$

7. Compute $v' \equiv -(h + s_1 l + v_2) \bmod u'$, $v' = v_1' x + v_0'$: $w_1 = l_2 - U_1'$, $w_2 = u_1' w_1 + u_0' - l_1$,

$$v_1' = w_2 w_3 - v_{21} - h_1 + h_2 u_1' = (u_1' w_1 + u_0' - l_1)w_3 - v_{21} - h_1 + h_2 u_1' = \left(\frac{U_1'}{s_3^2}\left(\frac{l_2}{s_3} - \frac{U_1'}{s_3^2}\right) - \frac{U_0'}{s_3^2} - \frac{l_1}{s_3}\right)\frac{s_1}{R} - \frac{V_{21}}{Z_1^3 Z_2} - h_1 + h_2 \frac{U_1'}{s_3^2} =$$

$$\frac{U_1'(l_2 s_3 - U_1') - U_0' s_3^2 - l_1 s_3^3}{s_3^3 Z_1^2 R} - \frac{V_{21}s_3^3 r}{s_3^3 Z_1^2 R} - h_1 + \frac{h_2 U_1'}{s_3^2} = \left| \overline{\overline{R}} = s_3 R \right| = \frac{U_1'(l_2 s_3 - U_1' + h_2 \overline{\overline{R}}) + s_3^2(U_0' - l_1 s_3 - V_{21}s_3 r - h_1 \overline{\overline{R}})}{s_3^3 \overline{R}} ,$$

$$v_0' = w_2 w_3 - v_{20} - h_0 - h_2 u_0' = \frac{U_0'}{s_3^2 R}\left(l_2 s_3 - U_1' + h_2 s_3 \overline{R}\right) - \frac{s_3^2}{s_3^2 \overline{R}}\left(V_{20}s_3 r + h_0 \overline{\overline{R}} + l_0 s_3\right) .$$

8. Compute Z_1' and Z_2': $Z_1' = s_3$, $Z_2' = \overline{R}$.

Table 3. Derivation of expressions in different steps of A.3

Specify the modifications carried out in A.2, which allowed reducing quantity of field operations.

Step A.2.1. This step is to be omitted due to existence of the same denominator of all coordinates, which allows saving 8 multiplications in $\mathbf{GF}(p)$.

Step A.2.2. While computation of y_1 and y_2, reduction to common denominator of coordinates U_{1j} and U_{2j} is not required, saves 2 multiplications in $\mathbf{GF}(p)$.

Moreover, the calculation of resulting Z_1' and Z_2' coordinates is also significantly simplified, that helps to save 5 multiplications and 3 squaring operations in $\mathbf{GF}(p)$.

Step A.2.4. While computation of w_1 and w_2, reduction to common denominator of coordinates V_{1j} and V_{2j} is also not required, saves 2 multiplications in $\mathbf{GF}(p)$.

Steps A.2.5 and A.2.6 The total number of multiplications remains unchanged, however the number of squaring operations in $\mathbf{GF}(p)$ decreases by 3, due to the interchange of calculations of coefficients l_0, l_1 and l_2 of polynomial l on the step A.2.6 (Table 4).

#	Expression	Cost
Input:	$\left[U_{11}, U_{10}, V_{11}, V_{10}, Z_1, Z_1^2, Z_2, Z_2^2\right],$ $\left[U_{21}, U_{20}, V_{21}, V_{20}, Z_1, Z_1^2, Z_2, Z_2^2\right]$	
Output:	$\left[U_1', U_0', V_1', V_2', Z_1', Z_1'^2, Z_2', Z_2'^2\right] =$ $\left[U_{11}, U_{10}, V_{11}, V_{10}, Z_1, Z_1^2, Z_2, Z_2^2\right] +$ $+\left[U_{21}, U_{20}, V_{21}, V_{20}, Z_1, Z_1^2, Z_2, Z_2^2\right],$ $weight(D_1) = weight(D_2) = 2$	
1	Compute resultant r of u_1, u_2: $y_1 = U_{11} - U_{21}$, $y_2 = U_{20} - U_{10}$, $y_3 = U_{11} \cdot y_1 + y_2 \cdot Z_1^2$, $r = y_2 \cdot y_3 + y_1^2 \cdot U_{10}$	1S, 4M
2	Compute almost inverse $inv = r/u_2 \bmod u_1$, $inv = inv_1 x + inv_0$: $inv_1 = y_1$, $inv_0 = y_3$	
3	Compute $s = (v_1 - v_2)inv \bmod u_1$, $s = s_1 x + s_0$: $w_0 = V_{10} - V_{20}$, $w_1 = V_{11} - V_{21}$, $w_2 = inv_0 \cdot w_0$, $w_3 = inv_1 \cdot w_1$, $s_0 = w_2 - U_{10} \cdot w_3$ $s_1 = (inv_0 + inv_1 \cdot Z_1^2) \cdot (w_1 + w_0) - w_2 - w_3 \cdot (Z_1^2 + U_{11})$, If $s_1 = 0$ then consider special case	6M
4	Precomputations: $R = r \cdot Z_1 \cdot Z_2$, $s_2 = s_0 \cdot Z_1^2$, $s_3 = s_1 \cdot Z_1^2$, $\tilde{R} = R \cdot Z_1^2$, $w_3 = s_1 \cdot y_1$, $w_5 = w_3 + s_1 \cdot U_{21} (= s_1 \cdot U_{11})$	7M
5	Compute $l = su_2$, $l = x^3 + l_2 x^2 + l_1 x + l_0$: $l_0 = s_0 \cdot U_{20}$, $l_2 = s_1 U_{21}$, $l_1 = (s_1 + s_0) \cdot (U_{21} + U_{20}) - l_0 - l_2$, $l_2 = l_2 + s_2$	2M
6	Compute $u' = (s(l + h + 2v_1) - k)u_1^{-1}$, $k = (f - v_1 h - v_1^2)/u_1$, $u' = x^2 + u_1' x + u_0'$: $U_1' = s_3 \cdot (2s_2 - w_5 + h_2 R) - \tilde{R}^2$ $U_0' = s_2^2 + w_3 \cdot (w_5 - 2s_2) + s_3 \cdot (y_2 \cdot s_1 + 2r \cdot V_{21} + h_1 \tilde{R}) +$ $+\tilde{R} \cdot \left[h_2(s_2 - w_5) + R \cdot (U_{11} + U_{21} - f_4 \cdot Z_1^2)\right]$	2S, 8M
7	Compute weights: $Z_1' = s_3$, $Z_2' = \tilde{R}$, $Z_1'^2 = s_3^2$, $Z_2'^2 = \tilde{R}^2$	1S
8	Compute $v' \equiv -(h + s_1 l + v_2) \bmod u'$, $v' = v_1' x + v_0'$: $V_1' = U_1' \cdot ((l_2 + h_2 R) \cdot s_3 - U_1') + s_3^2 \cdot (U_0' - s_3 \cdot (h_1 R + rV_{21} + l_1))$, $V_0' = U_0' \cdot ((l_2 + h_2 R)s_3 - U_1') - s_3^2 \cdot s_3 \cdot (l_0 + h_0 R + r \cdot V_{20})$	8M
		4S, 35M
9	Adjust: $Z_{1c} = s_1^2 \cdot Z_1^2$, $Z_{2c} = s_3^3 \cdot r$, $U_{20} = U_{20} \cdot Z_{1c}$, $U_{21} = U_{21} \cdot Z_{1c}$, $V_{20} = V_{20} \cdot Z_{2c}$, $V_{21} = V_{21} \cdot Z_{2c}$	1S, 6M
		5S, 41M

Table 4. Algorithm A.3. Co-Z reduced divisors addition

Step A.2.6. unlike algorithm A.2, the A.3 offers considering multiplier s_3 , present in each coefficient l_0 , l_1 and l_2 of polynomial l , when using coefficients l_i , $i = 0, 2$ on the steps A.2.7 and A.2.9. This allows us factor out s_3 , thus saving 3 multiplications in $\mathbf{GF}(p)$ (steps A.4.6-A.4.8).

Consider next the application of proposed algorithm for the mixed divisor addition $D_1 = \left[U_{11}, U_{10}, V_{11}, V_{10}, Z_1, Z_1^2, Z_2, Z_2^2 \right]$ and $D_2 = \left[U_{21}, U_{20}, V_{21}, V_{20}, 1, 1, 1, 1 \right]$ (mixed representation). Therefore it is necessary to reduce devisor D_2 to common Z -coordinate, i.e. $\left[U_{21} \cdot Z_1^2, U_{20} \cdot Z_1^2, V_{21} \cdot Z_1^3 \cdot Z_2, V_{20} \cdot Z_1^3 \cdot Z_2, Z_1, Z_1^2, Z_2, Z_2^2 \right]$, that requires 5 multiplications in $\mathbf{GF}(p)$. Hereinafter the provided algorithm A.4 should be used for addition of (prior formed) divisors with the same Z -coordinate.

6. Results

It is to be considered that after computing $D_3 = D_1 + D_2$ one of the items, for example D_2 , should be transformed so that it has the same Z -coordinate as divisor D_3 . For this purpose, at the step A.3.9, values Z_{1c} and Z_{2c} , where $Z_1'^2 = s_3^2 = Z_1^2 \cdot Z_{1c}$ and $Z_1'^3 Z_2' = s_3^3 \cdot \tilde{R} = Z_1^3 Z_2 \cdot Z_{2c}$, i.e. $Z_{1c} = s_1^2 \cdot Z_1^2$ and $Z_{2c} = s_3^3 \cdot r$, this requires 2 additional multiplications and 1 squaring in $\mathbf{GF}(p)$. In other words, reduction of a divisor to unified Z -coordinates takes 6 additional multiplications and 1 squaring. Ultimately, for the divisor addition step in A.1 46M+5S (M - multiplication, S - squaring) field operations are required. For the Fibonacci step 41M+5S field operations are required. If we reduce the obtained complexity estimations to the parameters of the curve (Lange, 2002c), we obtain that for the divisor addition step 40M+5S are required and for the Fibonacci step 45M+5S operations are required. In accordance to the computational complexity estimation, the approach described in this paper, is not effective, due to the complexity of mixed addition is 36M+5S (Lange, 2002c). However, the alternative approach to the divisor addition for the scalar multiplication implementation is proposed. Let us draw a computational complexity comparison between scalar multiplication algorithms described in (Kovtun and Zbitnev, 2004) and those suggested in this paper, based on idea (Meloni, 2007).

6.1 Comparison with other method

The results of known and proposed algorithms comparison are set out in Table 5.

The algorithm complexity represented in field operations (Table 6).

Assume that S=0,8M and scalar multiplier is an 80-bit integer and refer to the estimations (Kovtun and Zbitnev, 2004; Meloni, 2007) for the estimation of complexity of scalar multiplication algorithms. The results of comparison are set out in Table 7.

Computational complexity of Fibonacci-and-add scalar multiplication algorithm is by 23% greater than Binary left-to-right algorithm and by 12,5% greater than Window Fibonacci-and-add.

In other case computational complexity of Fibonacci-and-add scalar multiplication algorithm in weighted coordinates is by 23% greater than Binary left-to-right algorithm in mixed weighted coordinates and by 14,2% greater than Window Fibonacci-and-add.

Weighted coordinates with Co-Z approach are more effective than ordinary projective coordinates with Co-Z approach.

Alg. #	Curve description
1	$h(x) = 0$ (Harley, 2000)
2	$h_2 = 1$ (Lange, 2002a)
3	$h(x) = 0$ (Matsuo et al., 2001)
4	$h(x) = 0$, $f_4 = 0$ (Miyamoto et al., 2002)
5	$h(x) = 0$ (Takahashi, 2002)
6	$f_4 = 0$, $h_2 \neq 0$ (Lange, 2002a)
7	$\deg(h) = 2$, $h_i \in \mathbf{F}_2$ (Lange, 2002b)
8	$\deg(h) = 2$, $h_i \in \mathbf{F}_2$ (Kovtun and Zbitnev, 2004)
9	$h(x) = 0$, $f_4 = 0$ (Kovtun, 2006)
10	$h(x) = 0$, $f_4 = 0$ (Lange, 2002c)
11	$\deg(h) = 2$, $h_i \in \mathbf{F}_2$ (Kovtun, 2010)
12	$h(x) = 0$, $f_4 \neq 0$ [proposed]
13	$h(x) = 0$, $f_4 = 0$ [proposed]

Table 5. Algorithms and curve parameters

#	Addition					Doubling				
	General			Mixed		General			Mixed	
	$()^{-1}$	$\wedge 2$	$*$	$\wedge 2$	$*$	$()^{-1}$	$\wedge 2$	$*$	$\wedge 2$	$*$
Affine coordinates										
1	2		27			2		30		
2	2	3	24			2	6	26		
3	2		25			2		27		
4	1		26			1		27		
5	1		25			1		29		
6	1	3	22			1	5	22		
Projective coordinates $\left[U_1, U_0, V_1, V_0, Z \right]$										
7		4	47	3	40		6	40	5	25
8		4	46	4	39		6	39	5	25
9		4	46	4	39		6	35	5	24
Weighted coordinates $\left[U_1, U_0, V_1, V_0, Z_1, Z_2, Z_1^2, Z_2^2 \right]$										
10		7	47	5	36		7	34	5	21
Co-Z projective coordinates $\left[U_1, U_0, V_1, V_0, Z \right]$										
11		4	46				4	42		
Co-Z weighted coordinates $\left[U_1, U_0, V_1, V_0, Z_1, Z_2, Z_1^2, Z_2^2 \right]$										
12		5	46				5	41		
13		5	45				5	40		

Table 6. Computational complexity of group law in Jacobean of genus 2 HEC over $\mathbf{GF}(p)$

#	Scalar multiplication algorithm	Cost, M
Addition in mixed projective coordinates (Kovtun and Zbitnev, 2004)		
1	Binary (left-to-right)	5192
2	NAF	4629
3	w-NAF, $w = 4$	4349
Addition with Co-Z method in projective coordinates (Kovtun, 2010)		
7	Fibonacci-and-add	6773
8	Window Fibonacci-and-add	5970
Addition in mixed weighted coordinates (Lange, 2002c)		
4	Binary (left-to-right)	5104
5	NAF	4570
6	w-NAF, $w = 4$	4307
Addition with suggested method, alg. #13 from table 1		
7	Fibonacci-and-add	6629
8	Window Fibonacci-and-add	5829

Table 7. Computational Complexity Of Scalar Multiplication Algorithms

6.2 Another aspect of using the proposed approach

Yet another aspect of using a biometric authentication on based Co-Z approach to divisor scalar multiplication in Jacobian of genus 2 hyperelliptic curves over fields with odd characteristic in weighted coordinates is using it in the fight against Cyber terrorism. Using this approach and biometric authentication will significantly reduce financial losses of enterprises.

Primary solutions of the given problem are confirmed also with the data presented on figure 1 on which (Kavun, 2007) finance indexations of the put damage for some countries are shown. Apparently from the presented statistics, the state infrastructure is more developed; the larger it receives damage from cyber criminality.
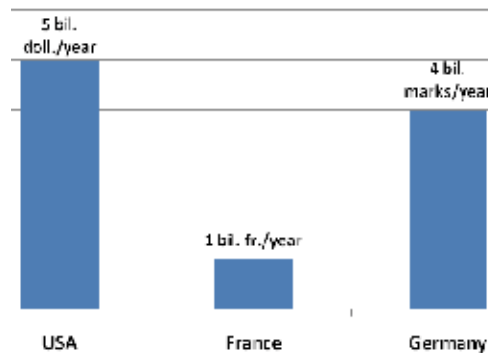

Fig. 1. Damage from cyber criminalities

During an epoch of the world economic crisis and modern transformations of economy of different countries the aspect of economic security becomes even more urgent . For example, the tendency of increase in crime in sphere of cyber terrorism, having an economic (money) basis is shown in figure 2.
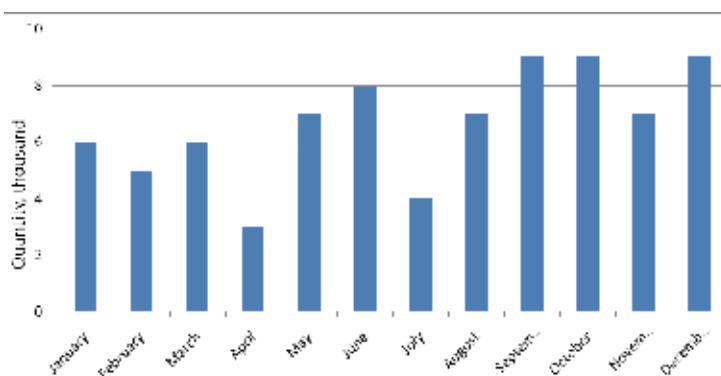
Fig. 2. Statistics of cyber criminality rate

From the presented data which has been average for some years (Kavun, 2007), it can be seen, that on the average the amount of incidents in the sphere of cyber criminality over a year increased by 50 %. It testifies to negative tendencies of cyber criminality development in the world.

Outcomes of the conducted analysis are presented in figure 3 from which it appears, that main sources of illegal operations are the USA and China which cover 50 % of all threats (Kavun, 2009).

7. Conclusion

In the paper, a new algorithm of weight 2 divisor addition with identical (shared) Z-coordinates (by the Co-Z approach) has been proposed, which requires more $\mathbf{GF}(p)$ operations then algorithm (Kovtun and Zbitnev, 2004), however it allows a decrease in computational complexity of Fibonacci-and-add scalar multiplication algorithm while approaching to the Binary left-to-right algorithm.

Biometrics are not secrets and are not revocable (Schneier,1999) while revocability and secrecy have been critical requirements of conventional cryptosystem design, one then wonders whether it is possible to design a secure authentication system from the system components which in themselves are neither secrets nor revocable—for example, whether the methods of ensuring liveness of biometric identifiers and challenge–response schemes (Maltoni et al., 2003) obviate fraudulent insertion of "stolen" biometric identifiers. Is it possible to nontrivially combine knowledge and biometric identifiers to arrive at key generation/release mechanisms where biometric identifiers are necessary but not sufficient for cryptographic key generation/release? Is it possible to require multiple biometrics to make it increasingly difficult for the attacker to fraudulently insert multiple biometrics into the system? Is it possible to make it unnecessary to revoke/update the cryptographic key in the event of a "stolen biometric"? Exploring challenges in designing such systems is a promising (yet neglected) avenue of research. When cryptobiometric systems eventually come into practical existence, there is a danger that biometric components may be used as an irrefutable proof of existence of a particular subject at a particular time and place. Mere incorporation of biometrics into a system does not in itself constitute a proof of identity. We need to understand how these foolproof guarantees can be theoretically proved in a deployed cryptosystem and how to institute due processes that will provide both

technological and sociological freedom to challenge the premises on which nonrepudiability is ascertained.

Genus 3 curves might save about a third in key lengths and so 180-bit ECC (which is beyond the usual range, as given in standards) is equivalent to 60-bit HCC, which can be implemented on a fast 64-bit computer. In practice, 160-bit ECC is typically used and so there is even room for added security in case of computational speed-ups in attacks. Stein, in particular, has pointed out how the use of "real" forms of hyperelliptic curves (i.e., with infrastructure) allows considerable speed-ups in implementation in some cases.

On the other hand, Gaudry (2000), showed that hyperelliptic curves of genus bigger than or equal to 5 and possibly 4 are less secure than hyperelliptic curves of genus g<4 (or 5) (ICCIT Biometrics Research Group, 2005). That means that the key-per-bit-strength of hyperelliptic curves of genus 2 and 3 is the same as for elliptic curves, and thus far better than conventional systems based on discrete logs or integer factoring. In fact, genus 2 is particularly interesting because the arithmetic appears to be only minimally slower than elliptic curve arithmetic and the bit size of the underlying finite field is half as big as for elliptic curves having the same security level. To our knowledge nobody has performed a down-to-earth implementation of genus 2 hyperelliptic curves.

We are considering hardware implementations of hyperelliptic curves of genus 2 and 3 . Among the currently available hyperelliptic curves there are, for instance, Koblitz curves and curves constructed by a complex multiplication method (CM-method). Both are natural extension of ideas from elliptic curves.

8. References

Basiri, A.; Enge, A. & Faugère, J. C. & Gurel, N. (2004). The arithmetic of Jacobian groups of superelliptic cubics, *Mathematics of Computation*. Retrieved 22.01.2011 on: http://www.ams.org/journals/mcom/2005-74-249/S0025-5718-04-01699-0/S0025-5718-04-01699-0.pdf

Brier, E.; Joye, M. (2002). Weirstrass elliptic curves ans side-channel attacks, *Proceedings of the International Workshop: Practice and Theory in Public Key Cryptosystems*, PKC 2002, LNCS 2274, Springer-Verlag, pp.335-345

Chudnovsky, D. V.; Chudnovsky, G. V. (1986). Sequence of number generated by addition in formal group and new primality and factorization test, *Advanced in Applied Math*, 8, pp.385–434.

Clavier, C.; Joye, M. (2001). Universal exponentiation algorithm – A first step towards provable SPA-resistance cryptosystems, *Proceedings of the International Workshop: Cryptographic Hardware and Embedded Systems*, CHES 2001, LNCS 2162, pp.300-308.

Cohen, H.; Miyaji, A. & Ono, T. (1998). Efficient elliptic curve exponentiation using mixed coordinates, *Proceedings of the International Conference on the Theory and Applications of Cryptology and Information Security: Advances in Cryptology*. CRYPTO'98. LNCS 1514. Berlin: Springer–Verlag, pp. 51–65.

Coron, J. (1999). Resistance against differential power analysis for elliptic curve cryptosystems, *Proceedings of the International Workshop: Cryptographic Hardware and Embedded Systems*, CHES 1999, LNCS 1717, pp.292-302.

Gaudry, P. (2000). An algorithm for solving the discrete log problem on hyperelliptic curves. In Bart Preneel, editor, *Advances in Cryptology - EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 19-34, Berlin, Springer-Verlag.

Goundar, R. R. (2008). *Addition chains in application to elliptic curve cryptosystems*: PhD thesis, Kochi University, Japan.

Hao, F.; Anderson, R.; Daugman, J. (2005). Combining cryptography with biometrics effectively, *Technical Report*, No. 640, University of Cambridge.

Izu, T.; Takagi, T. (2002). A fast parallel elliptic curve multiplication resistant against side channel attacks,*Technical Report CORR*, CORR 2002-03, University of Waterloo, 2002.

Jain, A. K.; Bolle, R. & Pankanti, S. (1999). *Biometrics: Personal Identification in Networked Society*. Norwell, MA: Kluwer.

Klein, D. V. (1990). "Foiling the cracker: a survey of, and improvements to, password security," in *Proc. 2nd USENIX Workshop Security*, 1990, pp. 5–14.

Koblitz, N. (1987). Elliptic curve cryptosystems. *Mathematics of Computation*, 48(177), pp. 203–209.

Koblitz, N. (1989). Hyperelliptic cryptosystems. *Journal of cryptology*, 1, pp.139-150.

Kovtun, V. Yu.; Zbitnev, S. I. (2004). Arithmetic operations in Jacobian of Genus 2 hyperelliptic curves in projective coordinates with reduced complexity, *East-Europian magazin of advanced manufacturing services*. 1 (13). pp. 14–22.

Lange, T. (2001). *Efficient arithmetic on hyperelliptic curves*: PhD thesis: Mathematics and Informatics. University of Essen: Institute for experimental mathematics. Germany: Essen.

Lange, T. (2002a). Efficient arithmetic on genius 2 hyperelliptic curves over finite fields via explicit formulae, *Cryptology ePrint Archive*. Report 2002/121. Available

Lange, T. (2002b). Inversion–free arithmetic on genius 2 hyperelliptic curves, *Cryptology ePrint Archive*. Report 2002/147.

Lange, T. (2003). Formulae for arithmetic on genius 2 hyperelliptic curves. September 2003. Retrieved 22.01.2011 on:
http://www.ruhr–uni–bochum.de/itsc/tanja/preprints/expl_sub.pdf.

Lange, T. (2002c). Weighted coordinates on genius 2 hyperelliptic curves, *Cryptology ePrint Archive*. Report 2002/153.

Maltoni, D.; Maio, D.; Jain, A. K. & Prabhakar, S. (2003). *Handbook of Fingerprint Recognition*. New York: Springer-Verlag.

Matsuo, K.; Chao J. & Tsujii S. (2001). Fast genius two hyperelliptic curve cryptosystem, *Technical report IEICE*. ISEC2001–31.

Miller, I. V. S. (1985). Use of elliptic curves in cryptography. In H. C. Williams, editor, *Advances in Cryptology - CRYPTO'85*, volume 218 of LNCS, Springer, pp. 417–426.

Menezes, A.; Wu, Y. H. & Zuccherato, R. (1998). An elementary introduction to hyperelliptic curves, In: Koblitz N. ed., *Algebraic aspects of cryptography*. Berlin, Heidelberg, New York: Springer–Verlag, pp. 28–63.

Meloni, N. (2007). New point addition formul. for ECC applications. In C. Carlet and B. Sunar, editors, *Arithmetic of Finite Fields (WAIFI 2007)*, LNCS 4547, Springer, pp. 189–201.

Miyamoto, Y.; Doi H.; Matsuo K.; Chao J. & Tsujii S. (2002). A fast addition algorithm of genius two hyperelliptic curve, *Symposium on cryptography and information security*. SCIS'2002. Japan: IEICE, pp.497–502.

NIST (2001). Advanced encryption standard (AES), Federal information processing standards publication 197. Retrieved 22.01.2011 on:

http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

Okeya, K.; Sakurai, K. (2000). Power analysis breaks elliptic curve cryptosystems even secure against timing attack, *Proceedings of the International Conference on the Theory and Applications of Cryptology and Information Security: Advances in Cryptology*, INDOCRYPT 2000, LNCS 1977, Springer-Verlag, pp.178-190.

Schneier, B. (1999). Biometrics: uses and abuses, *Commun. ACM,* 42(8), p. 136.

Sugizaki, H.; Matsuo, K.; Chao, J. & Tsujii, S. (2002). An extension of Harley addition algorithm for hyperelliptic curves over finite fields of characteristic two, *Technical report IEICE*. ISEC2002–09.

Stallings, W. (2003). *Cryptography and Network Security: Principles and Practices*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall.

Takahashi, M. (2002). Improving Harley algorithms for jacobians of genius 2 hyperelliptic curves, *Symposium on cryptography and information security*. SCIS'2002. Japan: IEICE, pp.155–160.

Wollinger, T. (2004). *Software and hardware implementation of hyperelliptic curve cryptosystems*: PhD dissertation: Electronics and informatics. Worchester Polytechnic Institute,Germany: Bochum.

Wayman, J. L. (2001). "Fundamentals of biometric authentication technologies," *Int. J. Image Graph.*, 1(1), pp. 93–113.

A New Fingerprint Authentication Scheme Based on Secret-Splitting for Enhanced Cloud Security

Ping Wang[1], Chih-Chiang Ku[1] and Tzu Chia Wang[2]
[1]Department of Information Management, Kun Shan University,
[2]Institute of Computer and Communication Engineering,
National Cheng Kung University,
Taiwan

1. Introduction

The number of commercially-available web-based services is growing rapidly nowadays. In particular, cloud computing provides an efficient and economic means of delivering information technology (IT) resources on demand, and is expected to find extensive applications as network bandwidth and virtualization technologies continue to advance. However, cloud computing presents the IT industry not only with exciting opportunities, but also with significant challenges since consumers are reluctant to adopt cloud computing solutions in the absence of firm guarantees regarding the security of their information.

Two fundamental issues arise when users applying cloud computing to software as a service (SaaS). First, if enterprise data is to be processed in the cloud, it must be encrypted to ensure its privacy. As a result, efficient key management schemes are required to facilitate the encryption (and corresponding decryption) tasks. Second, as the sophistication of the tools used by malicious users continues to increase, the data processed in the cloud is at increasing risk of attack. Consequently, there is an urgent requirement for robust authentication schemes to ensure that the data can be accessed only by legitimate, authorized users.

Network attacks such as phishing or man-in-the-middle (MITM) attacks present a serious obstacle to consumer acceptance of cloud computing services. According to reports released by privacy watchdog groups in the US, more than 148 identity theft incidents, affecting nearly 94 million identities, occurred in 2005 in the US alone (Mark, 2006). Identity theft is therefore one of the most severe threats to the security of online services. As a result, it is imperative that SaaS providers have the means to authenticate the identity of every user attempting to access the system. Due to the non-denial requirements of remote user identity authentication schemes, this is most commonly achieved using some form of biometrics-based method.

The term "biometrics" describes a collection of methods for identifying individuals based upon their unique physiological or behavioral characteristics (Furnell et al. 2008). Generally speaking, the physiological characteristics include the individual's fingerprint, vein pattern, DNA and shape of face, while the behavioral characteristics include the handwriting dynamics, voice and gait. Automated biometric recognition systems are now widely used

throughout the automotive; IT and banking industries (see Figs. 1 and 2). For example, Miura et al. (2005) developed a biometric authentication system based on the individual's finger vein for accomplishing secure online authentication over small devices such as notebook computers, cell phones, and so on.

Fig. 1. Fingerprint scanner and smart card reader
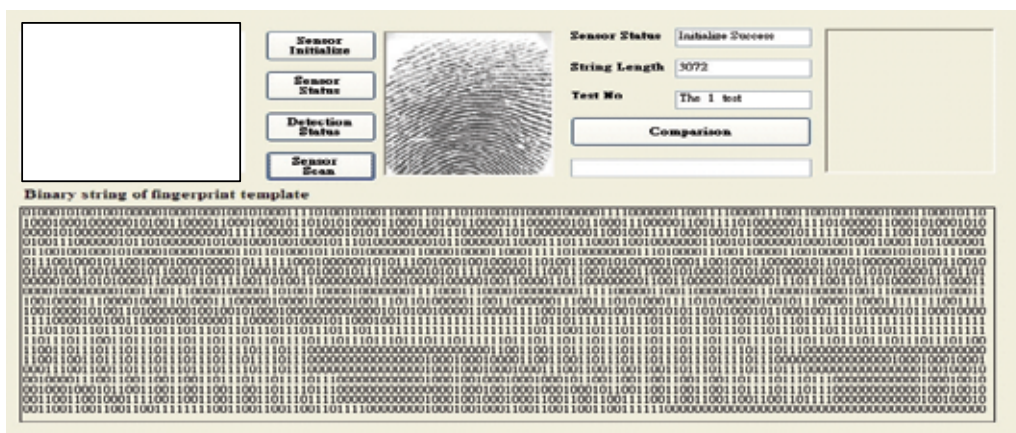
Fig. 2. Sensor-based scanning of user fingerprint template

However, existing biometric authentication systems cannot absolutely guarantee the identity of the individual. For example, biometric features such as the fingerprint may be acquired surreptitiously and then used by a malicious user. Similarly, even in multi-factor authentication methods such as smart cards, in which the biometric information is protected using a password, the password may be cracked by network hackers and the biometric information then copied and counterfeited. These proposals are invariably based on the assumption of employee honesty. Unfortunately, this assumption cannot also be guaranteed in practical applications, and many real cases have been reported in which dishonest interior staff have stolen users' authentication details from the authentication database and have then used these details to acquire the customers' private information for financial gain (see Fig.3).

Fig. 3. Theft of biometric features by dishonest staff

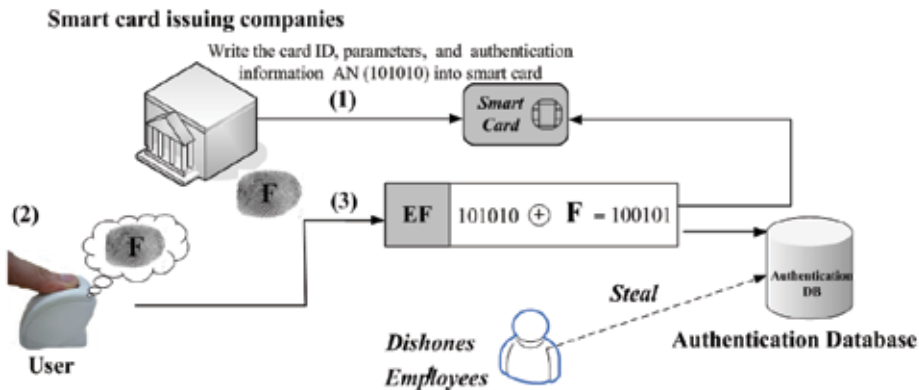To resolve the security issues described above, the present study proposes a new remote authentication scheme based on a secret-splitting concept. In the proposed approach, part of the biometric data is encrypted and stored on a smart card, while part of the data is encrypted and stored on a server (see Fig. 4). This approach not only resolves the problem of data abuse by interior staff, but also helps protect the users' information against malicious attack such as hacking into the Certificate Authority (CA) since to counterfeit the entire biometric information, dishonest staff or hackers must simultaneously decrypt two secret keys rather than just one.
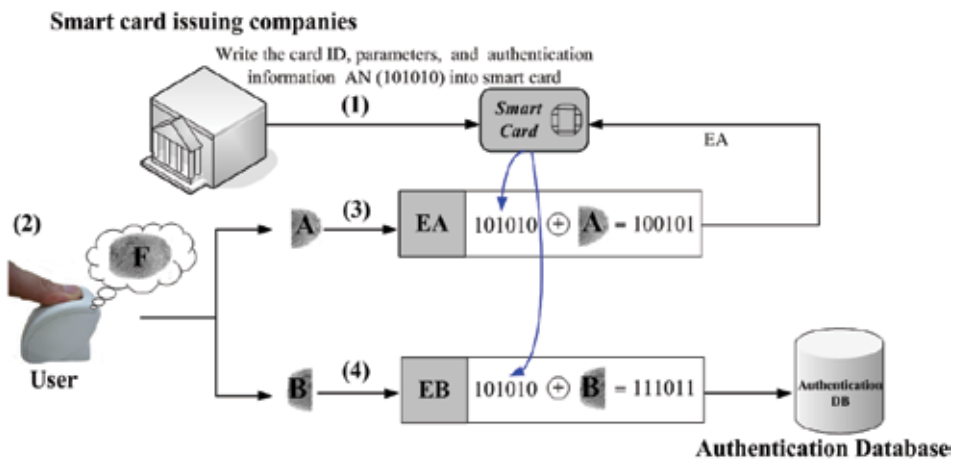
Fig. 4. Proposed authentication scheme based on secret-splitting

In addition to the secret-splitting concept, the proposed authentication scheme utilizes the Diffie-Hellman key exchange / agreement algorithm (Diffie & Hellman, 1976) to guarantee the security of the data transmissions between the terminal and the server. The main differences between the scheme proposed in this study and existing methods can be summarized as follows: (i) the smart card stores only part of the fingerprint template used in the identity authentication process. As a result, the user's identity is protected even if the card is lost or stolen. (ii) the template information stored on the smart card and server,

respectively, is independently encrypted. Consequently, the information obtained by a hacker or dishonest member of staff from a successful attack on the authentication database is insufficient to pass the liveness test.

A remote authentication scheme based on a secret-splitting concept is proposed for resolving the problem of user privacy in cloud-computing applications. In contrast to existing multi-factor authentication schemes, the proposed method minimizes the threat of attacks by dishonest interior employees since only a subset of the information required to pass the liveness test is stored on the user authentication database.

The remainder of this chapter is organized as follows. Section 2 briefly reviews the essential properties of identity authentication schemes and presents the related work in the field. Section 3 introduces the remote identity authentication scheme proposed in this study. Section 4 examines the robustness and computational efficiency of the proposed approach. Finally, Section 5 presents some brief concluding remarks.

2. Existing multi-factor authentication methods

Remote authentication is essential in ensuring that only legitimate individuals are able to access a network and make use of its resources. Typically, remote authentication is achieved using one of the following well-known schemes: (1) user account / password, (2) network address / domain name, (3) shared secret keys, (4) public keys, (5) digital signatures, (6) biometric authentication, (7) digital certificates, or (8) smart cards. Figure 5 shows a typical authentication procedure using a smart card in conjunction with the user's fingerprint.
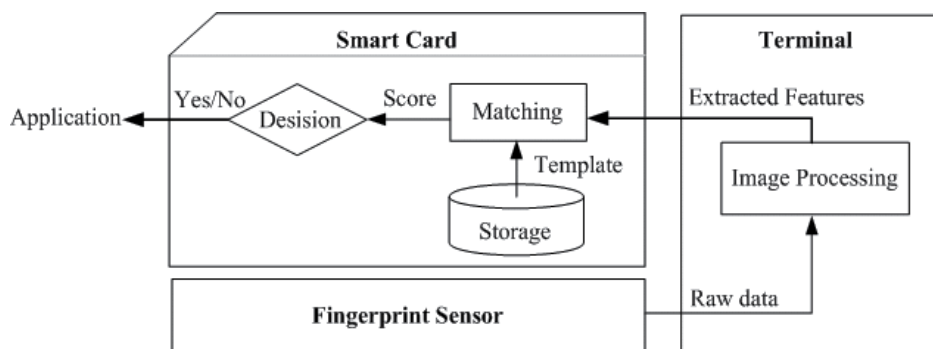
Fig. 5. Remote authentication using smart card and fingerprint

Password-based authentication schemes have the advantage of simplicity, but are reliant upon the user memorizing the password and remembering to modify it periodically in order to maintain the security of their account. Smart cards, in which the user's identity authentication information is encrypted on an embedded chip, have a number of benefits compared to traditional password-based methods, namely (i) the user's information is protected by a simple Personal Identity Number (PIN); (ii) the risk of identity theft is minimized by means of sophisticated on-chip defense measures; and (iii) single smart cards can be programmed for multiple uses, e.g. banking credentials, medical entitlement, loyalty programs, and so forth. Consequently, various multi-factor authentication schemes based upon smart cards and integrated biometric sensors have been proposed.

In a pioneering work of multifactor authentication schemes, For example, J.K. Lee et al. (2002) proposed a remote identity authentication scheme in which a smart card was

integrated with a fingerprint sensor. In the proposed approach, the smart card, a secret password, and the user's fingerprint were taken as inputs in the login process and the fingerprint minutiae, encrypted with the time stamp and the user's authentication template, were then compared with the authentication value stored on the card. To enhance the security of the ElGamal public key system (ElGamal, 1985) used in J.K. Lee et al. (2002), the encryption parameters were randomly generated in accordance with both the user's fingerprint minutiae and the time stamp. However, while the proposed method is therefore robust toward replay attacks, clock synchronization is required at all the hosts, which is a significant challenge in open network environments.

Kim et al. (2003) proposed an integrated smart card / fingerprint authentication scheme in which a password list was not maintained at the server such that the users were able to change their passwords at will. Moreover, protection against replay attacks was provided by means of Nonce technology, thereby avoiding the need for clock synchronization at the hosts. However, Scott (2004) showed that the Nonce-based design rendered the system vulnerable to imitation attacks given the collection by a hacker of a sufficient number of network packets to calculate the authentication value.

Later on, Lin and Lai (2004) showed that under certain circumstances, the method proposed by Lee et al. was unable to resist identity masquerade attacks. Accordingly, the authors proposed a new scheme for enhancing the security of the method presented in J.K. Lee et al. (2002) by allowing the users to choose and change their passwords at will. However, Mitchell and Tang (2005) showed that the method proposed by Lin and Lai also contained a number of serious flaws, most notably (i) hackers may simply copy the fingerprint from the imprint cup; (ii) the time stamp generated during a legal login procedure may be detected by a malicious user and then modified in order to login illegally at some future point in time; and (iii) the system contains no rigorous mechanism for preventing malicious users from using old passwords to perform an illegal login operation when the legitimate users change their passwords.

Recently, Fan et al. (2006) proposed a three-factor remote authentication scheme based on a user password, a smart card and biometric data. Importantly, in the proposed approach, the server only stores an encrypted string representing the user identity. That is, the biometric data is not revealed to any third party; including the remote server. The scheme is implemented using a two-step procedure. In the first step (the registration step), an encrypted user template is constructed by mixing a randomly-chosen string with the biometric characteristics of the user via an exclusive-or operation (XOR). In the second step (the login step), the fingerprint minutiae obtained via a sensor are encrypted using a second randomly-chosen string, and the two strings are then sent to a sensor for matching. The scheme has the advantage that all three security factors (the password, the smart card and the biometric data) are examined at the remote server, i.e., the system is a truly three-factor remote authentication system. Furthermore, the authentication process performed at the remote server does not require the exact value of the user's biometric data to be known. As a result, the security of the user's biometric information is improved. However, when the registration process is performed in a centralized way (e.g., a central Certificate Authority (CA) is used to issue certified users with a certificate), the scheme is vulnerable to the theft of the user's fingerprint minutiae by interior dishonest staff. In other words, while the scheme preserves the privacy of the users in the login and authentication phases, the security of the biometric data is not guaranteed during the registration phase.

3. A secret-splitting remote authentication scheme

This section proposes a novel remote authentication protocol for network services based on the secret-splitting concept. The proposed protocol comprises three phases, namely the initialization phase, the registration phase, and the authentication phase (see Fig. 6).
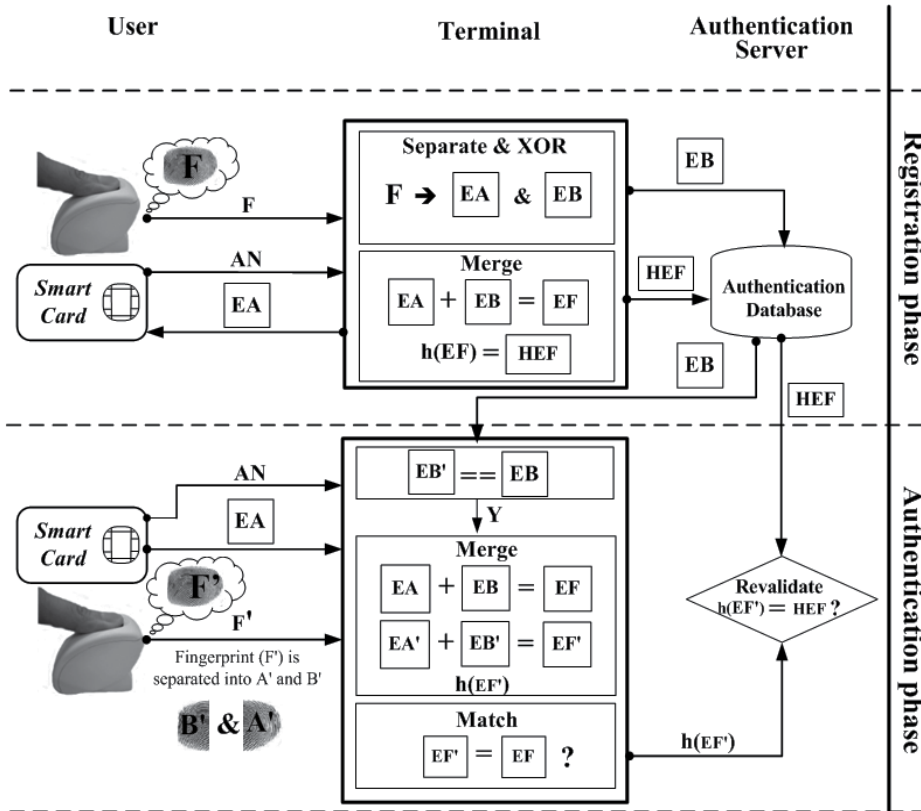
Fig. 6. Proposed fingerprint matching concept

In the initialization phase, the manufacturer produced a smart card and wrote a set of unique security parameters (AN) into it. In the registration phase, the user registers with a CA organization, and verifies their legal identity by means of traditional physical identity documents such as an identity card or a social security card. Encrypted fingerprint template A (EA') is generated from the information extracted by a fingerprint scanner (EA) and the smart card information obtained from a card reader (AN), and the remaining part of the encrypted fingerprint template B (EB') is directly extracted from the authentication database. Once the two templates (EA', EB') have been combined into a complete template by the terminal, the comparison results are sent to the server to verify the legality of the user, that is, (EF = EF'). Note that this step is designed to prevent counterfeit attacks in which a malicious hacker sends a "legal user" message directly from the terminal in order to deceive the server.

Fig. 7 presents the function flow diagram of the proposed remote authentication scheme. The details of each phase in the scheme are presented in Sections 3.1~3.3.
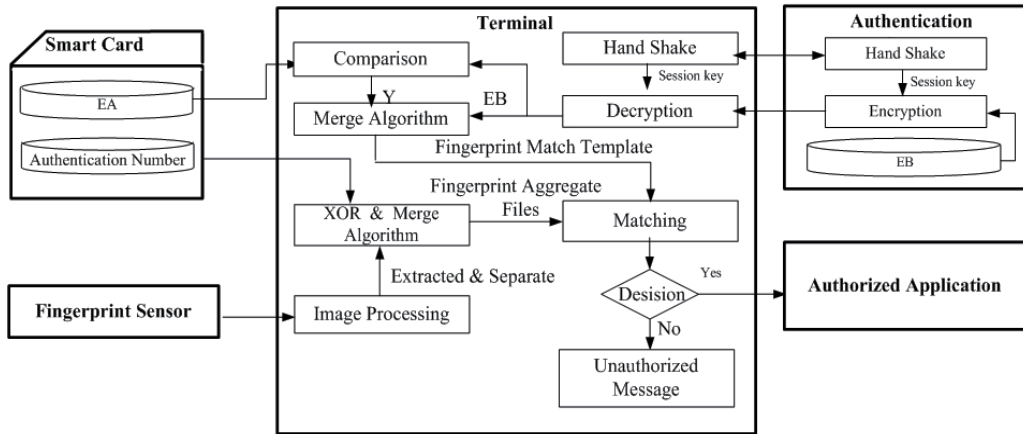
Fig. 7. Function flow diagram of proposed remote authentication scheme

3.1 Initialization phase

When the smart card manufacturer accepts an order from the CA, it writes various security parameters into the cards (e.g., the card number or authentication number (AN)) and then sends the cards to the CA. The detailed procedure is shown in Fig. 8

Step 1.1: Manufacturer randomly chooses a large prime number p and determines its root α.

Step 1.2: Manufacturer generates a unique Authentication Number (AN) based on a pre-defined coding rule.

Step 1.3: Manufacturer randomly selects a 128-bit string K as a key for symmetric encryption and keeps (p, α, AN, K) secret.

Step	Executor	Actions
1.1	*Manufacturer*	Randomly choose a large prime number *p* and determines its root α
1.2	*Manufacturer*	Generate a unique Authentication Number (*AN*)
1.3	*Manufacturer*	Randomly select a 128-bit string *K* as a key for symmetric encryption and keep [*p*, α, *AN*, *K*] Smart card secret.

Fig. 8. Initialization phase

3.2 Registration phase

The user registers with the CA and receives a smart card once he or she has confirmed their legal identity using some form of physical identity document. As shown in Fig. 9, the registration phase comprises five steps, namely:

Step 2.1: Let user U_i with identity ID_i be about to register with the server. The user chooses a card password, PW_i , the password is then saved to the smart card, and then protected by the encryption mechanism of smart card.

Step 2.2: The fingerprint image of User U_i is obtained via a sensor and the minutiae are extracted from this image to form a fingerprint template F_i. The terminal separates F_i into two parts, F_{iA} and F_{iB}, where F_{iA} and F_{iB} represents part A and part B of fingerprint template, respectively.

Step 2.3: The terminal computes $EA_i = h (F_{iA} \oplus AN)$, $EB_i = h (F_{iB} \oplus AN)$, and $HEF_i = h (EA_i \cup EB_i)$, where $\cup$ is a merge operation and h(.) is a public one-way hash function.

Step 2.4: The terminal sends (ID_i, hEB_i, p, α, K, HEF_i) to the server over a secure channel

Step 2.5: The terminal stores (ID_i, PW_i, hEA_i) in the smart card.

Step	Executor	Actions
2.1	U_i	Determine a card password PW_i
2.2	U_i *Terminal*	Form a fingerprint template F_i using the fingerprint minutiae obtained via a sensor. (Note that F_i represents the fingerprint template of user Ui.) Separate F_i into two parts, F_{iA} and F_{iB}.
2.3	*Terminal*	Compute $EA_i = F_{iA} \oplus AN$, $EB_i = F_{iB} \oplus AN$ $HEF_i = h(EA_i \cup EB_i)$
2.4	*Terminal➔ Server*	Send (ID_i, hEB_i, p, α, K, HEF_i)
2.5	*Terminal*	Store (ID_i, PW_i, hEA_i) on smart card
	Terminal ➔ U_i	[ID_i, PW_i, hEA_i, p, α, AN, K] $_{\text{Smart card}}$

Fig. 9. Registration phase

3.3 Authentication phase

Users insert their smart card, containing a partial authentication template into a card reader and a login request is then sent to the authentication server. The fingerprint information is checked using the following eight-step procedure (see Figs. 10 and 11):

Step 3.1: User U_i inputs his or her password PW_{i*} into the terminal. If the password is correct, the AN is extracted; else the login request is rejected.

Step 3.2: Users "provide their fingerprint via a sensor, and the fingerprint is then compared with that stored on the authentication server. Let F_i* represent the fingerprint minutiae extracted by the sensor. The terminal separates F_i* into F_{iA*} and F_{iB*}, and then computes $EA_{i*} = F_{iA*} \oplus AN$ and $EB_i*= F_{iB*} \oplus AN$. The two parts (i.e., EA_{i*}, EB_{i*}) are then merged to generate the full biometric template of the user, i.e., $EF_{i*} = EA_{i*} \cup EB_{i*}$. The server sends EBi to the terminal for comparison purposes in order to verify the user's legal identity. If a match is obtained (i.e., $EB_i == EB_{i*}$), the authentication process proceeds to Step 3.3; else it terminates.

Step 3.3: (Diffie-Hellman key exchange algorithm). The terminal randomly selects a number X_A such that $X_A < p$, and then computes $Y_T = \alpha^{X_A} \bmod p$ and $Y_A = ID_i \mid\mid Y_T$, where α and p are both stored on the smart card. The terminal then sends YA to the server. Similarly, the server randomly selects a number X_B such that $X_B < p$, computes $Y_B = \alpha^{X_B} \bmod p$, and then sends Y_B to the terminal.

Step 3.4: The terminal uses Y_B to compute the session key $SK = (Y_B)^{X_A} \bmod p$. Similarly, the server uses Y_A to compute the common session key, $SK = (Y_A)^{X_B} \bmod p$. Note that SK is a shared secret between the terminal and the server.

Step	Executor	Actions
3.1	U_i	Input password PW_i^*
	Terminal	Examine PW_i^* and gain AN
3.2	U_i	Scan finger to provide information required to construct fingerprint template F_i^*
	Terminal	Separate F_i^* into F_{iA}^* and F_{iB}^*, where $EA_i^* = F_{iA}^* \oplus AN$, $EB_i^* = F_{iB}^* \oplus AN$, and $EF_i^* = EA_i^* \cup EB_i^*$
	Server ➜ *Terminal*	Extract EB_i from server. If a match is obtained (i.e., $EB_i == EB_i^*$), go to Step 3.5; else terminate the authentication process
3.3	*Terminal*	Randomly select a number X_A such that $X_A <$p Compute $Y_T = \alpha^{X_A} \bmod p$.
	Terminal ➜ Server	$Y_A = ID_i \,\|\, Y_T$
	Server	Randomly select a number X_B such that $X_B <$p
	Server ➜ *Terminal*	Send Y_B to terminal
3.4	*Terminal*	$SK = (Y_B)^{X_A} \bmod p$
	Server	$SK = (Y_A)^{X_B} \bmod p$

Fig. 10. Authentication phase (Steps 3.1~3.4).

Step 3.5: The server generates a one-time symmetric key RK, computes $M = E_{sk}(EB_i \,\|\, RK)$, and then sends M to the terminal. Note that $E_{sk}(.)$ denotes a symmetric encryption function (such as the *AES* method) based on the session key SK.

Step 3.6: The terminal acquires EB_i and RK by performing the decryption process $D_{SK}(M)$, and extracts EA_i from the smart card. EA_i and EB_i, are then merged to obtain $EF_i = EA_i \cup EB_i$, where $D_{sk}(.)$ denotes a symmetric decryption function based on the session key SK.

Step 3.7: The terminal compares EF_i^* and EF_i. If a match is obtained, the legal user is successfully identified; else the terminal sends $RM = E_{RK}(h(EF_i) \,\|\, CM)$ to the server for reconfirmation purposes. Note that E_{RK} is a symmetric decryption function based on the key RK, and CM is a message indicating the matching result.

Step 3.8: The server re-verifies the match $h(EF_i) == HEF_i$. If a match is obtained, the server accepts the login request of U_i; else it rejects the request.

Step	Executor	Actions
3.5	*Server*	Generate a one-time private key *RK*
	Server➜Terminal	Compute $M=E_{sk}(EB_i||RK)$ and send *M*
3.6	*Terminal*	*Decrypt* $D_{SK}(M)$ to obtain EB_i and *RK*
		Extract EA_i using card reader, then merge two parts of template, i.e., $EF_i= EA_i \cup EB_i$
3.7	*Terminal*	If Match ($EF_i^*== EF_i$), then CM=true; else CM=false
	Terminal ➜ Server	Return the Comparison Message (*CM*) and EF_i with encryption $RM=E_{RK}(h(EF_i)||CM)$
3.8	*Server*	*Decrypt* $D_{RK}(RM)$ to obtain $h(EF_i)$ and *CM*
		Verify ($h(EF_i) == HEF_i$)
		Accept the login request of U_i if match is obtained; else reject login request.

Fig. 11. Authentication phase (Steps 3.5~3.8).

Summarizing the procedures shown in Figs. 9~11, the overall sequence diagram of the proposed remote authentication scheme can be illustrated as shown in Fig. 12.
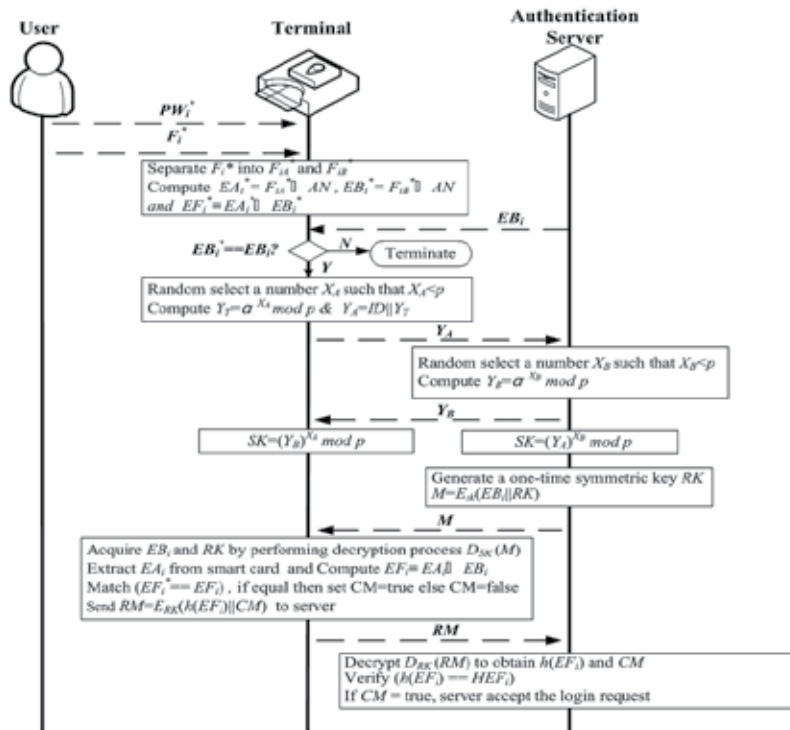

Fig. 12. Sequence diagram of proposed remote authentication scheme

4. The security performance and computational efficiency

This section discusses the security performance and computational efficiency of the proposed remote authentication scheme.

4.1 Security analysis

This sub-section demonstrates the robustness of the proposed authentication scheme toward three common forms of attack, namely (i) authentication factor attacks; (ii) network attacks; and (iii) interior attacks originating from the card-issuing organization.

4.1.1 Authentication factor attacks

Given a three-factor authentication scheme (i.e., password, smart card and user biometrics), a hacker requires all three factors in order to successfully complete the authentication process. It is possible that the smart card and password may be stolen or duplicated. However, in the scheme proposed in this study, the biometrics template is strongly protected using a secret-splitting technique. Thus, even if a hacker manages to obtain the partial fingerprint template EB_i, he or she cannot generate the partial template EA_i without possessing the knowledge of the authentication number (AN) stored on the smart card. Besides, hackers have no matters to generate the other part of biometrics template (EA_i), except they crack the program which is used for merging EA_i* and EB_i*, however, this program generally is an executive binary code and burn in the ROM of card issuing machine. In other words, it is extremely difficult for a hacker or a dishonest member of staff to obtain all three authentication factors for a particular user, and thus the proposed scheme is as safe as other multi-factor authentication schemes.

As described above, in the proposed approach, the biometric data of a user is separated into two parts (EA_i, EB_i), encrypted and stored on a smart card and a server, respectively. This approach not only preserves the privacy of the users in the login and authentication phases, but also helps protect the users' information against the theft of the user's fingerprint minutiae by interior dishonest staff.

4.1.2 Network attacks

This section demonstrates the robustness of the proposed authentication scheme toward three common types of network attack, namely (i) man-in-the-middle attacks, (ii) dictionary attacks, (iii) replay attacks.

Strong encryption authentication helps prevent *man-in-the-middle* attacks. In the proposed scheme, the authentication template, encrypted using a *128-bit* AES symmetric encryption algorithm, is split into two parts; stored on the smart card and the server, respectively. To prevent from the man-in-the-middle attacks, the data transmissions between the terminal and the server are protected by a session key generated using the Diffie-Hellman key exchange algorithm. Therefore, hackers are not easily able to steal the complete set of biometric data. Thus, the security of the biometric data is further enhanced since solving the Discrete-Logarithm Problem (DLP) in order to crack the Diffie-Hellman protected transmissions is extremely hard within a finite period of time [14]. In addition, the decryption process is further complicated (from the hacker's perspective) by the fact that the session key is changed on a periodic basis. Thus, a hacker not only faces a major challenge in determining the AN of the smart card and the coding used to construct the partial authentication template EA_i, but also encounters severe difficulties in cracking the encrypted transmission packets exchanged between the terminal and the server.

For *dictionary* attacks, cracking a password needs either weak password strength or large quantity of hash of the target password; two cases can be prevented by both strong hash function algorithms such as MD5 and the SHA family and long character password with numbers, mixed case, and symbols in Step 2.1.

Assume that a hacker has attained the formation of the terminal (AN, EA_i^*, EB_i^*) in Steps 3.1~3.4 from the terminal and smart card, and then launches a *replay attack* to counterfeit a legal user in the authentication process. In Step 3.5, a one-time session key (RK) is randomly generated by the server. This key is valid only for the current authentication process. In other words, old session keys cannot be re-used, and thus imitation attacks are thwarted.

4.1.3 Attacks originating within card issuing organization

This section demonstrates the robustness of the proposed scheme toward interior staff attacks in the registration phase and authentication phase, respectively.

Registration phase

In the registration process, the users scan their finger in order to provide the system with the fingerprint minutiae required to construct the finger template (see Step 2.2). The fingerprint template F_i, stored in the Random Access Memory (RAM) of the terminal is utilized only in the subsequent registration process. That is, to prevent exposure of the user's biometric data to any unauthorized third party, F_i and its related parameters are deleted as soon as the authentication process is complete. Therefore, interior staff and external hackers have little chance of acquiring F_i since it exists within the system for only a short period of time and, moreover, its location within the terminal RAM varies dynamically.

Authentication phase

As shown in Fig. 6, the three components of the authentication template generated in the proposed scheme are stored separately in the cards, terminal and servers ("on two different physical components, namely (i) EA_i is stored on the smart card; (ii) and (iii) EB_i is stored at the authentication server. Thus, even if the template data at the authentication server is stolen by a dishonest member of staff, the authentication process cannot be completed since the remaining template information is missing. In practice, a dishonest member of staff can only complete the authentication without password and smart card, except someone is capable of copying process by somehow copying the user's card and acquiring the user's fingerprint from the imprint cup.

4.2 Computational complexity

In this section, the computational complexity of the proposed scheme is compared with that of the schemes presented by Fan *et al.* (2006), Lin and Lai (2004), Kim *et al.* (2003) and J.K. Lee *et al.* (2002) (see Table 1). Among the various computations performed by the different schemes, the exponential operation in the decryption procedure (E) is the most time consuming. It is observed that the number of exponential operations in the schemes proposed by Lee *et al.* and Lin and Lai, respectively, is slightly higher than that in the proposed scheme and significantly higher than that in the scheme proposed by Fan et al. In addition, it is seen that the overall computational complexity of the scheme proposed in this study is slightly higher than that of the scheme proposed by Fan *et al.* due to the separation of the authentication data and the encryption of the symmetric keys during transmission.

Compared to the scheme proposed by Fan et al., the proposed scheme requires three additional exponential operations and two additional merge operations. However, the number of symmetric decryption operations is reduced by one, while the number of hash and *XOR* operations is reduced by three and one, respectively. Significantly, the scheme proposed by Fan *et al.* utilizes the Rabin algorithm (Rabin, 1979) to protect the symmetric keys during transmission. Whilst this approach reduces the number of exponential operations required, the security of the communications between the terminal and the

authentication server cannot be guaranteed in an open environment. By contrast, the scheme proposed in this study uses the Diffie-Hellman key exchange /agreement algorithm to protect the terminal-server communications. Thus, while a greater number of exponential operations are required (i.e., to solve the Discrete-Log Problem), the security of the transmissions is significantly improved relative to that in Fan *et al.*'s scheme.

It is acknowledged that the proposed scheme has certain limitations. For example, in the event that the user loses his or her smart card, the CA cannot immediately re-issue a new card since they do not possess the complete fingerprint template. In other words, the users must repeat the registration process in order to obtain a new card. Furthermore, the computational complexity of the proposed scheme is slightly higher than that of existing schemes. However, compared to existing methods, the proposed scheme ensures the security of the users' biometric information even if the contents of the authentication database are stolen. In other words, the proposed scheme achieves a compromise between the need to reduce the computational cost of the remote authentication process and the need to minimize the security threat posed by dishonest interior staff.

Scheme	Characteristic		
	Computational cost of login and authentication phases	Store complete biometric template	Clock synchronization
Proposed scheme	4E+3SE+3SD+H+2X+2M	No	No
Fan et al. (2006)	E+3SE+4SD+4H+3X	Client	No
Lin and Lai (2004)	5E+3H+4X	Client	Yes
Kim et al. (2003) (Timestamp-based)	4E+2H	Server	Yes
Kim et al. (2003) (Nonce-based)	4E+1H	Server	No
J.K. Lee et al. (2002)	7E+2H+2X	Server	Yes

Table 1. Comparison of related schemes (revised from Lee S.W. et al., 2005)

Note that E represents the computational time required to perform modular exponentiation; SE denotes the computational time required to perform modular symmetric encryption; SD is the computational time required to perform modular symmetric decryption computation; H denotes the computational time required to perform a one-way hash function; X is the computational time required to perform a modular exclusive-or operation; and M represents the computational time required to perform a modular merge operation.

5. Conclusions

This paper has presented a novel remote authentication scheme based on a secret-splitting concept for cloud computing applications. Compared to existing methods, the proposed scheme has a number of important advantages, namely (i) the users can choose passwords (PW_i) for their smart cards at will; (ii) the smart card and server each store a partial biometric template rather than the full template; and (iii) the partial templates are integrated only when the users have successfully completed the login process in the authentication phase. The proposed scheme is robust toward three common forms of attack, i.e., man-in-the-middle attacks, dictionary attacks and replay attacks. As a result, it provides an effective solution for enhancing the security of cloud computing applications, and is therefore beneficial to SaaS service providers in improving user acceptance of their services.

6. Acknowledgements

7. References

Diffie W. & Hellman M. E. (1976). Multiuser Cryptographic Techniques, *Proceedings of National Computer Conference*, New York, June 7-10, 1976

ElGamal T. (1985). A Public-Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms, *Proceedings of IEEE Transactions on Information Theory*, Vol.31, No. 4, pp. 469–472, ISSN 0018-9448

Fan C. I.; Lin Y. H. & Hsu R. H. (2006). Remote Password Authentication Scheme with Smart Cards and Biometrics, *Proceedings of 49th annual IEEE Global Telecommunications Conference (GLOBECOM)*, pp.1-5, San Francisco, California, USA, 27 Nov, 2006

Jeong J.; Chung M. Y. & Choo H. (2006). Secure User Authentication Mechanism in Digital Home Network Environments, *Lecture Notes in Computer Science (LNCS)*, Vol.4096, pp.345-354

Kim H. S.; Lee S. W. & Yoo K. Y. (2003). ID-based Password Authentication Scheme Using Smart Cards and Fingerprints, *ACM SIGOPS Operating Systems Review*, Vol.37, No.2, pp.32-41, ISSN 0163-5980

Lee J. K.; Ryu S. R. & Yoo K. Y. (2002). Fingerprint-based remote user authentication scheme using smart cards, *Electronics Letters*, Vol.38, No.12, pp.554-555, ISSN 0013-5194.

Lee S. W.; Kim H. S. & Yoo K. Y. (2005). Efficient nonce-based remote user authentication scheme using smart cards, *Applied Mathematics and Computation*, Vol.167, No.1, pp. 355-361, ISSN 0096-3003

Lin C. H. & Lai Y. Y. (2004). A flexible biometrics remote user authentication Scheme, *Computer Standards & Interfaces,* Vol. 27, No.1, , p.19-23, ISSN 0920-5489

Mark K. H. (2006). Data theft scandal - what we can learn from India Opinion, In: Offshoring, 6 Oct 2006. Available from http://services.silicon.com/offshoring/0,380000487 7,39163049,00.htm

Mitchell C. J. & Tang O. (2005). Security of the Lin-Lai smart card based user authentication scheme, Technical Report, Royal Holloway, University of London, 2005 Available from from http://www.rhul.ac.uk/mathematics/techreports

Miura N.; Nagasaka A. & Miyatake T. (2005). Extraction of Finger-Vein Patterns Using Maximum Curvature Points in Image Profiles, *Proceedings of the 9th IAPR Conference on Machine Vision Applications (MVA2005)*, pp.347-350, Tsukuba Science City, Japan, 2005.

Pfitzmann A. (2008). Biometrics---How to Put to Use and How Not at All, In: TrustBus 2008, Furnell S.M.; Katsikas S.K. & Lioy A. (Ed.), LNCS 5185, pp. 1–7, Springer-Verlag, ISSN 0302-9743

Rabin M. O. (1979). Digitalized Signatures and Public-key Functions As Intractable As Factorization, *Technical Report of MIT/LCS/TR212, MIT Labatory*, Computer Science Cambridge, MA, USA

Scott M. (2004). Cryptanalysis of an ID-based password authentication scheme using smart cards and fingerprints, *ACM SIGOPS Operating Systems Review*, Vol. 38, No. 2, pp.73-75, ISSN:0163-5980

Part 4

Other Application

Biometric Applications of One-Dimensional Physiological Signals – Electrocardiograms

Jianchu Yao[1], Yongbo Wan[2] and Steve Warren[3]
[1]East Carolina University, Greenville, NC,
[2]Oklahoma State University, Stillwater, OK,
[3]Kansas State University, Manhattan, KS,
USA

1. Introduction

In the last decade, affordable computing power has become available on platforms intended for low-power, mobile applications. For example, Gumstix Overo products integrate WiFi/Bluetooth connectivity, microSD storage, and 600 MHz Texas Instruments OMAP (Open Multimedia Application Platform) 35xx processors with up to 256 MB of flash memory/SDRAM, offering laptop-like resources and performance in a form factor of a stick of gum (Chaoui, et al., 2001). This speaks to the potential for wearable, wireless medical devices based on such products to process signals on-board; functionality that previously required expensive, bulky, benchtop/bedside equipment. These computational capabilities also show promise for smart devices that implement context-based intelligence and on-device expert systems for clinical decision making. This move toward highly-capable and intelligent mobile medical devices drives the need for verification tools, data integrity checkers, and role-based security mechanisms that can also be implemented at the embedded level, since the potential usage scenarios and associated protection needs are numerous in comparison with legacy medical device applications in controlled hospital and home care settings. For example, many implementations of personal and body area networks have been developed to facilitate ambulatory monitoring of health status, where physiological parameters such as heart rate, heart activity, blood oxygen saturation, and respiration rate can be gathered through the use of mobile, wearable electrocardiographs and pulse oximeters (Jovanov, et al., 2009; Galeottei, et al., 2008; Chuo, et al., 2010). These data need to be authenticated and checked for integrity before they are stored in electronic patient records, which implies the need for "owner-aware" devices that verify user identity as part of the data acquisition process (Warren, et al., 2005; Warren & Jovanov, 2006). Many biometric authentication protocols are computationally intensive and can well-utilize the emerging computational capabilities of low-power mobile devices.

A broad range of biomedical data, from physiological signals/images to behavioral traits, has been explored for its biometric authentication and identification potential (Biel, et al., 2001; Chan, et al., 2008; Doi & Yamanaka, 2004; Duc, et al., 1997; Elsherief, et al., 2006; Faundez-Zanuy, 2005; Irvine, et al., 2001; Israel, et al., 2005; Shen, et al., 2002; Sullivan, et al., 2007; Yao & Wan, 2010; G. H. Zhang, et al., 2009). Image-based mechanisms that assess

fingerprints (Doi & Yamanaka, 2004; Matsumoto, et al., 2002), retinal patterns (Elsherief, et al., 2006), facial features (Daugman, 1998; Koh, et al., 1999; Philips, et al., 2003), and vein/palm structures (Doi & Yamanaka, 2004) are the leading biometric modalities used today for identity verification. Recently, one-dimensional physiological signals such as electrocardiograms (ECGs) (Agrafioti & Hatzinakos, 2008a, 2008b; Biel, et al., 2001; Israel, et al., 2005; Kyoso & Uchiyama, 2001; Micheli-Tzanakou, et al., 2009; Nasri, et al., 2009; Plataniotis, et al., 2006; Saechia, et al., 2005; Shen, et al., 2002; Singh & Gupta, 2008; Yao & Wan, 2008, 2010), photoplethysmograms (PPGs) (Ludeman & Chacon, 1995; Ludeman & Chacon, 1996; Love, et al., 1997; Ma et al., 2006; Bao et al., 2005; Gu, et al., 2003; Gu & Zhang, 2003; Wan, et al., 2007; Yao, et al., 2007) and electroencephalograms (EEGs) (Marcel & Millan, 2007) have garnered attention as promising biometric candidates based on the following thoughts:

- In many cases, these signals are already acquired and stored as part of the healthcare delivery process. It is therefore sensible to utilize them as identification attributes because no additional user action or data gathering is required, as is the case with most other biometric modalities. In other words, authentication, identification, or verification can occur behind the scenes even without subject awareness (Warren & Jovanov, 2006; Warren, et al., 2005). Further, no additional hardware would be required to implement this feature, which implies that biometric features can be added to the system without incurring significant additional device cost. The efficiency of this approach in terms of care delivery workflow, coupled with the ease of use and economic sensibility of such tools, should lead to increased technology acceptance by patients and providers.

- Since they represent an individual's underlying physiological status, these signals may be less sensitive to environmental factors that affect other biometric parameters, thereby avoiding substantial deteriorations in biometric performance when fully controlled environments are unobtainable. For example, environmental noise unavoidably interferes with voice-based biometric systems (Ming, et al., 2007). In such cases, costly environmental improvements are often required to ensure that identification systems work properly.

- The use of physiological signals for identification or verification can help to prevent failure to enroll (FTE) issues that may occur when a subject does not possess a particular biometric. Most current biometric approaches are affected by this. In fingerprint identification, for example, it has been estimated that fingerprints from up to 4% of the population cannot be used for identification purposes due to the poor quality of the fingerprint ridges (Jain, et al., 2004).

- Some current biometric approaches are subject to forgery. Artificial fingerprints, for example, can be constructed to circumvent a fingerprint verification system. Unlike most current biometric data, which are extracted from "surficial" parts of the human body, physiological signals represent core internal behavior and are hard to emulate with tissue phantoms (Jain, et al., 2004). This "innerness" makes the identification process less prone to forgery, preventing imposters from disguising their true identity by changing these metric patterns.

A number of research efforts have attempted to address these thoughts within the context of one-dimensional biomedical signals. This chapter provides an up-to-date review of this research, summarized from the following four perspectives: (1) the signals used for identification, with an emphasis on ECGs, (2) signal processing methods, (3) classification

methods, and (4) identification algorithm performance. As noted in (Matyas & Riha, 2003), "the main issue in biometric authentication systems is performance." High identification accuracy is critical for practical biometric technologies.

This chapter also presents some of the authors' research findings geared toward quantitatively evaluating the identification accuracy of ECG data as population size increases. It presents the design of a wearable electrocardiograph and the associated signal processing algorithms, followed by an assessment of identification-algorithm performance for this application. In the analysis, three distance measures were defined in the wavelet domain: the Distance of Discrete Wavelet Coefficients (DDWC), the Distance of Continuous Wavelet Coefficients (DCWC), and the Ratio of the volume of Intersection to the volume of Union (RItU). Evaluation results for all three distance measures demonstrated consistent declination as the population grew to a size of 50. Possible causes for this performance drop are discussed. These experiments also recognized that distinguishable information from these signals may not be as prevalent as the unique data acquired using more popular modalities. In other words, the number of possible combinations for the patterns of the statistical attributes that can be extracted from these signals is limited. Based on these findings, the chapter suggests scenarios that ECGs can be utilized as the sole modality for biometric purposes or those they can serve as a supplemental tool to other modalities.

2. ECG as a biometric modality: a systematic review

Researchers have recognized for some time that ECGs contain innate human attributes (i.e., they reflect the electro-myocytic properties of the heart), so it seems sensible that each individual's ECG may demonstrate his or her uniqueness and therefore be useful for identity verification. This section provides a detailed review of research geared toward the usefulness of ECGs as biometric indicators.

Since the first such effort was reported in 1999 (Biel, et al., 1999), approximately 20 different groups have researched this interesting area (Israel, et al., 2005; Chan, et al., 2008; Biel, et al., 2001; Yao & Wan, 2008; Zhang & Wei, 2006; Biel, et al., 1999; Boumbarov, et al., 2009; Chan, et al., 2006; Chiu, et al., 2008; Fatemian & Hatzinakos, 2009; Gahi, et al., 2008; Israel, et al., 2003; Kim, et al., 2005; Kyoso, 2003; Kyoso & Uchiyama, 2001; Micheli-Tzanakou, et al., 2009; Nasri, et al., 2009; Plataniotis, et al., 2006; Shen, et al., 2002; Sufi & Khalil, 2008; Wang, et al., 2006; Yao & Wan, 2010) – see Table 1. Most of these published works use a similar approach to present research findings. They first start with a brief description of the physiological origin of the ECG and its characteristics (e.g., P wave, QRS complex, and T wave) (Webster, 1998). They then present the three primary steps in a typical classification process—feature selection, pre-processing, and classification (which usually includes enrollment/training and identification/testing). Finally, they draw optimistic conclusions from the classification results. Despite this consistency in presentation format, the projects themselves differ with regard to (a) ECG data sources, (b) data collection processes, (c) classification feature selection, and (d) classification methods adopted to realize the final identification results. Each is detailed below:

- **ECG Data Sources:** Multiple groups experimentally acquired ECG data from volunteers (Biel, et al., 1999 2001; Chan, et al., 2006; Chan, et al., 2008; Gahi, et al., 2008; Israel, et al., 2005; Israel, et al., 2003; Kim, et al., 2005; Kyoso, 2003; Kyoso & Uchiyama, 2001; Sufi & Khalil, 2008; Wan & Yao, 2008; Wang, et al., 2006; Yao & Wan, 2008, 2010).

In these cases, experimental conditions were often well controlled: subjects were requested to rest for a period of time prior to data collection (Biel, et al., 2001; Chan, et al., 2008; Yao & Wan, 2008, 2010), with the exception of studies that examined the performance of the identification methods under conditions where heart rate was increased (Kim, et al., 2005). Some of these groups developed their own devices (Kyoso, 2003; Kyoso & Uchiyama, 2001; Wan, et al., 2007; Yao & Wan, 2008), while others collected data with off-the-shelf ECG products (Sufi & Khalil, 2008; Chan, et al., 2008). Other non-experimental data sources were employed, as in (Agrafioti & Hatzinakos, 2008a; Plataniotis, et al., 2006; Singh & Gupta, 2008; Agrafioti & Hatzinakos, 2008b), where ECG data were extracted from existing databases (e.g., PTB ("The PTB Diagnostic ECG Database") and MIT-BIH ("MIT-BIH Database Distribution"); both of these databases are available through the Internet for public research use.

- **Data Collection:** Some publications provide subject demographic data, including gender (Biel, et al., 2001; Kim, et al., 2005; Yao & Wan, 2008), age range (Biel, et al., 2001; Yao & Wan, 2008; Chan, et al., 2008; Kim, et al., 2005; Yao & Wan, 2010), and heart condition (Agrafioti & Hatzinakos, 2008a; Chiu, et al., 2008; Kim, et al., 2005; Plataniotis, et al., 2006; Singh & Gupta, 2008; Sufi & Khalil, 2008). However, few report complete demographic or health-condition information for participants. Furthermore, the time interval between subject enrollment and data collection, a critical element when determining the effectiveness of a biometric modality, is frequently overlooked (Chiu, et al., 2008; Gahi, et al., 2008; Israel, et al., 2005; Kim, et al., 2005; Kyoso, 2003; Plataniotis, et al., 2006; Saechia, et al., 2005; Sufi & Khalil, 2008). Even studies that record this information often mention it vaguely (Chan, et al., 2008; Fatemian & Hatzinakos, 2009; Singh & Gupta, 2008) (see Table 1).

- **Selection of Classification Features:** Most investigators assess time domain features (e.g., time intervals between P, Q, R, S, and T waves, along with their amplitudes) (Biel, et al., 2001; Boumbarov, et al., 2009; Gahi, et al., 2008; Israel, et al., 2005; Kyoso, 2003; Z. Zhang & Wei, 2006) and angle information (Singh & Gupta, 2008). Others believe that post-transform features are more distinctive and will therefore improve identification performance. For example, wavelet transformation was used in (Chan, et al., 2006; Chan, et al., 2008; Chiu, et al., 2008; Yao & Wan, 2008, 2010) to find the wavelet coefficients and distances in the wavelet domain that optimally quantify the similarity between two ECGs. Autocorrelation coefficients are a third type of statistical feature under investigation (Agrafioti & Hatzinakos, 2008a; Plataniotis, et al., 2006). In addition to these three types of analytical information, the appearance of the ECG waveforms was added as a classification feature in (Wang, et al., 2006). Finally, after recognizing the difficulties encountered when delineating ECG cycles, some investigators extracted classification features without the need to detect fiducial points (Plataniotis, et al., 2006; Agrafioti & Hatzinakos, 2008a), where the DCT (Discrete Cosine Transform) approach did not rely on the accurate location of each ECG cycle.

- **Classification Algorithms:** As in other pattern recognition domains, numerous classification algorithms have been created for human identification based on ECGs, where algorithm performance varies widely. While most of these approaches used variations of a "distance" concept (e.g., Euclidean distance (Israel, et al., 2003; Plataniotis, et al., 2006) or Mahalanobis' distance (Kyoso, 2003; Kim, et al., 2005) to quantify the similarities between the unknown data and the waveforms enrolled in the

database, the classification algorithms they adopted were different, including (a) classic linear discriminate analysis (LDA) (Agrafioti & Hatzinakos, 2008a; Chan, et al., 2008; Kim, et al., 2005; Kyoso, 2003; Wang, et al., 2006), (b) neural networks (Saechia, et al., 2005; Shen, et al., 2002; Wan & Yao, 2008; Boumbarov, et al., 2009), and/or (c) voting after initial results were available from the first classification level (Israel, et al., 2005; Agrafioti & Hatzinakos, 2008b). Rather than use the intuitive distance concept, other investigators employed a sequential approach that employs a hidden Markov model (Boumbarov, et al., 2009) and a probabilistic, Bayesian-theorem-based approach (Z. Zhang & Wei, 2006). Both approaches obtained results comparable to distance-based methods.

Group	Year	Subjects	Time Span	Success Rate
Biel	2001	20	6 weeks	90-100%
Kyoso	2003	9	N/A	Wide range
Shen	2002	20	N/A	80-95%
Israel	2005	29	N/A	100%
Kim	2005	10	N/A	N/A
Saechia	2005	N/A	N/A	97%
Zhang	2006	502 records	Same datasets	82-97%
Wang	2006	13	A few years	84.6%
Plataniotis	2006	14	N/A	92.8-100%
Chan	2008	50	> 1 day	95%
Yao	2008	20	Hours to weeks	91.5%
Agrafioti	2008	27	Mixed length	96%-100%
Agrafioti	2008	14	A few years	85.6%-100%
Sufi	2008	15	N/A	93-95%
Gahi	2008	16	N/A	100%
Singh	2008	25	Same time or unclear	98.5-99%
Chiu	2008	45	N/A	100% and 81%

Table 1. Summary of research on ECG analysis as a biometric modality

- **Other Endeavors:** Unlike most of the aforementioned research, which sought better identification rates, other investigators wished to improve computational efficiency. They tried to reduce the number of necessary features by (1) selecting the most meaningful features after observing how each feature changed the classification results (Biel, et al., 2001; Agrafioti & Hatzinakos, 2008b) or (2) using methods such as principle component analysis (PCA) (Yao & Wan, 2008; Sufi & Khalil, 2008; Z. Zhang & Wei, 2006; Yao & Wan, 2010).

3. ECG as a bioidentification modality: performance evaluation

This section presents the authors' recent research on the performance and limitations of ECGs as biometric indicators, as well as other potential application fields.

3.1 Data acquisition

ECG data for this study were collected with an "in-house" device (see Fig. 1), and MATLAB scripts processed these data using wavelet-based approaches. Data collection and pre-processing details were described in (Wan & Yao, 2008; Yao & Wan, 2010). Thirty participants (26 males and 4 females) with ages ranging from 18 to 51 years were recruited for data collection. A total of 121 datasets were collected from these subjects, where each subject participated in multiple ($2 \leq N_i \leq 5$) data collection sessions; consecutive sessions were a few weeks apart.
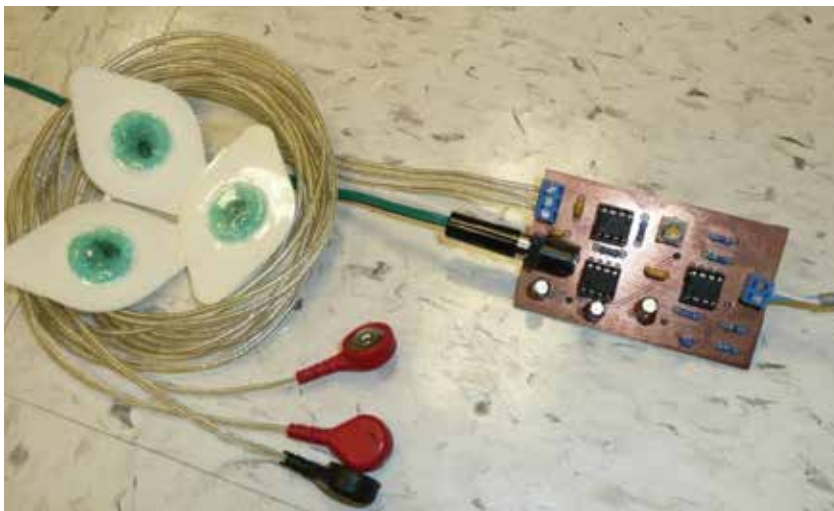

Fig. 1. An "in-house" ECG module for bioidentification data acquisition

3.2 Signal preparation and feature extraction

As specified in (Wan & Yao, 2008; Yao & Wan, 2010), the raw ECG signals were pre-processed to remove signal noise, detect R waves, and normalize each signal to a pre-defined length and amplitude range. Specifically, two major noise sources (low frequency signal drifts at around 0.06 Hz and higher frequency signal spikes at 60 Hz) were first filtered with "hard thresholding" after a scale 12 Daubechies's db6 wavelet transform was applied to all of the heart beat cycles. Detailed wavelet parts at scales 2, 3, and 4 were reconstructed so that the R peaks could be located as the fiducial points to identify ECG cycles. Identified ECG cycles were interpolated to a pre-defined length for the convenience of future steps. Sixty consistent heartbeat cycles from each of the datasets were selected and their amplitudes were normalized to the range of [-1, 1]. In this step, data consistency was examined by calculating the Euclidean distance between the mean of each cycle and the mean of all cycles.

A wavelet transform (Yao & Wan, 2008), similar to (Chan, et al., 2008; Chiu, et al., 2008), was applied to each processed time-domain signal, and then wavelet coefficients were calculated

for each cardiac cycle within that signal. Depending on the continuous or discrete wavelet transform applied, the number of coefficients varied, as discussed earlier when the specific measures were introduced. Six sets of wavelet coefficients (corresponding to sixty heart beats) were saved for each of the 121 ECG datasets. From this point on, the wavelet coefficients served as the "statistical features" and were manipulated for subsequent classification decisions.

Out of the N_i coefficient sets obtained from each subject, one coefficient set (corresponding to one heart interval) was enrolled in the database, creating a database of 30 coefficient sets. The other N_i-1 coefficient sets (corresponding to N_i-1 heart intervals of the same subject) were used for classification tests: 121-30 = 91 coefficient sets.

3.3 Measures of signal similarity/difference

The goal of this exercise was to explore identification-algorithm performance as a function of test population size. Three distance measures were utilized to represent the level of similarity between the unknown wavelet coefficient set and the enrolled coefficient sets: (1) Distance of Discrete Wavelet Coefficients (DDWC), (2) Distance of Continuous Wavelet Coefficients (DCWC), and (3) Ratio of Intersection to Union of continuous wavelet coefficients (RItU). The following paragraphs describe the three distance measures in detail.

3.3.1 Distance of Discrete Wavelet Coefficients (DDWC)

The wavelet distance proposed in (Chan, et al., 2008) was examined first (it was referred as WDIST in (Chan, et al., 2008)). This distance is notated here as DDWC to distinguish it from the distance obtained from a continuous wavelet transform. In this case, coefficients from a discrete wavelet transform were utilized for distance measure calculations. The DDWC is defined by

$$DDWC_n = \sum_{p=1}^{P}\sum_{q=1}^{Q}\frac{\left|c_0^{p,q} - c_n^{p,q}\right|}{max(\left|c_0^{p,q}\right|, T.H.)} \tag{1}$$

where $c_0^{p,q}$ is the qth wavelet coefficient at the pth scale of the unknown coefficient set; $c_n^{p,q}$ is the qth wavelet coefficient at the pth scale of the enrolled coefficient set; P is the number of scales of the wavelet transform; and Q is the number of coefficients at a specific scale. $T.H.$ is a pre-selected normalization constant. To obtain the DDWC measure, a scale 6, Bior1.1 wavelet transform (Mallat, 1999) was applied to the pre-processed ECGs, yielding coefficient structures of 256 elements. The 'Bior1.1' basis function belongs to the Biorthogonal Wavelet Pairs wavelet family. The orthogonal discrete wavelet transform functions have excellent localization properties in both the time and frequency domains (Kharate, et al., 2007), and the coefficients obtained contain distinctive information. Note that the basis function chosen here is different from that in (Chan, et al., 2008), which used a db3 function.

3.3.2 Distances of Continuous Wavelet Coefficients (DCWC)

The discrete wavelet transform is usually implemented as a dyadic-orthogonal transform where a signal can be presented as a combination of elements in the orthogonal basis set without information redundancy. The continuous wavelet transform decomposes time-domain signals into temporal-spectral components with continuous scale factors and

translation parameters. The coefficients obtained from a continuous wavelet transform depict the detailed, smooth transitions of the signal energy distribution along the time and frequency dimensions (see Fig. 2).
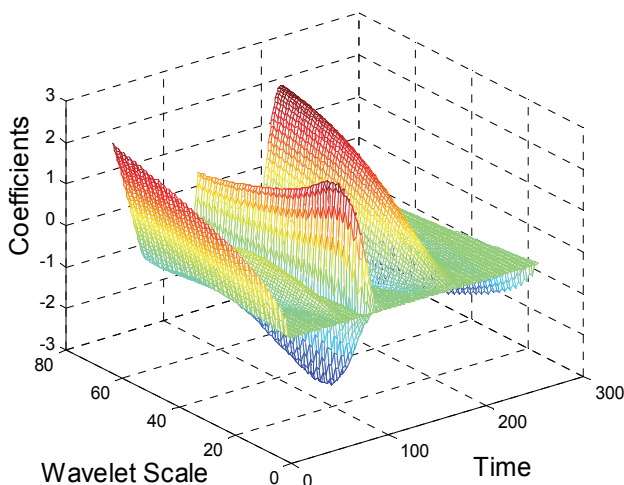

Fig. 2. A 3-D coefficient surface obtained from a continuous wavelet transform

Assuming that the inclusion of the smooth transition of coefficients at different scales will yield better identification results, a distance of continuous wavelet coefficients (DCWC) is defined by

$$DCWC_n = \frac{\sum_{p=1}^{P}\sum_{q=1}^{Q}\left|c_0^{p,q} - c_n^{p,q}\right|}{max\left(ABS(C_0)\right)} \tag{2}$$

where $c_0^{p,q}$ is the qth wavelet coefficient at the pth scale of the unknown coefficient set; $c_n^{p,q}$ is the qth wavelet coefficient at the pth scale of the enrolled coefficient set; $P = 64$ is the number of scales of the wavelet transform; and $Q = 256$ is the number of coefficients at a given scale. In this experiment, the continuous wavelet transform used the same basis function, Bior1.1, as was used in the discrete wavelet transform. Note that the denominator in Eq. (2) contains the maximum of the absolute value of the coefficients of the unknown subject. Experiments showed that this normalization could obtain better classification results than using the denominator in Eq. (1) and avoided the process of finding the threshold.

3.3.3 Ratio of Volume of Intersection to Volume of Union (RItU)
The waveform coefficients, when plotted as a mesh, form a 3-dimensional spatial surface as shown in Fig. 2. A more intuitive way to quantify the similarity of two signals is the ratio of the volume under the intersection of the two signals to the volume under the union of the two signals (see Fig. 3 for a graphical depiction of the intersection and union of two 2-D curves). The more two compared signals differ, the smaller the ratio. When the two signals are identical, the ratio is 1, and when the two signals do not overlap, the ratio is 0, implying that they are separate from each other and that the similarity between them is minimal. In addition to taking into account the distance between two coefficient sets, as in the other two

measures, the RItU measure also considers the coefficient locations over the temporal and frequency dimensions. This volume ratio is mathematically defined as

$$RItU = \bigcap\left(C^T, C^E\right) / \bigcup(C^T, C^E)$$ (3)

where C^T is the coefficient set to be tested, and C^E is one of the coefficient sets enrolled in the database. The intersection and union of the two coefficient sets are further defined by

$$\bigcap\left(C^T, C^E\right) = \sum_{p=1}^{P}\sum_{q=1}^{Q} c_i ,$$

where
$$c_i = \begin{cases} min(\,ABS(c_0^{p,q}), ABS(c_n^{p,q})), & SIGN(c_0^{p,q}) = SIGN(c_n^{p,q}) \\ 0, & SIGN(c_0^{p,q}) \neq SIGN(c_n^{p,q}) \end{cases}$$ (4)

and

$$\bigcup\left(C^T, C^E\right) = \sum_{p=1}^{P}\sum_{q=1}^{Q} c_u , \ where \ c_u = max(\,ABS(c_0^{p,q}), ABS(c_n^{p,q}))$$ (5)

where the notation follows from Eq. (2) because the spatial surfaces used to calculate RItU were also determined from continuous wavelet transforms.
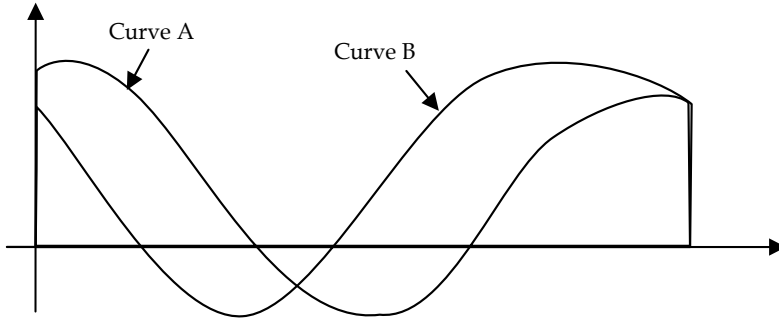

Fig. 3. The intersection (gridded area) and union (all shaded areas) of two curves

3.4 Evaluation of identification performance changes as population size increases

The classification method used in this experiment finds the distances (D) (as defined in the previous section) from the to-be-tested coefficient set C_j^T $(1 \leq j \leq 91)$ to the coefficient sets C_k^E $(1 \leq k \leq 30)$ enrolled in the database and uses these distances as the quantitative measure of signal difference/similarity. After all of the distances are compared, C_j^T is classified to the closest enrolled subject S_i. i.e., the unknown coefficient set:

$$C_j^T \rightarrow S_i , \ where \ i = arg \ i \ min D_{ij}$$ (6)

To evaluate the deterioration in accuracy as the test population size increases, a varied number (5, 10, 15, 20, 25, and 30) of subject waveforms were tested with the wavelet distance approach using the three difference/similarity measures introduced above. Coefficient sets

stored for testing purposes were randomly selected to perform identification tests. When a certain number (again, 5, 10, 15, 20, 25, or 30) of subject waveforms were tested, the total number of coefficient sets selected for testing could vary since the number of coefficient sets N_i for subject t could be different. The identification accuracy rate (AR) is defined as

$$AR = D_S \, / \, D_T \tag{7}$$

where D_S is the number of coefficient sets that have been successfully identified and D_T is the total number of coefficient sets selected for testing.

Repeated random sub-sampling was implemented to eliminate possible classification biases. A total of 20 trials with randomly selected unknown datasets were conducted for each case with a specific subject number (5, 10, 15, 20, and 25); only one test was conducted for the 30-subject case since all of the subjects were examined. In each trial, wavelet coefficient sets were selected randomly from those set aside for testing and classified according to the three measures. The average accuracy and standard deviation for all trials, using the three difference/similarity measures, was examined to analyze the biometric performance trend.

4. Experimental results

Fig. 4 illustrates identification performances when the three distance definitions, DDWC, DCWC, and RItU are utilized to measure subject similarity/difference. Comparing the three approaches, it is obvious that DDWC outperforms the other two distance measures. The latter two methods (DCWC and RItU) generate similar results, where the accuracy rate from the DCWC method is slightly higher than the RItU method. More importantly, these plots demonstrate that the classification accuracies for all three measures decline consistently by 12% as the number of test subjects increases from 5 to 30. Note also that, as the number of subjects grows, the standard deviation of the accuracy rate decreases (e.g., the DDWC method yields standard deviations of 6.6 and 2.3 for 5 and 25 subjects, respectively). This is true because more repeated datasets (and therefore a larger percentage of subjects) existed when larger numbers of subjects were incorporated.
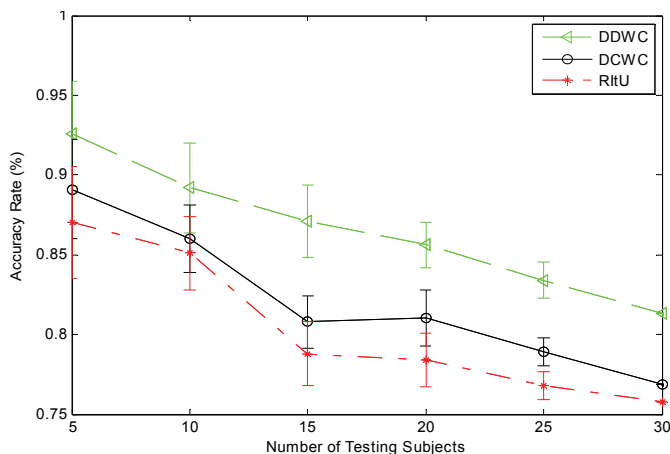

Fig. 4. Identification accuracy rate with the three difference/similarity measures versus the number of test subjects

5. Discussion

Possible causes for this performance drop can be identified. First, since signals such as ECGs are collected to provide patient health status and diagnose suspected illness, the signals are expected to show variations over time. Even for individuals whose health status does not change significantly over a period of time, normal circadian rhythms (on both a cycle-to-cycle basis and over a longer time interval) coupled with changes in stress level, emotions, and activity will in aggregate create variations that a robust identification algorithm must be able to tolerate. For some physiological signals, the detection environment may be another critical factor. PPGs, for example, which are based on light intensity either transmitted through or reflected by tissue, are extremely sensitive to motion artifacts, in spite of multiple existing approaches to help remediate these artifacts. ECGs are also sensitive to motion artifact and can be easily corrupted by electromagnetic interference that exists in most mobile patient environments. Without compensation, such variations and artifacts ultimately make one-dimensional signals less than ideal for identification or verification. More consistent attributes uniquely associated with patients are then desired.

These experiments also recognized that distinguishable information from these signals may not be as rich as the unique data acquired using popularly adopted modalities. In other words, the number of possible combinations for the patterns of the statistical attributes that can be extracted from these signals is limited. As the number of subjects increases beyond a certain number (20 to 30 in this case), the likelihood of having subjects whose signals are very similar increases significantly.

Therefore, despite the advantages that one-dimensional physiological signals may hold with respect to biometric identity assessment, performance assessments from previous research remind one that caution is required when such data are utilized for identification, especially when the subject population is large. The authors believe that these signals hold clear potential for this purpose, with the following qualifiers:

- The class of one-dimensional signals discussed here should be used with caution as a sole source of blind identification, primarily due to the less than desired uniqueness of the signal shapes and their time-dependent variations. However, when these signals are used as supplemental traits combined with other biometrics (e.g., in a data fusion approach), they are desirable due to the natural physical coupling between these various signal modalities, which is expected to improve the overall performance of the affiliated identification algorithms.

- While these nontraditional biometric modalities may not offer sufficient identification accuracy as required for legitimate authentication (i.e., where the goal is to identify an unknown subject given a large number of existing datasets), they may be better suited for individual verification, where the newly gathered signal is only compared to a recent set of data, with the assumption that the subject's identity is already inferred. Current verification processes (e.g., the two-stage process that requires something you have and something you know) usually seek information such as a password/passphrase, date of birth, home address, mother's maiden name, etc. A verification approach with one-dimensional signals, such as the ones proposed here, circumvents this process by employing non-transferrable datasets already native to the user.

- As the demand for long term state-of-health monitoring increases, medical sensors implemented on personal, wearable, or implanted platforms demand strict rules of

engagement to improve system interconnectivity and reliability so that they can be seamlessly woven into the user environment without requiring additional user intervention. Owner-aware sensors, or devices that recognize their owners based on an assessment of the data sets acquired from those individuals, are an appealing idea because they bolster security in the environment, minimizing their impact on normal human behavior, and increase the viability of the monitoring, diagnosis, and treatment process.

- Although the performance of these identification algorithms requires improvements for large populations, some of these one-dimensional signals do offer fairly accurate classifications when the subject population is relatively small. This points to their feasibility for environments such as homecare settings or community health centers, both of which are vital to an aging population. In these applications, health data could be constantly or periodically collected, so identification performance deteriorations caused by long-term signal alterations are expected to be minimized. Indeed, these signal alterations may themselves provide trend data as an additional means to distinguish individuals.

6. Conclusion

This chapter recognizes several important questions that arise upon completion of a comprehensive review of existing research work that explores the possibility of using ECGs as waveforms for human identification. It answers one of these questions by investigating how identification performance changes as a function of subject population size. Using three wavelet coefficient-based distances to measure the similarity/difference between unknown datasets and those in a database, consistent performance trends were obtained from the three discrimination cases, confirming that accuracy declines as the population grows. This finding is a reminder that, although ECG-based authentication holds potential for applications where ECG data have already been acquired or stored, caution is needed when the population size is large.

7. References

Agrafioti, F., & Hatzinakos, D. (2008a). ECG Based Recognition Using Second Order Statistics. *The 3rd International Symposium on Communications, Control and Signal Processing*, 2008.

Agrafioti, F., & Hatzinakos, D. (2008b). Fusion of ECG sources for human identification. *The 3rd International Symposium on Communications, Control and Signal Processing*, 2008.

Bao, B., Zhang, Y., & Shen, L. (2005). Physiological Signal Based Entity Authentication for Body Area Sensor Networks and Mobile Healthcare Systems. *Proceedings of the 27th Annual IEEE Engineering in Medicine and Biology Conference*, 2455-2458.

Biel, L., Pettersson, O., Philipson, L., & Wide, P. (1999). ECG analysis: a new approach in human identification. *Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference*, 1999.

Biel, L., Pettersson, O., Philipson, L., & Wide, P. (2001). ECG analysis: A new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(3), 808–812.

Boumbarov, O., Velchev, Y., & Sokolov, S. (2009). ECG personal identification in subspaces using radial basis neural networks. *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, 2009.

Chan, A. D. C., Hamdy, M. M., Badre, A., & Badee, V. (2006). Person Identification using Electrocardiograms. *Canadian Conference on Electrical and Computer Engineering*, 2006.

Chan, A. D. C., Hamdy, M. M., Badre, A., & Badee, V. (2008). Wavelet distance measure for person identification using electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 57(2).

Chaoui, J., Cyr, K., Gregorio, S. d., Giacalone, J.-P., Webb, J., & Masse, Y. (2001). Open multimedia application platform: enabling multimedia applications in third generation wireless terminals through a combined RISC/DSP architecture. *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*.

Chiu, C.-C., Chuang, C.-M., & Hsu, C.-Y. (2008). A Novel Personal Identity Verification Approach Using a Discrete Wavelet Transform of the ECG Signal. *International Conference on Multimedia and Ubiquitous Engineering*, 2008.

Chuo, Y., Marzencki, M., Hung, B., Jaggernauth, C., Tavakolian, K., Lin, P., et al. (2010). Mechanically Flexible Wireless Multisensor Platform for Human Physical Activity and Vitals Monitoring. *IEEE Transactions on Biomedical Circuits and Systems*, 4(5), 281.

Daugman, J. (1998). Phenotypic Versus Genotypic Approaches to Face Recognition *Face Recognition: From Theory to Applications* (pp. 108–123): Springer-Verlag.

Doi, J., & Yamanaka, M. (2004). Biometric authentication using finger and palmar creases. *2004 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement System*.

Duc, B., Bigun, E. S., Bigun, J., Maitre, G., & Fischer, S. (1997). Fusion of audio and video information for multi-modal person authentication. *Pattern Recognition Letters*, 18(9), 835–843.

Elsherief, S. M., Allam, M. E., & Fakhr, M. W. (2006). Biometric Personal Identification Based on Iris Recognition. *2006 International Conference on Computer Engineering and Systems*.

Fatemian, S. Z., & Hatzinakos, D. (2009). A new ECG feature extractor for biometric recognition. *16th International Conference on Digital Signal Processing*, 2009.

Faundez-Zanuy, M. (2005). Biometric verification of humans by means of hand geometry. *39th Annual 2005 International Carnahan Conference on Security Technology*.

Gahi, Y. L., M.; Zoglat, A., Guennoun, M., Kapralos, B., & El-Khatib, K. (2008). Biometric Identification System Based on Electrocardiogram Data. *The New Technologies, Mobility and Security*, 2008.

Galeottei, L., Paoletti, M., & Marchesi, C. (2008). Development of a low cost wearable prototype for long-term vital signs monitoring based on embedded integrated wireless module. *Computers in Cardiology*, 905 - 908.

Gu, Y., Zhang, Y., & Zhang, Y. T. (2003). A Novel Biometric Approach in Human Verification by Photoplethysmographic Signals. *4th Annual IEEE Conference on Information Technology Applications in Biomedicine*, UK.

Gu, Y., & Zhang, Y. T. (2003). Photoplethysmographic authentication through fuzzy logic. *The IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*.

Irvine, J., Wiederhold, B. K., Gavshon, L. W., Israel, S., McGehee, S. B., Meyer, R., et al. (2001). Heart rate variability: A new biometric for human identification. *Proceedings of International Conference on Artificial Intelligence*, 1106–1111.

Israel, S., Irvine, J. M., Cheng, A., D.Wiederhold, M., & K.Wiederhold, B. (2005). ECG to identify individuals. *Pattern Recognition Society*, 38, 133-142.

Israel, S., Scruggs, W. T., Worek, W. J., & Irvine, J. M. (2003). Fusing Face and ECG for Personal Identification. *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop (AIPR)*.

Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4- 20.

Jovanov, E., Poon, C., Yang, G.-Z., & Zhang, Y. T. (2009). Guest Editorial Body Sensor Networks: From Theory to Emerging Applications. *IEEE Transactions on Information Technology in Biomedicine*, 13(6), 859 - 863.

Kharate, G. K., Ghatol, A. A., & Rege, P. P. (2007). Selection of mother wavelet for image compression on basis of image. *IEEE Signal Processing, Communications and Networking*, Anna University, Chennai, India.

Kim, K.-S., Yoon, T.-H., Lee, J.-W., Kim, D.-J., & Koo, H.-S. (2005). A Robust Human Identification by Normalized Time-Domain Features of Electrocardiogram. *27th Annual IEEE International Conference of the Engineering in Medicine and Biology Society*.

Koh, L. H., Ranganath, S., Lee, M. W., & Venkatesth, Y. V. (1999). An integrated face detection and recognition system. *International Conference on Image Analysis and Processing*.

Kyoso, M. (2003). A technique for avoiding false acceptance in ECG identification. *Proceedings of IEEE EMBS Asian-Pacific Conf. Biomedical Engineering*, 190–191.

Kyoso, M., & Uchiyama, A. (2001). Development of an ECG identification system. *Proceedings of the 23rd IEEE Engineering in Medicine and Biology Confonerence*, 4, 3721-3723.

Love, J., Warren, S., Laguna, G., & Miller, T. (1997). Personal Status Monitor, SAND97-0418, DOE Category UC-706, February 1997, 199 pages.

Ludeman, L. & Chacon, M. (1995). Evaluation of Blood Pulse Signature for Potential Application in a Multisensor Biometric Identity Verification System. Final Report, Sandia National Laboratories, Contract No. AM-5506 Amendment No. 1, September 14, 1995, 16 pages.

Ludeman, L. & Chacon, M. (1996). Use of Blood Pulse Signature for Identity Verification. *Proceedings of the 7th International Conference on Signal Processing Applications & Technology*, October 7-10, 1996, Boston, Massachusetts, USA, pp. 1608-1612.

Ma, T., Zhang, Y., & 7hang, Y. (2006). *Biometrics, Wyle Encyclopedia of Biomedical Engineering*, April 14, 2006, pp. 1-13, http://onlinelibrary.wiley.com/doi/10.1002/9780471740360.ebs0164/pdf.

Mallat, S. (1999). *A wavelet tour of signal processing* (2 ed.): Academic Press.

Marcel, S., & Millan, J. D. R. (2007). Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. *IEEE Transactions on Information Technology in Biomedicine Pattern Analysis and Machine Intelligence*, 29(4), 743 - 752.

Matsumoto, T., Matsumoto, H., Yamada, K., & Hoshino, S. (2002). *Impact of artificial gummy fingers on fingerprint systems.* Paper presented at the Proc. SPIE.

Matyas, V. J., & Riha, Z. (2003). Toward reliable user authentication through biometrics. *IEEE Security Privacy*, 1(3), 45–49.

Micheli-Tzanakou, E., Plataniotis, K., & Boulgouris, N. (2009). Electrocardiogram (ECG) Biometric for Robust Identification and Secure Communication. *Biometrics : Theory, Methods, and Applications* (pp. 383 - 427).

Ming, J., Hazen, T. J., Glass, J. R., & Reynolds, D. A. (2007). Robust speaker recognition in noisy conditions. *IEEE Transactions on Information Technology in Biomedicine Audio, Speech, and Language Processing*, 15(5), 1711 - 1723.

MIT-BIH Database Distribution. from http://ecg.mit.edu/ Harvard-MIT Division of Health Sciences and Technology.

Nasri, B., Guennoun, M., & El-Khatib, K. (2009). Using ECG as a measure in biometric identification systems. *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*.

Philips, P. J., Grother, P., Micheals, R. J., Blackburn, D. M., Tabassi, E., & Bone, J. M. (2003). Face recognition vendor test 2002: Overview and summary, from http://www.frvt.org/FRVT2002/documents.htm

Plataniotis, K. N., Hatzinakos, D., & Lee, J. K. M. (2006). ECG Biometric Recognition Without Fiducial Detection. *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*.

The PTB Diagnostic ECG Database. from http://www.physionet.org/physiobank/database/ptbdb/, Physikalisch-Technische Bundesanstalt (PTB)

Saechia, S., Koseeyaporn, J., & Wardkein, P. (2005). Human Identification System Based ECG Signal. *TENCON 2005*, 2005 IEEE Region 10.

Shen, T. W., Tompkins, W. J., & Hu, Y. H. (2002). One-lead ECG for identity verification. *Proceedings of the 2nd Joint EMBS/BMES Conference*, 62–63.

Singh, Y. N., & Gupta, P. (2008). ECG to Individual Identification. *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008.

Sufi, F., & Khalil, I. (2008). An automated patient authentication system for remote telecardiology. *International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2008.

Sullivan, T. J., Deiss, S. R., & Cauwenberghs, G. (2007). A low-noise, non-contact EEG/ECG sensor. *IEEE Biomedical Circuits and Systems Conference*, 2007, Montreal, Canada.

Wan, Y., Sun, X., & Yao, J. (2007, Oct. 17-20). Design of a photoplethysmographic sensor for biometric identification. *International Conference on Control, Automation and Systems 2007*, Seoul, Korea.

Wan, Y., & Yao, J. (2008, 22-24 October). A neural network to identify human subjects with electrocardiogram signals. *The World Congress on Engineering and Computer Science (WCECS) 2008 of International Association of Engineers (IAENG)*, San Francisco, CA.

Wang, Y., Plataniotis, K. N., & Hatzinakos, D. (2006). Integrating Analytic and Appearance Attributes for Human Identification from ECG Signals. *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*.

Warren, S., & Jovanov, E. (2006). The need for rules of engagement applied to wireless body area networks. *3rd IEEE Consumer Communications and Networking Conference*, 2006 (CCNC 2006).

Warren, S., Lebak, J., Yao, J., Creekmore, J., Milenkovic, A., & Jovanov, E. (2005). Interoperability and Security in Wireless Body Area Network Infrastructures. *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2005. IEEE-EMBS 2005. .

Webster, J. G. (1998). *Medical instrumentation: Application and design*: Wiley.

Yao, J., Sun, X., & Wan, Y. (2007, 23rd - 26th August). A pilot study on using derivatives of photoplethysmographic signals as a biometric identifier. *29th International Conference of the IEEE Engineering in Medicine and Biology*, Lyon, France.

Yao, J., & Wan, Y. (2008, June 1-3). A wavelet method for biometric identification using wearable ECG sensors. *5th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2008)*, Hong Kong.

Yao, J., & Wan, Y. (2010). Improving Computing Efficiency of a Wavelet Method Using ECG as a Biometric Modality. *International Journal of Computer and Network Security*, 2(1), 15-20.

Zhang, G. H., Poon, C. C. Y., & Zhang, Y. T. (2009). A biometrics based security solution for encryption and authentication in tele-healthcare systems. *2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, 2009 (ISABEL 2009).

Zhang, Z., & Wei, D. (2006). A new ECG identification method using Bayes' Theorem. *TENCON IEEE Region 10 Conference*, Hong Kong.

Electromagnetic Sensor Technology for Biomedical Applications

Larissa V. Panina

School of Computing & Mathematics, University of Plymouth,
United Kingdom

1. Introduction

Magnetic bio-detection constitutes a large area of research and development driven by its potential to provide versatile diagnostic tools in biology and medicine. Specific sensing technology is used depending on the applications which can be subdivided in two main groups: measuring a magnetic field from people and detecting magnetically labelled bio substances. The human body is mostly composed of what is normally regarded as nonmagnetic materials. In reality, every substance has some magnetic sensitivity, however small, being paramagnetic or diamagnetic. Their response is greatly limited by thermal fluctuations. In addition to this there is a further source of a magnetic field due to the neural activity which operates continuously throughout a living body. This neural activity involves movement of electric charges and, as such gives rise to magnetic fields. In principle, these fields represent a description of the neural activity and can be studied to help understanding the workings of the human body as well as provide an aid to diagnosis. On the other hand, they can be predicted and quantified by the fundamental laws of electromagnetic. In the other stream of applications, the use of magnetic labels allows the detection of various bio molecular reactions in immunoassays. It also results in a number of additional functionalities, such as transport of bio-molecules to a specific location, on-chip magnetic immuno-separation and testing or accelerating bio-molecular binding events.

The magnitudes of the magnetic fields involved are very small, being in the sub-nanoTesla region and their detection requires very sensitive instrumentation. We have extremely sensitive magnetic technology: SQUID magnetometer (superconducting quantum interference device). The noise level of the SQUID detection is in femto-Tesla so it could become an ideal instrument for studying magnetic fields from biological subjects (see, for example, Sternickel & Braginski, 2006). However, the cost involved and the complex cryogenic technology present huge hurdles that have prevented SQUID (including high transition temperature requiring liquid nitrogen) from becoming widely used. Several field measurements from various parts of the body were published and summarised in (Wikswo , 1999). Figure 1 presents comparison of some common magnetic fields and those generated by different parts of the human body. The signals from the brain are at about 1 picoTesla or less but from other parts of the body (such as the eyes and the stomach) are at levels an order of magnitude larger. One of the largest signals results from the heart which is at the level of about 25pT. The detection of this level of a magnetic field does not require all the

potential of SQUIDs and other high performance magnetometers (Ripka, 2001) can be used. Here we will discuss the potential of using as biosensors such magnetometers as giant magnetoresistive (GMR) and giant magnetoimpedance (GMI) sensors placing emphasis on a relatively new GMI sensing technology which could overcome many limitations of SQUID and magnetoresistive sensing platforms. As far as the detection of magnetic labels is concerned, we will consider the advantages of the detection method based on non linear magnetisation of magnetic labels through generation of high frequency harmonic spectra often referred to as magnetic particle spectral method (MPS). In both GMI and MPS the dynamic magnetisation processes are involved and we categorise them as electromagnetic bio sensing platform.

	B(Tesla)	
Common noise	10^{-7}	Magnetic fields from people
	10^{-8}	
	10^{-9}	Lung particles
Car at 50 m	10^{-10}	Heart
Screwdriver at 5 m	10^{-11}	
	10^{-12}	Fetal Heart Eye
IC transistor chip at 2 m	10^{-13}	Brain
	10^{-14}	

Fig. 1. Comparison of some common magnetic fields with those from human body.

Magnetoresitive sensing technology benefits from recent research and technological advances aiming the design of ultra high density magnetic memory hard discs and write/read heads (Prinz, 1998; Wolf 2001). However, limited sensitivity prevents the use of GMR sensors for measuring human magnetic fields, but they have been used to detect a variety of commercially available magnetic micro and nano beads as a basis for biochip development. The detection of molecular recognition (the interaction of complementary or affinity-linked biomolecules) with GMR has been shown. The sensor element can be made of a fraction of micron in size and can in principle provide the detection at the level of a single molecular interaction (Gaster et al, 2009). Yet, the GMR technology for biosensing is regarded as long-term development owing to the need to improve either sensitivity or highly spatial resolution. It requires precise manipulation of magnetic beads that must be transported to the sensor location. This approach makes the technology very specific, not suitable for a wider range of application, and in fact prevents its transfer from laboratory to end-user. GMI offers at least an order of magnitude higher magnetic field resolution (down to 10 PicoTesla) than GMR but is larger in size with a single sensing area limit of about 5μm×50μm at frequencies of 20-250MHz. Integration of both GMR and GMI sensing concepts could be a possibility.

The discovery of GMI effect in 1993 (Panina & Mohri, 1994; Beach & Berkowicz, 1994) had a strong impact on the development of micro magnetic sensors operating in the nano-Tesla

range (Zhukova et al, 2009; Mohri et al, 2001, 2002). In certain soft magnetic materials, such as composites of amorphous thin wires, the impedance change (GMI ratio) is in the range of more than 100 % in the MHz frequency band for the external magnetic fields of 0.1mTesla. One of the main activities in improving GMI technology is devoted to developing miniature GMI systems preserving high sensitivity. This could be realised in ultrathin amorphous/nanocrystalline wires and multilayer structures with inner silver/gold lead. In particularly, the modelling results for CoFeSiB/Au/CoFeSiB multilayers with in-plane size of 20μm×100μm and thickness of 1μm demonstrate the GMI ratio of more than 300% at 350 MHz (Morikawa et al, 1997; Hika et al, 1994). Further miniaturisation and high GMI ratios are proposed to be achieved employing structures having a special spiral type of anisotropy (Panina et al, 2001). GMI sensors and particular their arrays provide the needed sensitivity for magnetic cardiography (MCG) and some laboratories have already reported successful detecting cardiac signals with the use of GMI (Uchiyama et al, 2009). An MCG allows details of the localised electrical activity of the heart to be revealed, enabling accurate diagnosis of heart conditions. The development of room temperature MCG would be of great importance for wide practical use.

We further will discuss the use of non linear magnetisation of magnetic labels for their sensitive detection. Magnetic labels or carriers, also referred to as microspheres, microbeads and nanoparticles, have found wide-ranging scientific and clinical application in biotechnological and biomedical research (Haukanes & Kvam, 1993), most notably in the areas of bioseparations, molecular biology and drug delivery. The labels used are non-remanent paramagnetic or superparamagnetic beads. The magnetic material within the label exists as small particles (usually iron oxide), having random moments. A conventional magnetic sensor system relies on the alignment of these moments within the label to produce a measurable fringe field. On the other hand, the magnetisation of such magnetic particles is a nonlinear function of external magnetic field and capable of generating high frequencies harmonics having high signal to noise ratio, thus, high detection sensitivity.

Therefore, the purpose of this chapter will be to promote relatively new magnetic sensing technologies for use in biomedical field. GMI and MPS can offer very high sensitivity (sub nano-Tesla range), low cost, with the benefit of a portable operation. In many areas of medical applications, such as MCG and magnetic immunoassays, the importance of these sensing technologies cannot be overestimated as they may provide relatively simple room-temperature tests, which could be realised in any general hospital, unlike such expensive methods as SQUID magnetometry.

2. SQUID magnetometers for biosensing applications

SQUID instrument depends on the effects of magnetic fields on circuits containing superconducting (Josephson) junctions. The temperature at which these operate is traditionally at about 4° Kelvin which is the temperature of liquid Helium. A Josephson junction, which is the essential part of a SQUID device, is effectively a weak link between two superconductors that is capable of carrying supercurrents below a critical value. There are two types of SQUIDs: rf SQUIDs and dc SQUIDs. In both types, the device consists of a superconducting ring interrupted by one (rf) or two (dc) Josephson junctions. The difference between the two is in the nature of the biasing current being an rf or a dc. In either type, the properties of the Josephson junction cause the impedance of the ring to vary periodically as a function of the magnetic flux crossing the ring. With the use of a lock-in detector to

measure the impedance, SQUIDs can operate as flux-to-voltage converters with highest magnetic sensitivity of any magnetometer in existence. Recently, with the discovery of high-temperature superconductors, SQUID systems at liquid nitrogen temperatures (e.g. 70° Kelvin) have been built although the noise level is significantly higher than that of the liquid helium systems (Mahdi & Mapps, 2000). A typical principle of operation is to detect two magnetic signals at different distances from the source and arrange for these to be in opposition around a local supercooled circuit. This 'gradiometer' approach eliminates most of the noise – caused by spurious magnetic fields originating from, for example, electrical devices or natural (geomagnetic) sources. A typical unshielded laboratory has a noise level in the dc to ten hertz frequency region of about 10^{-7} Tesla so a magnetically screened environment is usually required for these studies. The basic requirement has been for a noise level of the measuring environment to be lower than the sensor limit and this has caused the design and construction of sophisticated screened rooms in hospitals environments (Ter Brake, 1991). The screened rooms currently in use employ large amounts of mu-metal and are generally built as hospital rooms where patients can walk in and sit under comparatively large SQUID detector systems. For smaller systems this will not be necessary since screening is actually only needed in the detection zone, for example around the head. For this a screened 'hat' is envisaged which will be at much lower cost. Also various techniques are available for providing noise cancellation such as at the National Physical Laboratory, Teddington, U.K. where 'negative' noise is generated and added to the noise in the measurement zone by a Helmholz coil system (Hall, 2001).

A single measurement of a field, for example, near the head has only a very limited value in neural analysis of the brain. Therefore, multiple SQUID systems have been devised and used in an attempt to build up a picture of the increased neural brain activity that results, for example, when the biological subject of the study has been stimulated in some way. The development of multiple SQUID systems has also led to an acceleration of research on MEG systems having, typically, 150 SQUID sensors arranged in a framework around the head. A great deal of research on this has been done by Professor Ueno in Japan (Iramina et al, 2001; Iramina et al, 1997). Similar multiple detection systems have also been developed for EEG in the past and an interesting comparison can be made with MEG.

3. GMR for biosensing applications

Giant magnetoresistive (GMR) sensors as biosensors are mainly developed in conjunction with magnetic immonoassays. Magnetoresistive-based biochips were first introduced in 1998 by the Naval Research Laboratory (NRL) and since then an increasing number of research laboratories and companies have been developing such systems (Graham et al., 2004; Megens & Prinz, 2005; Tamanaha et al., 2008; Xu et al., 2008, Martinsa et al, 2009). The use of small magnetic particles or beads in biomedical sciences has increased significantly in recent years (Haukanes & Kvam, 1993). Integration of sensor design with magnetic-particle and assay development is a significant part of this research. Magnetic labels usually are superparamagnetic or non-remanent ferromagnetic in nature, with nano- or micrometer dimensions, and can attach to the target biomolecules. Under an applied magnetic field these particles or beads acquire a moment and their fringe field can induce a change in resistance of the magnetoresistive sensor, enabling biomolecular recognition detection.

The change in material resistance, which occurs when the magnetisation changes from parallel, with respect to the direction of current flow, to transverse is known as anisotropic

magnetoresistive effect (AMR). AMR is present in ferromagnetic alloys such as NiFe, NiFeCo, but the resistance change is small. A large change in resistance up to 70% (Baibich et al, 1988) is based on the spin dependent interfacial and bulk scattering asymmetry that is found for spin-up and spin-down conduction electrons crossing ferromagnetic–nonmagnetic–ferromagnetic multilayer structures, where the parallel or antiparallel alignment of the ferromagnetic layers can be engineered. Then, the resistance of two thin ferromagnetic layers separated by a thin nonmagnetic conducting layer can be altered by changing the moments of the ferromagnetic layers from parallel to antiparallel. Layers with parallel magnetic moments will have less scattering at the interfaces, longer mean free paths, and lower resistance. Layers with antiparallel magnetic moments will have more scattering at the interfaces, shorter mean free paths, and higher resistance. For spin-dependent scattering to be a significant part of the total resistance, the layers must be thinner than the mean free path of electrons in the bulk material. For many ferromagnets the mean free path is tens of nanometers, so the layers themselves must each be typically thinner than 10 nm. There are various methods of obtaining antiparallel magnetic alignment in thin ferromagnet-conductor multilayers. The structures currently used in GMR sensors are unpinned sandwiches, antiferromagnetic multilayers and spin valves.

Ti10W90	15 nm
Ta	2 nm
Mn76Ir24	2.5 nm
Co90Fe10	2.5 nm
Cu	2 nm
Co90Fe10	2.5 nm
Ni80Fe20	3 nm
Ta	1.5 nm

Fig. 2. Spin-valve multilayer system.

The structure of the spin-valve sensor used in (Martinsa et al, 2009) is shown in Fig. 2. The pinned and free layers are deposited with the easy axes in parallel orientation. The shape anisotropy is then used to rotate the easy axis of the free-layer at 90◦ to get a linear response. The sensor element is patterned by direct write laser photolithography and ion milling, resulting in U-shaped sensors with a final active area of 2.5×80 μm². Patterned sensors have an average magnetoresistance of 7.5%. A differential sensor set-up uses a reference sensor in the Wheatstone bridge architecture to enable thermal and electrical (mains) drift compensation between a biologically active sensor and a biologically inactive sensor.

The fields required to change the magnetisation direction in one of the layers could be in kOe range resulting in a reduced sensitivity. When used as biosensors, the needed sensitivity was obtained thanks to the combination of detection and manipulation of magnetic particles. The secret of the sensitivity improvement lies within the location of the magnetic markers on the chip via active guiding of magnetic particles using on-chip generated magnetic forces. First, a sandwich assay is built up on the device surface, followed by labelling with magnetic particles. Then, the bound magnetic particles are released and transported to the position that theoretically gives rise to the maximal signal, ensuring the most sensitive detection. This process is schematically demonstrated in Fig. 3.

Magnetic label detection has been accomplished by using different types of integrated GMR sensor designs, such as stripes, meander, spirals, and serpentine-shaped GMR sensors. Different shapes of GMR sernsors were tried in an attempt to optimize the sensor active surface, averaged stray magnetic fields and on-chip manipulation and transport of magnetic beads.
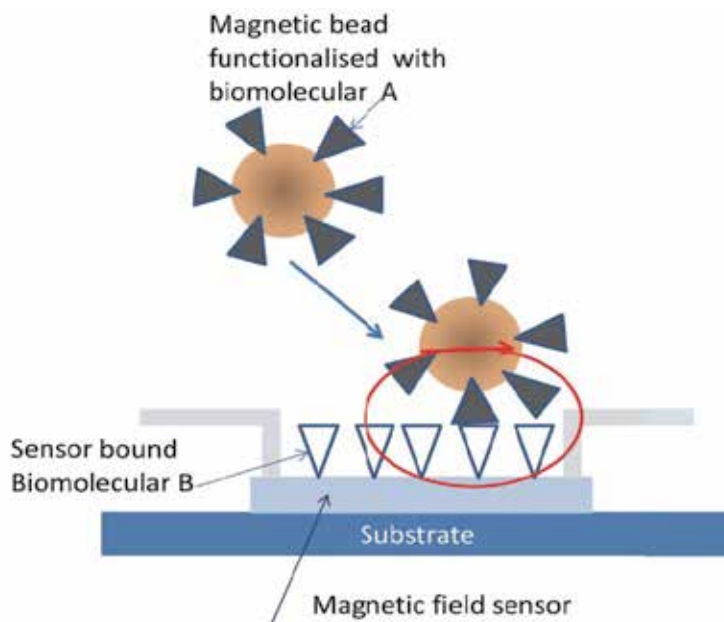

Fig. 3. Schematic of a biochip which is composed of GMR sensors, an array of probe biomolecules (biomolecules B of known identity) immobilized onto the surface of the sensors, magnetic labels functionalized with target biomolecules (A) that bind to the sensor surface through biomolecular recognition. The magnetic fringe field resulting from the magnetic moment of the label induced by an on-chip applied magnetic field changes the resistance of the sensor, resulting in a voltage signal.

Typical target analyte resolutions today are near 1pM for passive detection (INESC MN) using spin valve sensors (few $nT/Hz^{1/2}$) at thermal background. The sensor element covers partially the probe immobilization area (few hundred μm^2). Increase in sensitivity is being pursued by moving towards low noise magnetic tunnel junction based platforms. INESC MN has demonstrated field resolutions down to few tens $pT/Hz^{1/2}$ (Martinsa et al, 2009). Although single molecule process detection is within reach by reducing the sensor size towards the magnetic label dimension, for practical applications, the challenge resides in increasing sensitivity to allow detection of few sub 100 nm labels with a dynamic range up to few thousand labels on a point of care portable device. Furthermore, the use of labelled targets has allowed the use of current lines for magnetically assisted hybridization. INESC-MN has been exploring this ability to enhance the sensitivity and specificity via active guiding of magnetic beads using on-chip generated magnetic forces, hoping to reach detection sensitivity into the fM range (Martinsa et al, 2009). This approach makes the technology very specific, not suitable for a wider range of application, and in fact prevents its transfer from laboratory to end-user.

4. Biosensor based on non linear magnetization detection

One of the most promising approaches in developing magnetic biosensors associated with the use of magnetic labelling is based on non-linear magnetisation of magnetic beads. Non-linear magnetisation processes result in the generation of higher order harmonics which can be detected and discriminated for use in remote sensing and monitoring (Ong & Grimes, 2002). Typically, as biological labels, nanosized magnetic particles are used which are in a single domain state and demonstrate superparamagnetic behaviour with an essential non-linearity. The non-linear magnetisation permits easy discrimination of these particles from surrounding paramagnetic materials and their reliable detection by high harmonic spectrum techniques for many biological applications. The magnetization of non-interacting particles obeys the Langevin function as shown in Fig. 4. If a system of such magnetic particles is excited by a harmonically oscillating magnetic field of frequency ω its magnetisation response $M(t)$ is non-linear and contains integer multiples of ω, so-called higher harmonics. This non linear magnetisation response can be detected inductively and evaluated spectroscopically. Therefore, this detection method is often referred to as magnetic particle spectroscopy (MPS). The induced inductive voltage is proportional to the rate of the magnetisation change:

$$V(t) \propto \frac{\partial M}{\partial t}$$

This represents the measurable quantity which also reflects the magnetisation response's degree of distortion: it is clearly not harmonic anymore. The excitation magnetic field, the response functions and the voltage signals are depicted in Fig. 5.
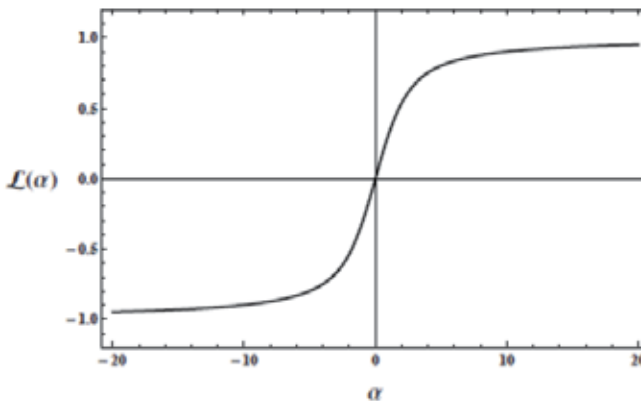

Fig. 4. Plot of the Langevin function representing the relative magnetization of a system of noninteracting and freely rotatable magnetic moments.

The next step is to obtain the detected signal's frequency spectrum, i.e. resolving the harmonic contributions and their amplitudes. Therefore, the time signal has to be transformed into the frequency domain by a Fourier transformation. Generally, the exact calculation of Fourier series coefficients is too difficult or impossible. Instead, $V(t)$– no matter if it was gained synthetically (analytically/ numerically) or in an experiment – is analysed numerically using a discrete Fourier transformation (DFT). Therefore, $V(t)$ has to be sampled with a specific rate f_{sample}, yielding a set of data points $V_i(t_i)$. Then, DFT transforms $V_i(t_i)$ into $A_i(\omega_i)$, i.e. it resolves the complex amplitude A_i for a specific frequency ω_i.
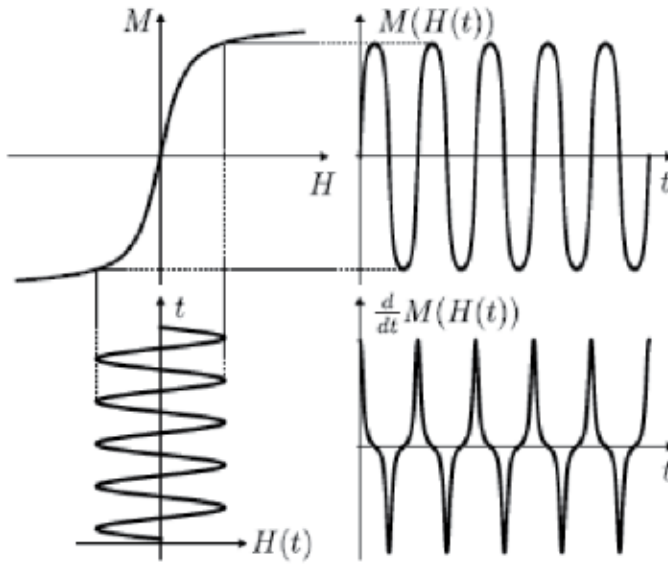
Fig. 5. Harmonic magnetic excitation *H(t)* of superparamagnetic particles. Transfer function *M(H)(*top left). Magnetization response *M(t)* (top right). Response's time derivative, which is proportional to measurable voltage *V(t)* (bottom right).
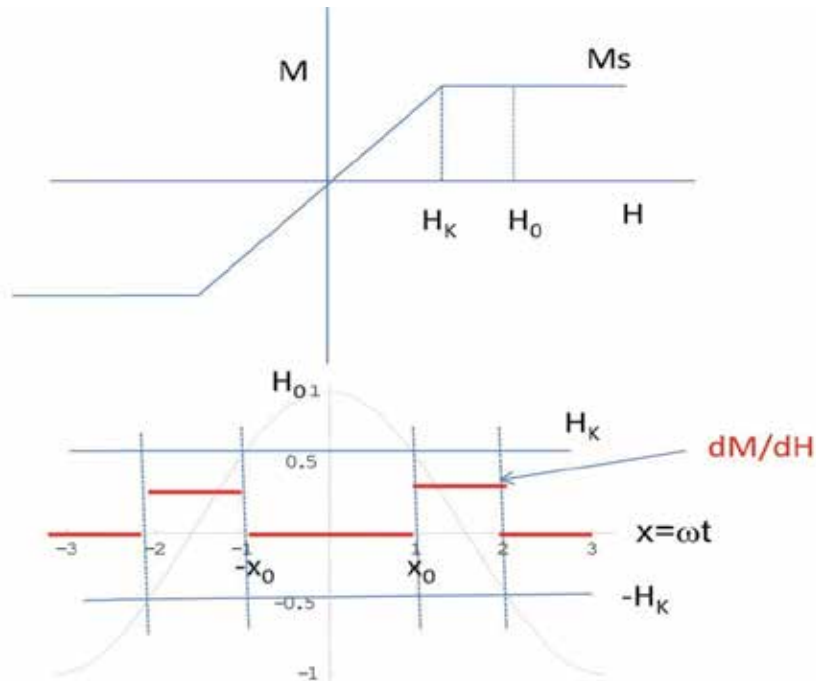

Fig. 6. Idealised response function $M(H)$: linear for $|H| < H_K$ and constant for $|H| > H_K$ (top). Magnetic excitation and step-wise susceptibility $\partial M/\partial H$ (bottom).

For illustration of the method, an idealised response function: linear with saturation, as shown in Fig. 6 is considered. The magnetization saturates when the magnetising field exceeds a characteristic value, denoted by H_K. Suppose, that the magnetising field is changing as $H = H_0 cos\omega t$. Then, the voltage signal is proportional to

$$V \propto \frac{\partial M}{\partial H}\frac{dH}{dt} = H_0\omega \sin \omega t \frac{\partial M}{\partial H} \tag{1}$$

The case of interest is when the amplitude of the magnetising field is sufficient to reach the saturation: $H_K < H_0$ for which $\partial M/\partial H$ is a step-wise function:

$$\frac{\partial M}{\partial H} = \begin{cases} \dfrac{M_S}{H_K} = \chi, & x_0 < x = \omega t < \pi - x_0, \ \ x_0 = \cos^{-1}(H_K/H_0) \\ 0, & otherwires \end{cases} \tag{2}$$

Since the voltage signal in (1) is an odd function of time, the Fourier spectra will be represented by sin-series:

$$V(x) = \sum_{n=1}^{\infty} A_n \sin nx , \qquad -\pi \le x = \omega t < \pi. \tag{3}$$

In this case, the harmonic coefficients A_n can be expressed analytically for any n:

$$A_1 \propto \frac{2H_0\omega\chi}{\pi} \int_{x0}^{\pi-x0} \sin x \ \sin x dx = \frac{H_0\omega\chi}{\pi}[\pi - 2x_0 + (\sin 2x_0)] . \tag{4}$$

$$A_{2k+1} \propto \frac{H_0\omega\chi}{2\pi}\left[\frac{\sin 2k(\pi - x_0) - \sin 2k(x_0)}{k} - \frac{\sin 2(k+1)(\pi - x_0) - \sin 2(k+1)(x_0)}{(k+1)}\right]. \tag{5}$$

In (5), $= 1,2,...$. The spectra calculated from (4) and (5) for different values of parameter $h = H_0/H_K$ which reflects the degree of non-linearity are shown in Fig. 7.
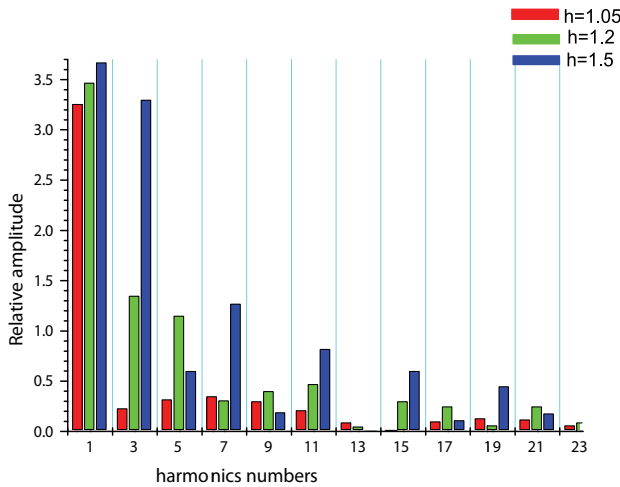

Fig. 7. Harmonic spectra for magnetisation response function of Fig. 6 and described by Eqs. (1), (2).

It is seen that certain high –order harmonics preserve relatively large values (as harmonics 19 and 21 for $h = 1.5$) and can be used for determination of the magnetic particle concentration. It is also interesting to notice that the spectra have specific characteristics depending on parameter h, such as non monotonic decrease with increasing n, which can be used for magnetic particle discrimination in multiparameteric analysis. The spectra will change dramatically in the presence of off-set dc magnetic field. Utilising spatially dependent off-set field, imaging techniques with non-linear magnetic particles can be developed (Gleich & Weizenecker, 2005; Weizenecker et al, 2009). The achieved resolution is well below 1 mm and the method has the potential to be developed into an imaging method characterised by both high spatial resolution as well as high sensitivity.
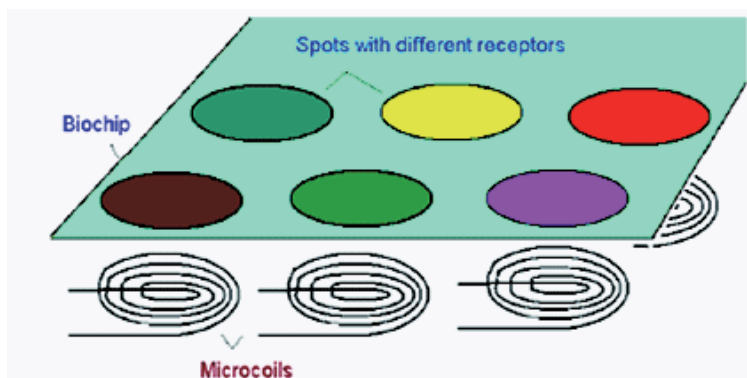

Fig. 9. Arrays of inductive coils and immobilization zones for multi component immunoassay.

Further, magnetic particles with non-linear response can be subjected to multi tone excitation. It was proposed to use the interrogation magnetic field with two frequencies, ω_1 and ω_2, the first one being considerably higher (Nikitin et al, 2007, 2008). The amplitude of the lower frequency tone is chosen high enough to get a strong non-linearity of $M(H)$. For example, this component may periodically 'switch' on and off the capability of magnetic particles to change the magnetisation. When the particle can be further magnetised, the higher frequency component ω_1 contributes to the resulting induction signal. As a result, the response signal $V(t)$ is non-linearly modulated by both frequencies. The spectral response is measured at combinatorial frequencies $\omega_{nm} = m\omega_1 + n\omega_2$, where m and n are integers. This technique results in much higher SNR than in the case of a single-tone excitation. It was successfully applied for multi-component analysis using coil-arrays and zones with different immobilised agents as shown in Fig. 9. Several sensing configurations for superparamagnetic particle detection have been developed in a wide linear dynamic range (3 ng–70 mg) in the volume of 0.1–0.4 cm³. The sensitivity of this type of magnetic immunoassays at the level of 0.1 ng/ml for soluble proteins of LPS from F. tularensis has been demonstrated (Nikitin et al, 2007).

5. Giant magnetoimpedance (GMI)

Giant magnetoimpedance (GMI) has at least one order of magnitude higher sensitivity than GMR and can be developed for measurements of magnetic fields from human body such as the fields from heart and muscles, as well as for magnetic immunoassays.

5.1 Basic principles of GMI

First experiments on GMI dated to 1993 were obtained with amorphous magnetic wires and ribbons utilizing a simple concept of measuring an ac voltage in the presence of a dc magnetic field H_{ex} applied in parallel with the current (Panina & Mohri, 1994; Beach & Berkowicz, 1994). For wires with the composition $(Co_{0.94}Fe_{0.06})_{72.5}Si_{12.5}B_{15}$ having almost zero magnetostriction of -10^{-7} the impressive sensitivity (the impedance ratio per a unit of magnetic field) of up to 100%/Oe at MHz frequencies was quickly realised. Since then, the range of materials exhibiting large and sensitive GMI rapidly expanded including nanocrystalline wires and ribbons, as well as more sophisticated materials as multilaered wires and films (Vazquez et al, 2011; Panina, 2009; Phan & Peng, 2008; Knobel & Pirota , 2002). GMI can be considered as a high frequency analogy of giant magnetoresistance. However, the origin of GMI lies in classical electromagnetism and can be understood in terms of the skin effect in conjunction with the transverse magnetization induced by a passing ac current $i = i_0 exp(-j\omega t)$. For some simple geometries and magnetic structures the impedance can be approximated by an analytical form valid for any frequency. Thus, for a wire with a circular cross section and circular domain structure placed in an axial magnetic field H_{ex} and carrying a current i, the impedance is given by (Panina et al, 1994):

$$Z = \frac{R_{dc}(ka)J_0(ka)}{2J_1(ka)} \tag{6}$$

$$k = \frac{1+j}{\delta_m}, \quad \delta_m = \frac{c}{\sqrt{2\pi\sigma\omega\mu_\varphi}} \tag{7}$$

In (6) and (7), R_{dc} is the dc resistance of the wire, a is the wire radius, J_0, J_1 are the Bessel functions, σ is the conductivity, c is the velocity of light (cgs units are used), δ_m is the penetration depth depending on the wire circular permeability μ_φ, $j = \sqrt{-1}$. The circular permeability should be defined considering the full ac permeability tensor $\hat{\mu}$ from the constituent equation: $b_\varphi = (\hat{\mu}h)_\varphi = \mu_\varphi h_\varphi$, which connects the ac magnetic field h and induction b. Equation (6) is valid if the induced axial induction $b_z = 0$. For high frequencies, when the skin depth is smaller than the wire radius, the impedance becomes inversely proportional to δ_m, and hence, proportional to $\sqrt{\mu_\varphi}$.
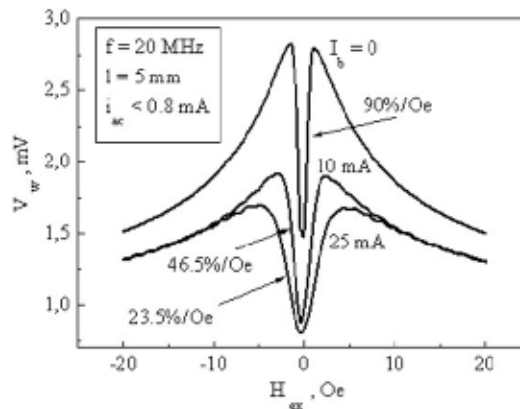

Fig. 10. MI plots in glass-coated CoFeNiSiB wires having a metal core diameter of 48 μm and the total diameter of 50 μm with a dc current as a parameter

Therefore, the changes in permeability which could be few orders of magnitude when the dc magnetisation is rotated from circular direction to the axial direction can be detected by measuring the wire impedance (Garcia et al, 2000). The example of GMI characteristics in amorphous wires is given in Fig. 10 showing the sensitivity up to 90%/Oe. The samples are glass coated wires with well established circumferential anisotropy.

In the MHz frequency range, the parameter μ_φ in general has contributions from domain wall displacements and moment rotation. For sensor applications, it is preferable to eliminate the domain structure in order to avoid Barkhausen jumps noise. For wires with a helical type of anisotropy, this can be done by applying a dc bias current I_b as shown in Fig. 10. However, this current contributes to magnetic hardness so the sensitivity reduces and the value of I_b should be carefully optimized.

For a complicated magnetic configuration which support modes with axial induction b_z, equation (6) becomes invalid and asymptotic or numerical methods should be used to calculate the impedance at arbitrary frequency even for a wire with a circular cross-section (Makhnovsky et al, 2001). However, at high frequencies when the skin effect is strong $(a \gg \delta_m)$, the following asymptotic form for the local surface impedance parameter $\varsigma_{zz} = e_Z/h_\varphi$, which is the ratio of the tangential components of electric and magnetic fields at the surface, can be used:

$$\varsigma_{zz} = \frac{c(1-j)}{4\pi\sigma\delta_0}\left(\sqrt{\tilde{\mu}}\,cos^2\theta + sin^2\theta\right), \quad \delta_0 = \frac{c}{\sqrt{2\pi\sigma\omega}} \tag{8}$$

In (8), the angle θ is the angle between the dc magnetisation and the current direction. The magnetic parameter $\tilde{\mu}$ is the effective transverse permeability with respect to the dc magnetisation. Equation (8) demonstrates clearly the role of the dc magnetisation in determining the high frequency impedance and is very useful for designing the materials with required GMI characteristics.

Substantial amount of works on GMI have been devoted to the asymmetric effects (Panina et al, 2004; Ueno et al, 2004; Delooze et al, 2003; Panina et al, 1999). In the case of sensor applications, the linearity of GMI is an important feature. On the other hand, the GMI characteristics presented in Fig. 10 are not only non-linear, but also shaped in a way that the operation near zero-field point can present serious problems. Generally, a dc bias field is used to set properly the operating point on the GMI characteristics, which can be regarded as producing asymmetry with respect to the sensed field H_{ex}. Therefore, for linear sensing the asymmetrical magnetoimpedance (AMI) is of great importance. There are mainly two ways to realise AMI. The first one is due to asymmetrical static magnetic structure, which can be established in multilayers involving a hard magnetic layer or a layer with a helical anisotropy. The other method is based on the dynamic cross-magnetisation processes. The linear GMI characteristics can be also obtained when detecting the induced voltage from the coil mounted on GMI element when it is excited by a high frequency current (Sandacci et al, 2004). This is based on off-diagonal component of the impedance. This configuration was adopted by Aichi Steel for the development of miniature compass for mobile communication (Mohri & Honkura , 2007; Honkura, 2002).

The condition of a strong skin effect to obtain large GMI may not be required for multilayered systems having an inner conductive lead. If its resistance is considerably smaller than the resistance of the magnetic layers the current mainly flows along the conductive lead. With these assumptions, the expression for the impedance can be written in the form (Hika et al, 1996):

$$Z = R_c - \frac{j\omega\Phi}{ci} \tag{9}$$

where R_c is the resistance of the inner conductor and Φ is the total transverse magnetic flux generated by the driving current i in the magnetic layers. The second term in (9) can be made much larger than R_c in a wide frequency range from MHz to GHz bands in structures with Cu, Ag, Au inner leads and soft magnetic amorphous outer layers of submicron cross section. In particular, multilayered thin films would be of interest for sensing applications in the context of miniaturisation and compatibility with integrated circuit technology.

5.2 Magnetic wires for GMI

Thin amorphous ferromagnetic wires of Co-rich compositions having a negative magnetostriction are very popular for GMI applications (Vazquez et al, 2011; Mohri et al, 2009; Zhukov & Zhukova, 2009). In the outer layer of the wire, an internal stress from quenching coupled with the negative magnetostriction results in a circumferential anisotropy and an alternate left and right handed circular domain structure (Takajo et al, 1993). In this case, the circular magnetization processes determining the GMI behaviour are very sensitive to the axial magnetic field. Along with this, special types of anisotropy as a helical one can be established in the outer layer by a corresponding annealing treatment, which results in unusual asymmetric GMI behaviour. Many experimental results on GMI and designed magnetic sensors utilise amorphous wires of $(Co_{1-x}Fe_x)SiB$ compositions with $x < 0.06$ to decrease the magnetostriction down to -10^{-7} and the characteristic saturation magnetic fields down to few Oe.
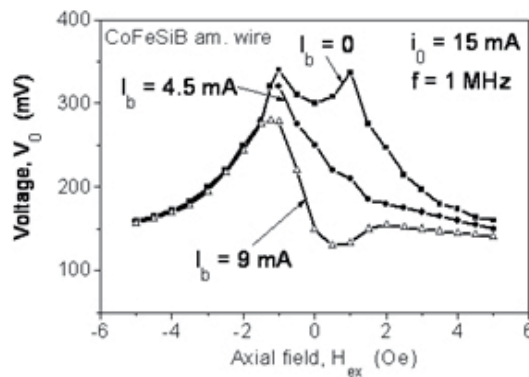

Fig. 11. AMI plots in torsion annealed wires (280 turn/m, 500C) with I_b as a parameter.

Currently, there are basically two methods of wire fabrication techniques. The first one utilises in-water-spinning method for which as-cast wires have a diameter of 125 microns (Ogasawara & Ueno, 1995). The wires then cold drawn down to 20 - 30 μm, and finally annealed under stress to established a needed anisotropy. In the case of negative magnetostrictive wires, annealing under a tension induces a circumferential anisotropy for which the GMI plots shown in Fig. 10 are typical. If the wire is annealed under torsion stress a helical type of the anisotropy is introduced which is important for asymmetrical GMI as shown in Fig. 11 (Panina, 2004).
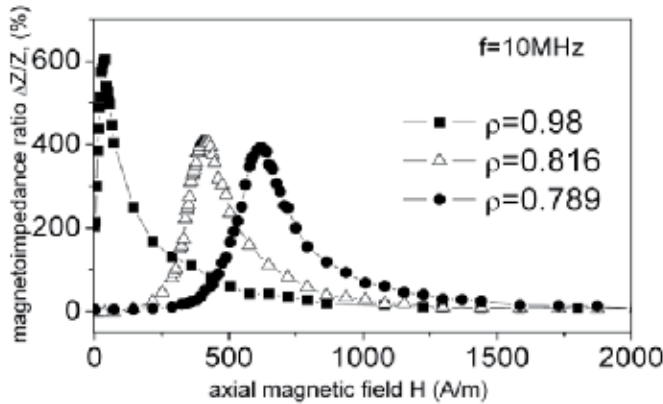
Fig. 12. Sketch of a glass-coated wire.

Fig. 13. GMI plots in $Co_{67}Fe_{3.85}Ni_{1.45}B_{11.5}Si_{14.5}Mo_{1.7}$ glass-coated wires for different values of ρ. d_w is about 22 μm.

The other technique produces amorphous wires in a glass cover (see Fig. 12) by modified Tailor method which is also referred to as Talor-Ulitovsky method (Zhukov et al, 2009; Larin et al, 2002; Zhukov et al, 2000; Chiriac et al, 1996). The method is based on drawing a thin glass capillary with molten metal alloy. The diameter of the metal core is ranging between 1-50 microns and even submicron cross section size is possible (Zhukov et al, 2008). In this case, different temperature expansion coefficients of glass and metal alloy result in a longitudinal tensile stress, which is needed for the circumferential anisotropy. The value of this stress and induced anisotropy depends on the wire composition and ratio $\rho = d_w/D_w$ of the metal core diameter d_w to the total diameter D_w. This is a simple one-step process allowing a strict control of properties in as-cast state and optimisation of the GMI characteristics. Figure 13 (Zhukova et al, 2002) shows the GMI ratio vs. external field for different values of ρ. For a very thin glass layer, the GMI ratio reaches 600% for a field of about 1 Oe at a frequency of 10 MHz. This is the best result obtained so far for any GMI system.

5.3 Multilayered films for GMI
Magnetic/metallic multilayers are especially important for miniaturisation of GMI elements and realising arrays of GMI sensors. The basic structure consists of an inner conductive lead (M) and two magnetic layers (F) as shown in Fig. 14. The magnetic layers could be made of the same alloy, yet, they can be produced with different magnetic anisotropies. In particular, for asymmetric magnetoimpedance the anisotropy axes have to be directed at an angle $\pm\alpha$ to the long (current) axis, respectively for the upper and lower magnetic layers. Such anisotropy can be induced, for example, by current annealing in the presence of a longitudinal field.
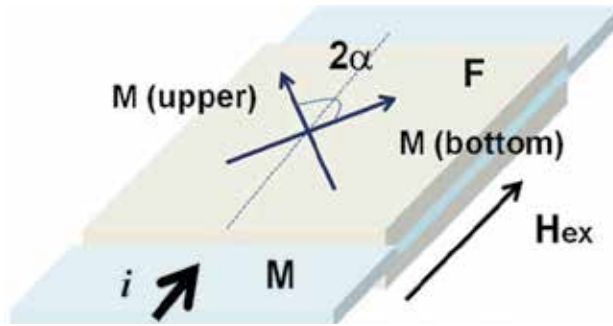
Fig. 14. Sketch of multilayered films. The case of crossed magnetic anisotropy axes in the bottom and top layers makes it possible to obtain almost linear GMR characteristics .

If the edge effects are neglected (i.e., the film is treated as infinitely long in the plane directions) equation (9) for the impedance becomes (Hika et al, 1996):

$$Z = R_c(1 - 2j\tilde{\mu}\frac{d_1 d_2}{\delta_c^2})$$ (10)

Here $2d_1, d_2$ are the thicknesses of the inner and outer layers, respectively, δ_c is the skin depth in the metallic inner lead. Expression (10) shows that the GMI ratio in the sandwich film can be very large even at relatively low frequencies when the skin effect is not substantial, and has a linear dependence on the permeability. This conclusion can be illustrated as follows. At a frequency of 10 MHz, taking $d_1 = d_2 = 0.5\,\mu m$ and $\sigma_{Cu} = 2 \cdot 10^{18}s^{-1}$(conductivity of Copper), we get $d_{1,2}/\delta_c = 0.045$ (so, the skin effect is weak). A typical low-frequency change in the permeability $\tilde{\mu}$ (having a rotational mechanism) under the application of the field equal to the anisotropy field is from 1 to 500; then, the impedance varies over 200% according to (10). In the case of a similar magnetic layer of submicron thickness, the change in the impedance would not be noticeable at these conditions since the skin effect is week. A considerable enhancement of the GMI effect in multilayers can be achieved by insulator separation between the conductive lead and the magnetic films, which prevents layer diffusion and further decreases the DC resistance. A CoSiB/SiO$_2$ /Cu/SiO$_2$/CoSiB multilayer of total thickness of 7 microns exhibited the GMI ratio of 620% for 11 Oe (Morikawa et al, 1997).

For a practical device design, the effect of in-plane sandwich width on GMI has to be studied. If the edge effect is neglected (approximation of an infinite width), the magnetic flux generated by the current flowing along the inner lead is confined within the outer magnetic layers. In a sandwich of finite width 2b, the flux leaks across the inner conductor (Panina et al, 2001). This process eventually results in the considerable drop in GMI ratio, if the film width is smaller than some critical value b^* depending on the transverse permeability and the thicknesses of the magnetic and conductive layers. This process is similar to that resulting in a drop in the efficiency of inductive recording heads and magnetoresistive thin-film devices. In the low-frequency limit, $b^* = \sqrt{d_1 d_2 \tilde{\mu}}$. Typical parameters for the structures of interest are $d_1 \sim d_2 \sim 0.1$-0.5 μm, 2$b \sim$10-50 μm and $\tilde{\mu} \sim 10^3$. This gives the value of $b^* \sim$3-15 μm, which is comparable to the half-width, suggesting that the size effects cannot be neglected. The results of modelling the maximum of the GMI ratio vs. frequency, with a width b as a parameter are shown in Fig. 15. It can be concluded, that even for films narrower than 25 microns very high values of the GMI ratio can be obtained at higher frequency region which was confirmed experimentally for NiFe/Au/NiFe films (De Cos et al, 2005).

The most interesting results on thin film GMI are obtained in amorphous films CoFeB/Cu/ CoFeB with cross-anisotropy (Delooze et al, 2003; Ueno et al, 2000). The films were made on a glass substrate by dc sputtering. During the deposition, a constant magnetic field of 200 Oe was applied in the transverse direction to the GMI element in order to add uniaxial anisotropy. Finally, crossed anisotropy was induced in the sample by current annealing (30 mA) in a longitudinal field of 11.8 Oe at a temperature of 215°C. The anisotropy axes in the upper and bottom layers are at approximately 67° to the long axis, which was estimated from the DC magnetization loop measurements. Applying a bias current produces highly asymmetrical GMI plots as shown in Fig. 16 (Delooze et al, 2004). The differential characteristic from two oppositely biased GMI films is also given, demonstrating a near linear region in the field interval of ± 5 Oe.
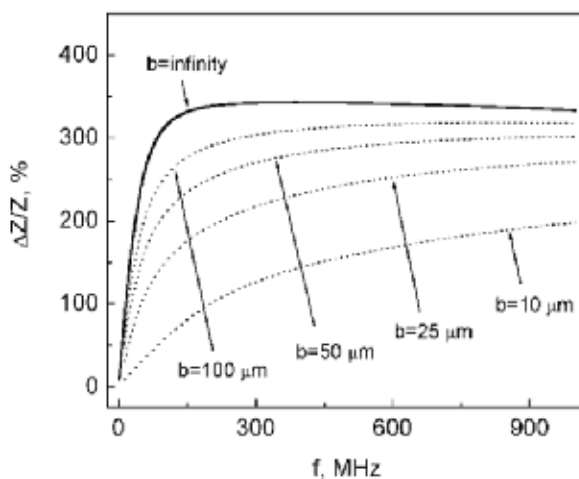

Fig. 15. Maximum of GMI ratio calculated at the anisotropy field, vs. frequency at different values of width b. $2(d_1+d_2) = 1\mu m$, conductivity ratio is 50.
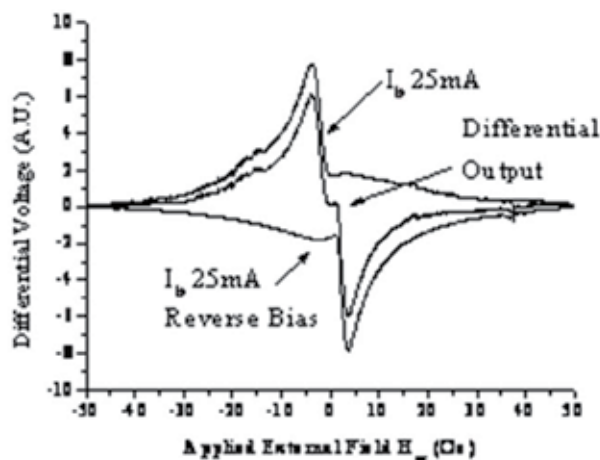

Fig. 16. GMI characteristics in $Co_{70.2}Fe_{7.8}B_{22}$ /Cu/ $Co_{70.2}Fe_{7.8}B_{22}$ films with a DC bias current of 25 mA and their differential output. Magnetic films are amorphous. $d_1 = d_2 = 0.5\ \mu m$.

6. GMI sensor design and biomedical applications

There is a number of high frequency circuits designed to drive GMI elements. One of the most successful solutions utilises a pulse excitation of the GMI element with the help of C-MOS digital circuits(Mohri & Honkura, 2007; Mohri et al, 2002, Shen et al, 1997). The circuit with a C-MOS IC multivibrator as shown in Fig. 17 produces sharp-pulsed current of duration 5–20 ns.
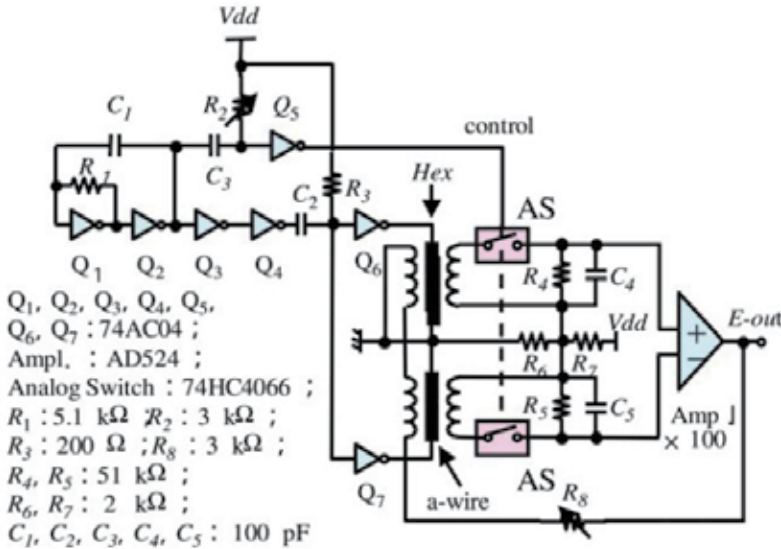
Fig. 17. High-stability CMOS IC with analog-switch type GMI sensor circuit
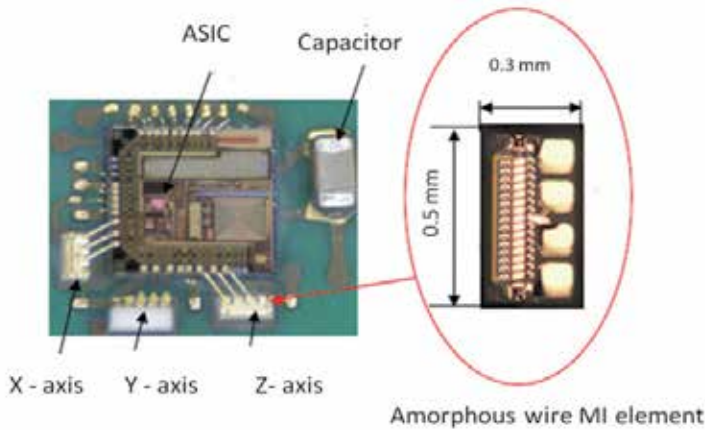
Fig. 18. Three axis electric compass using amorphous wire GMI element.

Pulse excitation provides: simplicity of electronic design, low cost components, and reasonably good stability since C-MOS multivibrator oscillation frequency almost does not depend on the impedance characteristics of the GMI elements. Power consumption of this circuit is also small (10mW). In addition, such pulsed current involves both high frequency

(20–100 MHz) and low (quasi-DC) harmonics. Therefore, it can be ideally used for the asymmetrical GMI requiring a dc or ac bias. Generally, the method provides the field detection resolution of 10^{-4} Oe (10nT) for dc fields and 10^{-6} Oe (100pT) for ac fields. Recently, Aichi Steel Co. has developed GMI-sensor IC-chip (shown in Fig. 18) for mobile phone electronic compass for mass production (Mohri & Honkura, 2007). The restrictions in the field resolution at that level are not due to the intrinsic limitations of advanced GMI elements, but related with the circuit performance. For biomagnetic sensing with GMI further improvements in circuitry are needed, which was achieved with the use of CMOS timer circuit as the multi-vibrator oscillator to effectively reduce circuit noise. In a shielded environment and differential signal amplification, the rms noise was estimated as 3 pT/Hz $^{1/2}$ at 1 Hz (Uchiyama et al; 2009).

In another approach to GMI sensor design, a single frequency excitation is used (Delooze et al, 2005; Yabukami et al, 2004; Yabukami et al, 2001). The sensor measurement system represents a single frequency network analyzer to measure the magnitude of the incident reflected power produced by a mismatch in the complex impedance between the source and load. Typically, GMI thin-film elements are used in this scheme. The incident power or carrier is produced by a Surface Acoustic Wave (SAW) filtered crystal oscillator designed for low power, portable applications. The reflected power is separated from the incident power by means of a directional coupler based on an active op-amp design which provides non-magnetic coupling approach to lower noise. When the impedance of the GMI element matches the impedance of the source (50Ω) the maximum amount of power is transferred to it. With no external field the carrier is suppressed by 60dB. An AC external field causes variation in the 50Ω impedance of the sensor element, which is measured as an AM modulation on the suppressed carrier in the reflected incident power. This is then demodulated, filtered, amplified and measured. For high frequency field detection (>10MHz), the resolution is in the range of pT. Low frequency phase noise (1/f) of the oscillator limits the performance of the sensor at frequencies lower than 1 kHz. A technique to overcome this problem is to firstly modulate (chop) the low frequency AC field to be measured with a locally produced high frequency field (1 to 5 kHz). The second local modulation field shifts the measurement field of interest to a higher frequency offset from the local modulation. This allows the measurement of the low frequency field in the spectrum of the oscillator that is not affected by the phase noise. The achieved ac biased field performance is as following: a 20 Hz field has a resolution of 5.27 x 10^{-6} Oe, and at 10 Hz it is 9.33 x 10^{-6} Oe (Delooze et al, 2004). An improvement of the phase noise of the oscillator, electronics and the use of screening could further increase the performance of the sensor. With these improvements, the GMI sensor technique will be suitable for a wide range of bio-medical applications.

After 10pT resolution of GMI sensors was confirmed in a number of laboratories, GMI-wire and GMI-multilayers sensors were tested for use in the fields of advanced intelligent transport systems, public automation systems, and security systems. It further was recognised that GMI sensors may represent a viable alternative for the conventional biosensors. Firstly, GMI- wire sensors were used in magnetic immunoassays (Chiriac. & Herea, 2007). The sensor system contained a pair of glass-coated amorphous microwires one of them was copolymer-functionalized. The sensor response to the presence of different magnetic microparticles as perturbating agents for the external dc magnetic field was studied. This type of magnetic biosensor prototype was then used for biomolecule detection.

More importantly, the GMI sensor technology was applied for biological magnetic field detection. The GMI sensor based on C-MOS IC with a pair of amorphous wires for differential output was successfully applied for measurements of biocell magnetic fields (Mohri et al, 2009; Uchiyama et al, 2009). This kind of detection can be developed to become, for example, an organ prediction method for iPS cell growth. When compared with the microelectrode method of biocell measurement, the magnetic method benefits from non-invasion operations. The resolution of 10 pT was achieved when measuring the magnetic field generated by a smooth muscle tissue sample (4 mm width, 7 mm length, and 0.3 mm thickness) prepared from a guinea-pig bladder. The sample was dipped in an extracellular solution during the experiment. The pulsed wave was obtained due to Ca^{2+} flow through a membrane. The sample and differential sensor heads were separated by a cover glass. The distance from the sample to the upper sensor head was approximately 1mm. It is known that spontaneous electrical activities in smooth muscle cell clusters have a high temperature dependence (Nakayama et al, 2006). This was confirmed with magnetic detection. Maximum field strength of approximately 1 nT and a cyclic pulse wave can be observed at 33 degrees centigrade. Alternatively, the amplitude of the magnetic field is less than 50 pT and cyclic pulse wave cannot be observed at 27 degrees centigrade.
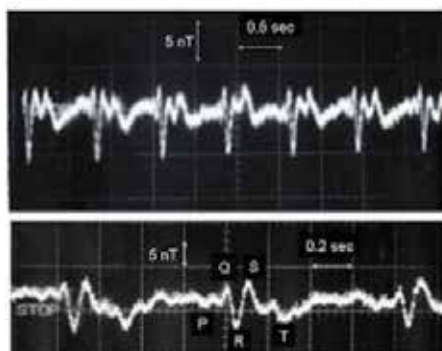

Fig. 19. Waveform of the cardiogram signal measured by CMOC GMI differential sensor in non-shielded environment.

Both GMI-wire and GMI multilayers sensors were used for detecting cardiac magnetic field. The long axis of the sensor head is aligned parallel to the chest surface. The cardiogram signal was clearly detected under the no shield environment when the sensor head is very close to the chest surface. The waveform of the cardiogram is shown in Fig. 19 obtained with GMI-wire sensor(Uchiyama et al, 2009). The features of QRS and T waves correspond with the electrocardiogram. The maximum field strength of several nT is almost 50 times larger than the magneto-cardiogram measurement by a SQUID system using Sensor-to-Chest spacing of approximately 50 mm (Fong et al; 2004).

7. Conclusions

This Chapter summarises some magnetic technology used for detecting and measuring magnetic fields from human subjects and from magnetically tagged bio substances. It develops the argument that there will be an increasing desire for more 'magnetic' information about the human body in the future. The desire for this information will need to further the development of modern magnetic sensor techniques to compete with the now

well-established SQUID magnetometer technology. Measuring and monitoring magnetic fields from the human body is becoming a rapidly increasing subject of research. The very expensive use of the SQUID magnetometer up until now has produced many advances in the understanding of magnetic fields from our bodies. The advance of much cheaper room-temperature sensor technologies offers the prospect of much greater use of magnetic field monitoring in medicine. In such cases it is also important to use low-cost shielding environments or carefully designed noise-suppressed detection systems. Two relatively new magnetic sensing technologies, namely, magnetoimpedance and magnetic particle spectroscopy, have a potential to replace the SQUID magnetometry in such areas as MCG and immunoassays.

8. References

Baibich, M.N.; Broto, J.M.; Fert, A.; Nguyen Van Dau, F.; Petroff, F.; Etiene, P; Creuzet, G.; Friederich, A.; Chazelas, J. (1988). Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* Vol. 61, No 21, November 1988, pp. 2472–2475.

Beach, R.S. & Berkowicz, A.E. (1994). Giant magnetic field dependent impedance of amorphous FeCoSiB wire. *Appl.Phys.Lett.*, Vol. 64, No. 26, Jun 1994, pp. 3652-3654, ISSN 0003-6951.

Chiriac, H. M. & Herea, D.D. (2007) Magneto-impedance sensor for biomedical applications. Journal of Applied Electromagnetics and Mechanics, Vol.25, No.1-4, pp. 453-459, ISSN: 1383-5416.

Chiriac, H. & Ovari, T.A. (1996). Amorphous glass-covered magnetic wires: preparation properties, applications. *Progress in Material Science*, Vol. 40, No. 5, February 1999, pp. 333-407, ISSN: 0079-6425.

D. De Cos, L. V. Panina, N. Fry, I. Orue, A. García-Arribas, and J. M. Barandiarán. (2005). Magnetoimpedance in narrow NiFe/Au/NiFe multilayer film systems. *IEEE Trans. Magn.* Vol. *41*, No 10, October 2005, pp. 3697-3699, ISSN 0018-9496.

Delooze, P.; Panina, L. V.; Mapps, D. J.; Ueno K.; Sano,H. (2004). Sub-nano tesla differential magnetic field sensor utilizing asymmetrical magneto impedance in multilayer films. *IEEE Trans. Magn.* Vol. *40*, No 4, pp. 2664-2666.

Delooze, P.; Panina, L. V.; Mapps, D. J.; Ueno K.; Sano,H. (2003). CoFeB/Cu layered film with crossed anisotropy for asymmetrical magneto-impedance. *IEEE Trans. Magn.* Vol. 39, No 5, pp. 3307-3309.

Fong, L.E.; Holzer, J.R.; McBride, K.K.; Lima, E.A.; Baudenbacher, F.; Radparvar, M. (2004) High-resolution imaging of cardiac biomagnetic fields using a low-transition temperature superconducting quantum interference device microscope, *Appl. Phys. Lett.*, Vol. 84, pp. 3190-3192.

Garcia Prieto, M.J.; Pina, E.; Zhukov, A.P.; Larin, V.; Marin, P.; Vazquez, M., Hernando, A. (2000). Glass-coated Co-rich amorphous microwires with enhanced permeability. *Sensors and Actuators A: Physical*, Vol. 81, No. 1, April 2000, pp. 227-231, ISSN 0924-4247.

Gaster, R. S.; Hall, D. A.; Nielsen, C.H.; Osterfeld, S. J.; Yu, H.; Kathleen E. M., K. E.; Wilson, R. J.; Murmann, B.; Liao, J. C.; Gambhir, S. S.; Wang, S. X. (2009). Matrix-insensitive protein assays push the limits of biosensors in medicine. *Nature Medicine*, Vol. 15, No. 11, November 2009, pp. 1327-1333, ISSN 1078-8956.

Gleich, B. & Weizenecker, J. (2005) Tomographic imaging using the nonlinear response of magnetic particles. *Nature*, Vol. 435, 30 June 2005, pp. 1214-1217.

Graham, D.L.; Ferreira, H.A.; Freitas, P.P.(2004). Magnetoresistive-based biosensors and biochips. *Trends Biotechnol*. Vol. 22, No 9, September 2004, pp. 455–462.

Hall M. (2001). Low field measurements. U.K. Magnetics Society Conference on Magnetic Measurement Techniques and Applications, National Physical Laboratories, U.K., Wed 10th October 2001.

Haukanes, B.I.; Kvam, C. (1993) Application of magnetic beads in bioassays. *Nature Biotechnology*. Vol. 11, No 1, pp. 60–63.

Hika, K.; Panina, L.V.; Mohri, K. (1996). Magneto-Impedance in Sandwich Film for Magnetic Sensor Heads. *IEEE Trans Magn* , Vol. *32*, No. 5, pp. 4594-4596.

Honkura, Y. (2002) Development of amorphous wire type MI sensors for automobile use. *J. Magn. Magn. Mater.* Vol. 249, No. 1-2, August 2002, pp. 375-381, ISSN 0304-8853.

Iramina, K.; Kamei, H.; Yumoto, M.; Ueno S. (2001). Effects of repetition rate of electric stimulation on MEG and fMRI signals. *IEEE Trans. Mag.*, Vol 37, No. 4, July 2001, pp. 2918-2920.

Iramina, K.; Yumoto, M.; Yoshikawa, K.; Kamei, H.; Ueno S. (1997) Measurement of Somatosensory evoked response using functional MR images and MEG. *IEEE Trans. Mag.* Vol. 33, No. 5, September 1997, pp. 4260-4262.

Knobel, K. & Pirota K. R. (2002) Giant magnetoimpedance: concepts and recent progress. *J Magn Magn Mater*, Vol. 242-245, Part 1, April 2002, pp. 33–40.

Larin, V. S.; Torcunov, A. V.; Zhukov, A. P.; Gonzalez, J.; Vazquez, M. & Panina, L. V. (2002). Preparation and properties of glass-coated microwires. *J. Magn. Magn. Mater.*, Vol. 249, No.1-2, August 2002, pp. 39-45, ISSN 0304-8853.

Mahdi, H. & Mapps. D.J. (2000). High-T$_c$ SQUIDS: the ultra sensitive sensors for non-destructive testing and biomagnetism. *European Journal of Sensors and Actuators*, Vol A81, No 1-3, April 2000, pp. 367-370.

Makhnovskiy, D.P.; Panina, L. V. & Mapps, D. J. (2001). Field-dependent surface impedance tensor in amorphous wires with two types of magnetic anisotropy: helical and circumferential, *Phys. Rev. B*, Vol. 63, No. 14, April 2001, pp. 144424-17, ISSN 1098-0121

Martinsa, V.C. ; Cardosoa, F.A.; Germanod, J.; Cardosoa, S.; Sousad, L.; Piedaded, M.; Freitas P.P.; Fonseca. L.P. (2009) Femtomolar limit of detection with a magnetoresistive biochip. *Biosensors and Bioelectronics*, Vol. 24, No. 8, pp. 2690–2695, ISSN 0956-5663.

Megens, M. & Prins, M. (2005). Magnetic biochips: a new option for sensitive diagnostics *J. Magn. Magn. Mater*. Vol.   293, No. 1, May 2005, pp. 702–708.

Mohri, K.; Humphrey, F.B.; Panina, L.V.; Honkura, Y.; Uchiyama, T.; Hirami, M. (2009). Advances of amorphous wire magnetics in 27 Years. *Phys. Stat. Solidi* A, Vol. 206, No. 4, April 2009, pp. 601-607.

Mohri K. & Honkura Y. (2007). Amorphous wire and CMOS IC based magneto-impedance sensors --- Orign, topics, and future. *Sensor Letters*, Vol. 5, No. 2, March 2007 pp. 267-270, ISSN 1546-198X.

Mohri, K.; Uchiyama, T.; Shen, L.P.; Cai, C.M.; Honkura, Y.; Aoyama, H.; (2002). Amorphous wire and CMOS IC-based sensitive micromagnetic sensors utilizing

magnetoimpedance (MI) and stress-impedance (SI) effects, *IEEE Trans Magn* Vol. **38**, pp. 3063–3068.

Mohri, K.; Uchiyama, T.; Shen, L.P.; Cai, C.M. & Panina, L.V. (2001). Sensitive micro magnetic sensor family utilizing magneto-impedance (MI) and stress-impedance (SI) effects for intelligent measurements and controls. *Sensors and Actuators, A: Physical*, Vol. 91 No. 1-2, June 2001, pp. 85-90, ISSN: 0924-4247

Morikawa, T.; Nishibe, Y.; Yamadera, H.; Nonomura, Y.; Takeuchi, M.; Taga, Y. (1997). Giant magneto-impedance effect in layered thin films. *IEEE Trans Magn* . Vol. *33* , No.5, September 1997, pp. 4367-4372, ISSN: 0018-9464.

Nakayama, S.; Shimono, K.; Liu, H.-N.; Jiko, H.; Katayama, N.; Tomita T., & Goto, K. (2006). Pacemaker phase shift in the absence of neural activity in guinea-pig stomach: a microelectrode array study, *J. Physiol.* Vol. 576, No. 3, pp. 727-738.

Nikitin, P.I. ; Vetoshko, P.M. ; Ksenevich, T.I. (2007) Magnetic Immunoassays. *Sensor Letters*, Vol. 5, No. 1, pp 296-298, ISSN 1546-198X.

Nikitin, P.I.; Vetoshko, P.M.; Ksenevich, T.I. (2007) New type of biosensor based on magnetic nanoparticle detection. *J. Magn. Magn. Mater*. Vol. 311, pp. 445-449.

Ogasawara, I. & Ueno, S. (1995). Preparation and properties of amorphous wires. *IEEE Trans. Magn*. Vol. *31,* No. 2, March 1995, pp. 1219-1223, ISSN 0018-9464.

Ong, K .G. & Grimes, C. A. (2002). Tracking the harmonic response of magnetically-soft sensors for wireless temperature, stress, and corrosive monitoring. *Sensors and Actuators A*, Vol. 101, pp. 49–61.

Panina, L.V. (2009). Magnetoimpedance (MI) in amorphous wires: new materials and applications. *Phys. Status Solidi A*, Vol.206, No.4, April 2009, pp. 656-662.

Panina, L. V.; Makhnovskiy, D. P.; Mohri, K. (2004) Magnetoimpedance in amorphous wires and multifunctional applications: from miniature magnetic sensors to tuneable microwave metamaterials, *JMMM*, Vol. 272, pp. 1452-1459.

Panina, L.V.; Zarechnuk, D.; D.P. Makhnovskiy, D. P.; Mapps, D.J. (2001). Two-dimensional analysis of magnetoimpedance in magnetic/metallic multi- layers. *J Appl Phys* , Vol. *89* , pp. 7221-7224.

Panina, L. V.; Makhnovskiy, D. P.; Mohri , K. (1999) Mechanism of asymmetrical magneto-impedance in amorphous wires. *J. Appl. Phys*, Vol. 85, pp. 5444-46.

Panina, L. V.; Mohri, K.; Bushida, K.; Noda, M. (1994). Giant Magneto-Impedance and Magneto-Inductive Effects in Amorphous Alloys. *J. Appl. Phys*, Vol. 76, pp 6198-6203.

Panina, L.V. & Mohri, K. (1994). Magneto-impedance effect in amorphous wires. Appl.Phys.Lett, Vol. 65, No. 9, August 1994, pp. 1189-1191, ISSN 0003-6951

Prinz, G.A. (1998) Magnetoelectronics. *Science,* Vol. 282, pp. 1660–1663.

Ripka, P. Magnetic sensors and magnetometers, Artech House Publishers (2001).

Phan, M. & Peng, H. (2008). Giant Magnetoimpedance Materials: Fundumentals and Applications. *Progress in Materials Science*, Vo.53, No 2, February 2008, pp 323-420.

Sandacci, S. I.; Makhnovskiy, D. P.; Panina, L. V.; Mohri, K. & Honkura, Y. (2004a) Off-diagonal impedance in amorphous wires and its application to linear magnetic sensors. *IEEE Trans Magn.*, Vol. 40, No. 6, November 2004, pp. 3505 – 3511, ISSN: 0018-9464.

Shen, L.P.; Uchiyama, T.; Mohri, K.; Kita, E. & Bushida, K. (1997). Sensitive stress-impedance micro sensor using amorphous magnetostrictive wire. *IEEE Transactions on Magnetics,* Vol. 33, No. 5, September1997, pp. 3355 – 3357, ISSN 0018-9464.

Sternickel, K. & Braginski A.I. (2006) Biomagnetism using SQUIDs: status and perspectives. Supercond. Sci. Technol. Vol. 19, No.3, pp. 3-6.

Takajo, M.; Yamasaki, J.; Humphrey, F.B.; (1993). Domain observations of Fe and Co based amorphous wires, IEEE Trans. Magn., Vol. 29, No 6, pp. 3484-3486, ISSN:0018-9464.

Tamanaha, C.R.; Mulvaney, S.P.; Rife, J.C.; Whitman, L.J. (2008). Magnetic labeling, detection, and system integration. *Biosens. Bioelectron.* Vol. 24, No. 1, September 2008, pp 1–13.

Ter Brake, H. J. M.; Wieringa H. J.; Rogalla, H. (1991). Improvement of the performance of a µ-metal magnetically shielded room by means of active compensation. Measurements Science & Technology. Vol. 2, No. 7, pp. 596-601, ISSN 0957-0233.

Uchiyama T.; Nakayama, S.; Mohri, K.; Bushida, K. (2009). Biomagnetic field detection using very high sensitive magneto-impedance sensors for medical applications. *Physica Status Solidi* (a). Vol. 206, No.4, april 2009, pp. 639-643.

Ueno, K.; Hiramoto, H.; Mohri, K.; Uchiyama, T.; Panina, L.V. (2000) Sensitive asymmetrical MI effect in crossed anisotropy sputtered films. *IEEE Trans. Magn.* Vol. *36*, No. 5, September 2000, pp. 3448-3450, ISSN 0018-9464

Vazquez, M.; Chiriac, H.; Zhukov, A; Panina, L. & Uchiyama T., On the state-of-the-art in magnetic microwires and expected trends for scientific and technological studies. *Phys. Status Solidi A*, Vol. 208, No. 3, March 2011, pp. 493–501 ISSN 1862-6300

Weizenecker, J; Gleich, B.; Rahmer, J.; Dahke, H.; Borgert, J. (2009) Three-dimensional real time in vivo magnetic particle imaging. *Phys. Med. Boil.*, Vol. 54, pp. L1-L10.

Wikswo, J.P. (1999). Application of SQUID magnetometers to biomagnetism and nondestructive evaluation. Applications of Superconductivity, H. Weinstock, ed., Kluwer Academic Publications, 1999.

Wolf, S.A. *et al.*(2001) Spintronics: A spin-based electronics vision for the future. *Science,* Vol. 294, pp. 1488–1495.

Xu, L.; Yu, H.; Akhras, M.S.; Han, S.-J.; Osterfeld, S.; White, R.L.; Pourmand, N.; Wang, S.X. (2008) *Biosens. Bioelectron.* Vol. 24, pp. 99–103.

Yabukami, S.; Mawatari, H.; Murayama, Y.; Ozawa, T.; Ishiyama, K.; Arai, K.I. (2004). High-frequency carrier type thin-film sensor using low-noise crystal oscillator. *IEEE Trans. Magn.* Vol. 40, No. 4, July 2004, pp. 2670-2672, ISSN 0018-9464

Yabukami, S.; Suzuki, T.; Ajiro, N.; Kikuchi, H.; Yamaguchi, M.; Arai, K.I. A high frequency carrier-type magnetic field sensor using carrier suppressing circuit. *IEEE Trans. Magn.* Vol. 37, No. 4, July 2001, pp. 2019-2022, ISSN 0018-9464

Zhukov A. & Zhukova V. (2009). *Magnetic properties and applications of ferromagnetic microwires with amorphous and nanocrystalline structure*, Nova Science Publishers, ISBN: 978-1-60741-770-400, Hauppauge, NY, USA

Zhukov, A.; Ipatov, M.; Zhukova, V.; Garcia, C.; Gonzalez, J. & Blanco, J. M. (2008). Development of ultra-thin glass-coated amorphous microwires for HF magnetic sensor applications. *Phys. Stat. Sol. (A)*, Vol. 205, No. 6, June 2008, pp. 1367-1372, ISSN 1862-6300

Zhukov, A.; Gonzalez, J.; Blanco, J.M.; Vazquez, M. & Larin, V. (2000). Microwires coated by glass: A new family of soft and hard magnetic materials. *J. Mat. Res.*, Vol. 15, No. 10, October 2000, pp. 2107-2113, ISSN: 0884-2914

Zhukova, V; Ipatov, M. & Zhukov A., (2009) Thin Magnetically Soft Wires for Magnetic Microsensors. *Sensors* Vol. 9, No. 11, November 2009, pp. 9216-9240, ISSN 1424-8220

Zhukova, V.; Chizhik, A.; Zhukov, A.; Torcunov, A.; Larin, V. & Gonzalez, J. (2002), Optimization of giant magnetoimpedance in Co-rich amorphous microwires. *IEEE Trans. Magn.*, Vol. 38 No. 5, September2002, pp. 3090-3092, ISSN: 0018-9464

Exploiting Run-Time Reconfigurable Hardware in the Development of Fingerprint-Based Personal Recognition Applications

Mariano Fons and Francisco Fons
Departament d'Enginyeria Electrònica, Elèctrica i Automàtica,
Univeristat Rovira i Virgili, Tarragona
Spain

1. Introduction

The current technological age brings the knowledge and the means to continuously improve the quality of life of human beings. One example can be seen in the recent advances done in the field of biometrics, where those physiological (fingerprints, iris, hand geometry, face, etc.) and/or behavioural (voice, gait, keystroke dynamics, signature, etc.) characteristics of human beings, unique and different to each individual, are used in order to either authenticate or identify individuals in a more reliable way, enhancing thus those existing personal recognition applications based on physical tokens (ID cards, keys, etc.), PINs or passwords. The deployment of automatic biometrics-based personal recognition systems and their acceptance by the society depends on several factors such as the ease of use, the non-intrusive methods of operation and their related privacy concerns; as well as their recognition accuracy, reliability and security levels, response time and system costs. All these factors will determine the successful spread of the biometric security in a wide range of daily use applications such as electronic payment, access systems, border control, health monitoring, etc. all over the world.

Among the different human traits analyzed in the field of biometrics, this work is focused on fingerprints. Fingerprints are the oldest and most deeply used signs of identity. Personal recognition based on fingerprints has been successfully deployed in law enforcement, government, and forensic applications for more than one century. The first recognition systems were based on human experts in charge of matching fingerprints. However, the current technological age demands the development of less expensive and fully automated fingerprint-based personal recognition systems, not only in the cited fields of application but also in many other daily use consumer applications (mobile phones, personal digital assistant devices, laptops, automatic teller machines, internet, e-commerce, etc.). Although big advances have been made in recent years, automatic and reliable biometric recognition is still an open research problem today. That ideal personal recognition algorithm able to unequivocally authenticate the identity of any user from his/her legitimate fingerprint features does not exist. The way to overcome the present limitations and improve the accuracy performance of current biometrics-based authentication systems consists of adding further processing stages into the recognition algorithms, which directly affects the

complexity, the processing power and the costs of the physical systems where to implement those applications.

This works focuses on the search of the proper system architecture able to face those demanding constraints for the application: a high computational power needed to achieve reliable recognition performances in terms of False Acceptance and False Rejection rates (FAR/FRR), a high security level in order to stand any kind of external attacks (cryptographic systems), real-time performance, and low cost. A novel approach of embedded system based on programmable logic devices such as field programmable gate arrays (FPGA), hardware-software co-design techniques, and the exploitation of run-time reconfigurable hardware is proven to successfully address the above requirements.

This chapter is split in nine sections and in each of the sections specific research topics are addressed. Section 2 provides a general overview of the proposed application to be dealt in this work: the development of an Automatic Fingerprint-based Authentication System (AFAS) in charge of verifying the identity of any individual based on the analysis of that distinctive information available in fingerprints. A description of the proposed personal recognition algorithm to be used as reference in this work and to be implemented under different processing platforms is presented. The accuracy performance achieved by the suggested algorithm when evaluated on a large database of fingerprints is addressed in Section 3. One public database composed of up to 800 fingerprint images corresponding to 100 different individuals is used for evaluation purposes. Impostor and Genuine distributions, as well as performance indicators such as FAR, FRR or EER (Equal Error Rate) are given in order to objectively compare the reached performance with the performance of other published algorithms evaluated with the same open database. After presenting the accuracy performance exhibited by the proposed recognition algorithm, Section 4 aims at defining the proper system requirements for the physical platform in charge of the authentication process. The main goal is to find a flexible and high-performance processing platform able to deploy the biometric security in a wide range of daily use applications at low cost, therefore an embedded system architecture is suggested. Two different implementations of the same recognition algorithm are carried out in this work. The first implementation, covered in Section 5, is based on purely software-based solutions. One high performance computing (HPC) platform under Windows operating system and three different embedded system platforms based on low-cost and mid-performance microprocessors are evaluated. The strengths and weaknesses of each of the architectures are pointed out, and based on that information, a different embedded system architecture is suggested in Section 6 to overcome the main limitations exhibited by the previous systems. An embedded system architecture based on a general-purpose microprocessor acting as application core processor, and a programmable and run-time reconfigurable logic region where to instantiate –multiplexed in time and under demand– application-specific hardware coprocessors in charge of the execution of those time-intensive tasks is proposed as alternative solution. Both the microprocessor unit and the hardware accelerators, together with memory blocks and other peripherals are all embedded under a System-on-Programmable-Chip (SoPC) device to provide a highly integrated and more reliable solution. The second implementation of the AFAS application under the proposed embedded system architecture is covered in Section 7. The performance achieved in this new scenario is compared against that of previous scenarios. An outstanding improvement in performance is achieved at a reasonable cost. The work ends with some concluding

remarks in Section 8, and the citation of some research references in Section 9. The reached results prove that the suggested system architecture based on hardware-software co-design techniques under run-time reconfigurable FPGA devices is a cost-effective alternative solution to those existing software-based processing platforms in the deployment of AFAS applications.

2. Fingerprint-based personal recognition algorithm

The personal recognition process is composed of two main phases, as depicted in Fig. 1: the enrolment phase and the authentication phase.
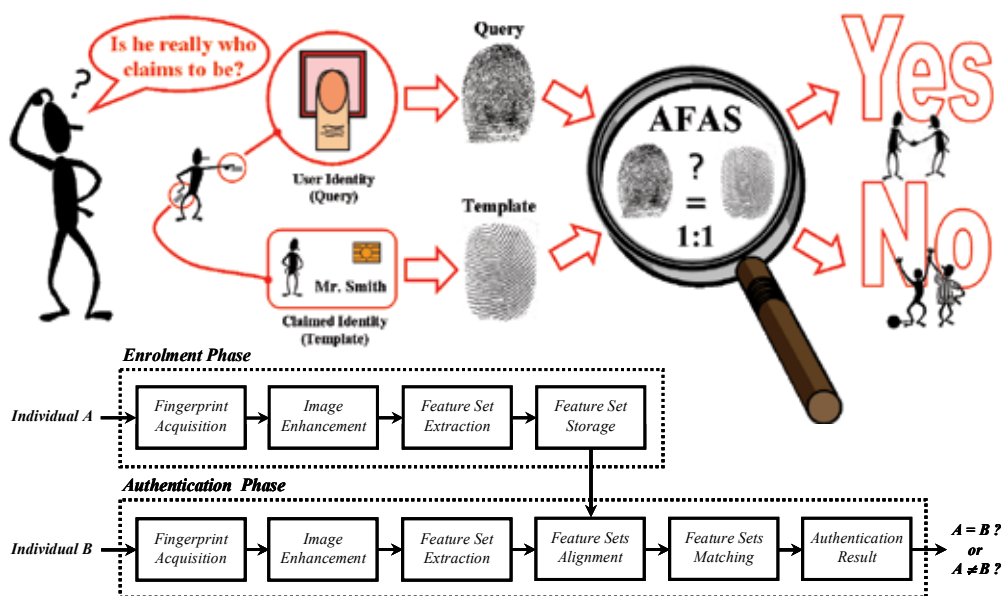

Fig. 1. Processing phases of an Automatic Fingerprint-based Authentication System

The enrolment phase is generally performed off-line, and consists in the registration of that set of biometric features extracted from the digital impression of the user's fingertip –known as *template*– together with any other relevant information of the user within the authentication system, either in a secure database or a personalized smart card. The authentication phase however is normally done on-line, and aims at validating the user's identity by comparing the set of on-line extracted biometric features –known as *query*– against those saved in the authentication system during the enrolment stage and linked to the legitimate individual claimed by the user –*template*–. The matching of both feature sets delivers a similarity score that is used to determine whether the user is really who claims to be, or on the contrary is an impostor who attempts to access the system fraudulently.

As it is indicated in Fig. 1, both phases –enrolment and authentication– are composed of a set of sequential stages. Each of the stages is, at the same time, split into smaller processing operations called tasks, and some of the stages/tasks carried out with the template and query fingerprints are common, as shown in Fig. 2. The aim of the authentication system is the execution of both phases of the processing; therefore the system has to be designed to afford any of the requested tasks along the application.

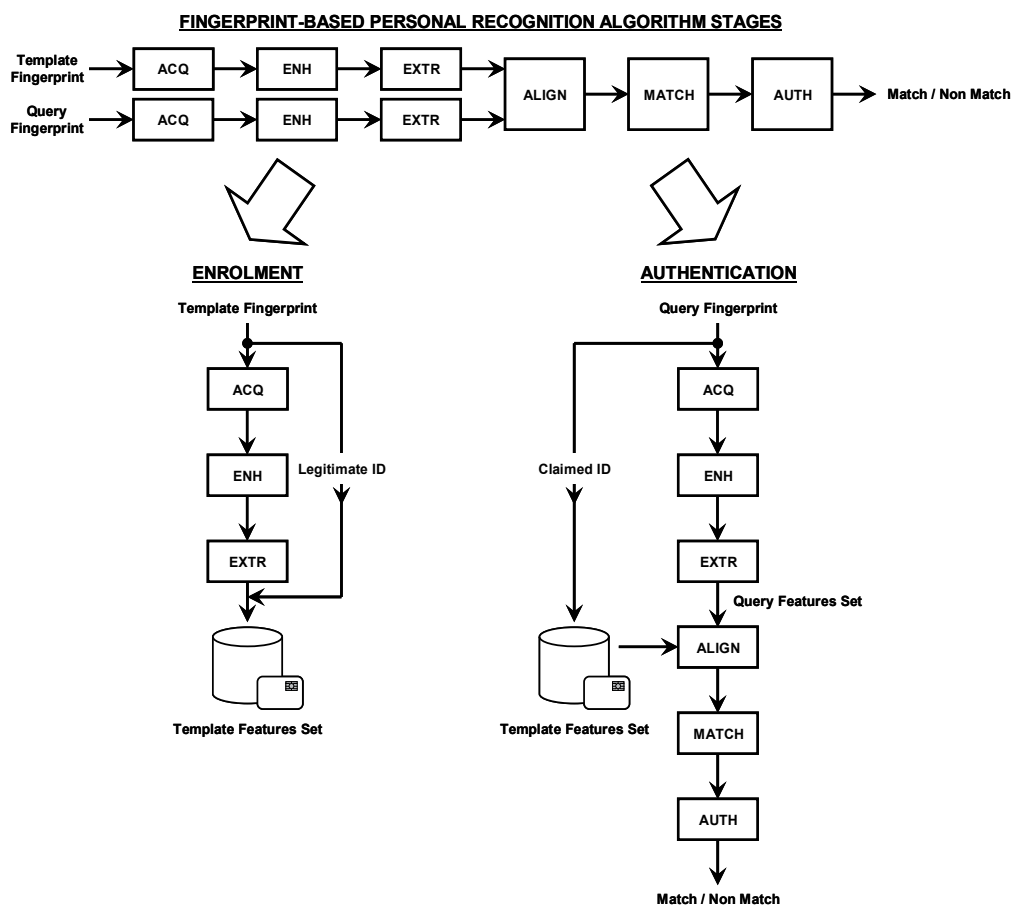**FINGERPRINT-BASED PERSONAL RECOGNITION ALGORITHM STAGES**

Fig. 2. Enrolment and authentication stages decomposition

The proposed recognition algorithm in charge of the enrolment and the authentication processes is not developed from scratch but based on some existing reference biometric algorithms and known techniques well described in the scientist literature. Specific image processing operations like convolutions, filters, etc. and other signal computations in the field of trigonometrics, statistics, etc. are performed on the acquired images in order to deduce that distinctive information available in the fingerprints. For a better understanding of the involved computational tasks refer to the authors' work (Fons et al., 2010). Fig. 3 shows the different processing steps that take place in the suggested fingerprint-based personal verification flow. A hybrid fingerprint matching algorithm that relies on the field orientation map and the set of minutia points extracted from the fingerprints is proposed for its physical implementation. Those classical biometric traits are considered as the genuine marks of identity of any individual. The computational load of the suggested algorithm is equivalent to those other similar or dissimilar algorithms that define the state of the art in fingerprint personal recognition today (Maltoni et al., 2009; Nanni & Lumini, 2009; Yang & Park, 2008).
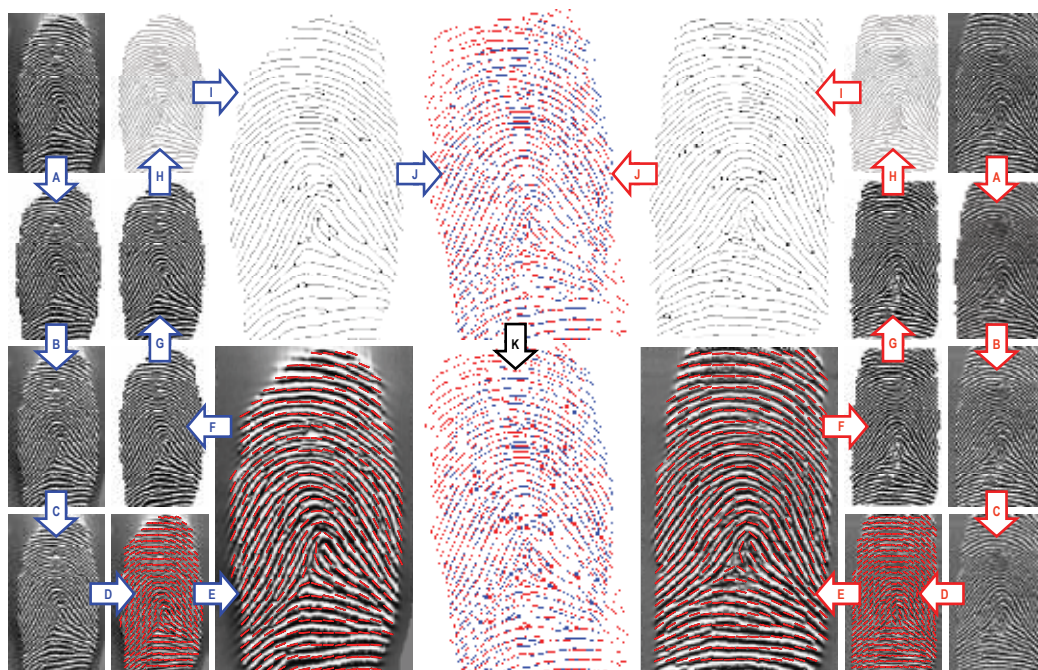
Fig. 3. Intermediate results in the processing of template (left side) and query (right side) fingerprints

A summary of the processing stages involved in the suggested personal recognition algorithm can be deduced from Fig. 3 when authenticating one query fingerprint (right side, red arrows) against one previously enrolled template fingerprint (left side, blue arrows). Up to 11 different tasks (A-K) are carried out along the processing, covering the image enhancement stage (tasks A-G), the feature sets extraction stage (tasks H-I), the feature sets alignment (task J) and the feature sets matching (task K) stages:

- Task A refers to the image segmentation process, which takes as input the acquired fingerprint impression and aims at isolating the valid fingerprint area, also known as foreground, from the rest of the image, also known as background.
- Task B refers to the image normalization process, which aims at adapting the variation of grey level intensities along ridges and valleys in the different regions of the fingerprint.
- Task C refers to the isotropic filtering of the image, which aims at removing some of the hazard noise that could be present in the fingerprint impression.
- Task D refers to the field orientation map computation, which consists in the calculation of the dominant direction of ridges and valleys in each local region of the fingerprint.
- Task E refers to the filtered field orientation map computation, which pursues the enhancement of the previously computed field orientation map.
- Task F refers to the image binarization process, which aims at discriminating ridges and valleys based on the directional filtering of the image according to the enhanced field orientation map.
- Task G refers to the image smoothing process, which aims at enhancing the black and white representation of the image by removing some of the noise that could be present in the binary version of the fingerprint image.

- Task H refers to the image thinning process, which aims at progressively removing the ridge pixels of the image preserving the geometric topology of the ridge-valley pattern till obtaining one skeleton of one single pixel wide to make easy the subsequent identification of minutia points.
- Task I refers to the minutia extraction process, which aims at deducing those salient features spatially distributed along the ridge-valley pattern such as the ridge endings and the ridge bifurcations. Those features will be used as discriminatory information of the fingerprint, together with the filtered field orientation map.
- Task J refers to the image alignment process, which aims at looking for any spatial correspondence between both template and query images based on the extracted feature sets. In case of positive alignment, the overlapped area between both fingerprint impressions is deduced. The overlapped area becomes the region of interest for comparison of template and query prints in the next stage.
- Task K refers to the image matching process and the authentication result (match/non-match) computation based on the comparison of the feature sets (field orientation maps and minutia points) previously aligned.

Most of the cited tasks deal with fingerprint images and/or big amounts of data so a high computational demand is expected for the physical platform in charge of the processing. Although a first implementation of the recognition algorithm under a personal computer platform has been developed in order to validate the accuracy performance reached by the suggested algorithm, more cost-effective system solutions have also been evaluated in this work in order to make easy the spread of those fingerprint-based biometric applications in the consumer arena, accessible to whomever, wherever and whenever.

3. Recognition accuracy performance

In order to prove the validity of the suggested fingerprint recognition algorithm it is needed to proceed with the evaluation of its accuracy performance when submitted to test under a large fingerprint database. The fingerprint recognition algorithm needs to be properly tuned to the environment conditions (fingerprint sensor, sensing technique, attended/unattended acquisition method, etc.) of the real application. The selected database corresponds to the database DB3 of the Fingerprint Verification Competition FVC2004 contest (Maio et al., 2004). This public database is 110 fingers wide, and 8 samples per finger in depth, which results in a total of 880 fingerprint images. All the images were collected by using a thermal sweeping sensor. The complete database is split in two subsets A and B. The subset A is composed of 100 fingers (800 images) and the subset B is composed of 10 fingers (80 images). The subset B is firstly used in order to adjust some of the parameters of the algorithm to the properties of the fingerprint images acquired with the selected sensor, and once the algorithm is properly tuned, the subset A is used in order to verify the real performance of the application. The performance evaluation procedure follows the same criteria than in FVC contests:

i. In order to get the impostor distribution, one sample of each finger in the subset A is collected. A total of 100 images are used, and each of the images is matched against the others to compute the False Match Rate –FMR– or False Acceptance Rate –FAR– distribution. If the matching of g against h is performed, the symmetric one (i.e., h against g) is not executed in order to avoid correlation. A total of 4950 matches are carried out.

ii.   In order to deduce the genuine distribution, each of the samples corresponding to one
      finger is matched against the other samples of the same finger. Similarly to the impostor
      distribution procedure, if the matching of g against h is performed, the symmetric one
      (i.e., h against g) is not executed in order to avoid correlation. The total number of
      genuine tests results in 2800, and from them it is possible to compute the False Non-
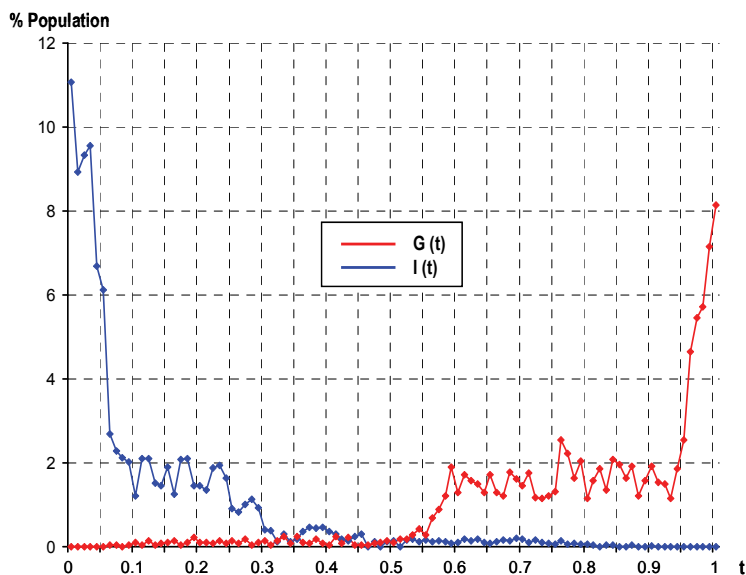   Match Rate –FNMR– or False Rejection Rate –FRR– distribution.

Fig. 4. Genuine and Impostor distributions

Given one template and one query fingerprints, the recognition algorithm provides a
similarity score between both images within [0,1]. Similar images, understood as images
belonging to the same finger, will have scores close to 1, while dissimilar images,
understood as images from different fingers, will present scores close to 0. After
performance evaluation with the subset A, the algorithm features an Equal Error Rate
EER=4.162%. The Genuine and Impostor distributions –I(t) and G(t)–, the representations of
the performance indicator rates FMR and FNMR as a function of the similarity threshold
score t –FMR(t) and FNMR(t)–, and the Receiver Operating Characteristic (ROC) curve of
the tested algorithm are shown in Figs. 4, 5 and 6 respectively.

The parameter EER is the main indicator used to evaluate the performance of the
recognition algorithms in FVC contests. If comparing the performance of the proposed
algorithm against those presented in FVC2004 with the same database, the proposed
algorithm would be ranked in 17th position from a total of 41 participants in the open
category (executed by one personal computer platform without resources constraints),
where the winner algorithm presented an EER=1.18% and the last classified algorithm an
EER=43.95%; or ranked in 5th position from a total of 26 participants in the light category
(executed by a personal computer platform with restrictions on the execution time and the
memory resources), where the winner algorithm presented an EER=2.92% and the last
classified algorithm an EER=54.28%.
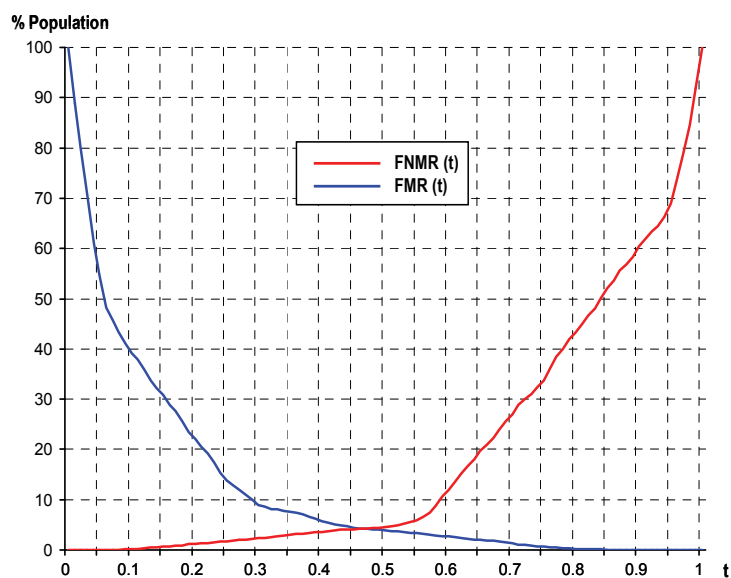
Fig. 5. False Match and False Non-Match distributions
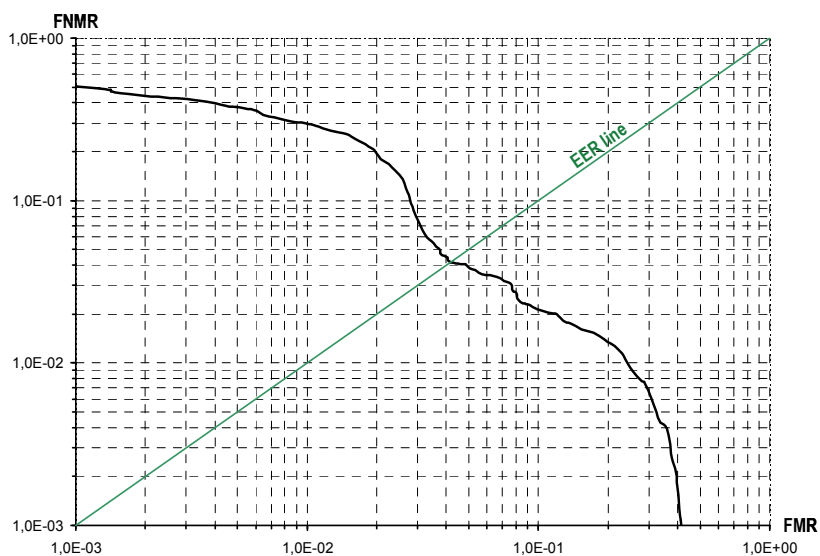
Fig. 6. Receiver Operating Characteristic curve

The first implementation of the recognition algorithm is carried out under a personal computer platform and uses floating point operations in order to be as much accurate as possible in the different computations (statistical analysis parameters like standard deviation, square root calculation, trigonometric computing, etc.) carried out along the recognition process. After proving the validity of the proposed algorithm, a new version of the algorithm is developed by replacing those floating point operations by fixed point

operations in order to reduce the complexity of the processing and the computational demands of the physical platforms where to implement the AFAS application. A new evaluation performance loop of the modified version of the algorithm is performed with very similar results –the EER evolves from 4.162% to 4.242%–. Therefore the new version of the algorithm is also accepted and used as reference to be implemented under low-cost and low-performance microprocessors without floating point units (FPU) on embedded system platforms in the next stage.

4. Application execution time requirements definition

Nowadays most of the applications that exploit biometrics-based personal recognition demand a fast response time to the physical systems in charge of the processing. In case of fingerprint-based authentication systems, soft real-time performance is normally required. In this specific context, soft real-time is understood as providing the proper recognition response within a reaction time short enough to be unnoticed by the user. This reaction time covers the interval elapsed since the user presents his identity credentials to the system and puts his finger on the sensing surface of the capture device till the moment when the automatic authentication system provides the result of the verification process. Reaction times in the range between 1.5s and 3.5s are usually accepted as normal and valid authentication response times for any AFAS application. Therefore, this work focuses on the evaluation of the execution time performance of the proposed fingerprint recognition algorithm when implemented on different computational platforms in order to determine those efficient architectures able to meet the execution time requirements at the lowest possible cost.

Fig. 7. Template and Query fingerprints used in the evaluation process

In order to perform a fair comparison between platforms, the same template and query fingerprints have to be used in all scenarios. Among the different images of FVC2004 DB3 database, two fingerprint impressions taken from the same finger have been selected as template and query fingerprints respectively thus it is possible to build some representative enrolment and authentication processes to be used as reference for evaluation purposes. The two greyscale images depicted in Fig. 7, of size 268x460 pixels and with a resolution of 8 bits and 500 dpi, are used as reference in order to properly compare the same processing effort in all scenarios.

5. Proof of concept I: software-only implementation

Different computational platforms addressing the execution of software-based applications have been selected for processing speed evaluation purposes. The scope covers from high-cost and high-performance personal computer platforms to low-cost and mid-performance embedded system platforms based on general-purpose hard-core or soft-core processors. One personal computer and three embedded system platforms have been evaluated, as indicated in Table 1. The evaluation procedure permits to point out in an easy way which advantages and disadvantages in performance are featured by each of the suggested architectures.

Technical Features	Personal Computer Platform	Embedded System Platform 1	Embedded System Platform 2	Embedded System Platform 3
Platform	Acer Aspire 9420	Altera Excalibur EPXA10	Xilinx Spartan 3AN	Xilinx Virtex4 ML401
Family	MPU Intel Core 2 Duo	SoPC EPXA10F1020C1	FPGA XC3S700AN	FPGA XC4VLX25
Processor	Intel Core 2 Duo T5600	ARM922T	MicroBlaze	MicroBlaze
Processor data bus	64 bits	32 bits	32 bits	32 bits
Number of cores	2	1	1	1
Type of core	Hard-core	Hard-core	Soft-core	Soft-core
Technology	65 nm	180 nm	90 nm	90 nm
Clock speed	1.83 GHz	200 MHz	66.667 MHz	100 MHz
Bus speed	667 MHz	200/100 MHz	133.3/66.6 MHz	200/100 MHz
Cache	2 MB L2	8 KB Inst. Cache	8 KB Inst. Cache 8 KB Data Cache	32 KB Inst. Cache 64 KB Data Cache
Operating system	Windows XP	–	–	–
AFAS program code	DDR2 SDRAM (2 GB)	SoPC SRAM (256 KB)	DDR2 SDRAM (64 MB)	DDR SDRAM (64 MB)
AFAS application data	DDR2 SDRAM (2 GB)	DDR SDRAM (128 MB)	DDR2 SDRAM (64 MB)	DDR SDRAM (64 MB)
SDRAM/SRAM data bus	64 bits	32 bits	16 bits	32 bits
SDRAM frequency	≥ 200MHz	125 MHz	133.333 MHz	100 MHz

Table 1. Computational platforms used in the execution time performance evaluation process

The execution time performance reached in each of the platforms, in both enrolment and authentication stages, is presented in Tables 2 and 3 respectively. The enrolment process of the template fingerprint and the authentication process of the query fingerprint with the enrolled template are evaluated. The authentication execution times are obviously longer

than the enrolment times. Special attention needs to be done to the authentication stage since, unlike the enrolment stage, the authentication process is normally carried out on-line in the real application so real-time response is usually requested. The enrolment stage tends to be less critical since it is normally carried out off-line –under the supervision of application staff to guarantee the reliable enrolment of the user in the system– so no real-time performance is usually demanded.

Task ID	Processing Stage	Personal Computer Platform	Embedded System Platform 1	Embedded System Platform 2	Embedded System Platform 3
Task 1	Image segmentation	2.810 ms	1083.219 ms	299.578 ms	227.035 ms
Task 2	Image normalization	0.470 ms	178.940 ms	46.960 ms	32.772 ms
Task 3	Image isotropic filtering	7.030 ms	5304.010 ms	719.703 ms	467.329 ms
Task 4	Field orientation	2.190 ms	834.062 ms	344.651 ms	244.916 ms
Task 5	Filtered field orientation	0.620 ms	97.061 ms	26.646 ms	17.294 ms
Task 6	Image directional filtering and binarization	13.440 ms	3792.712 ms	860.133 ms	609.518 ms
Task 7	Image smoothing	12.350 ms	1536.114 ms	360.012 ms	229.732 ms
Task 8	Image thinning	1.250 ms	1695.930 ms	547.847 ms	404.085 ms
Task 9	Minutiae extraction and minutiae filtering	0.630 ms	76.626 ms	35.404 ms	23.982 ms
Total Execution Time:		40.790 ms	14598.674 ms	3240.934 ms	2256.663 ms

Table 2. Enrolment process execution time performance

As it can be deduced from the tables, the real-time performance requested to the application is not achieved in all the scenarios. The personal computer platform is able to meet the requested performance, but those other scenarios based on low-cost and mid-performance embedded processors running at low operation frequencies are far away from the requested timing performance. The big latency exhibited by the embedded system platform 1 with regard to the other two embedded system platforms is justified by the fact that no data cache is enabled in that scenario, which severely affects the final performance of the application.

On the one hand, although the powerful processor embedded in the personal computer platform is able to reach the requested performance, its cost is excessive for those low-cost consumer applications demanding biometric recognition. On the other hand, although the embedded system platforms tested in this work are able to meet the system cost requirements of the consumer applications arena, the exhibited execution time performances are clearly insufficient. Therefore, it is needed to find alternative system architectures able to meet both key requirements: high performance and low cost.

Task ID	Processing Stage	Personal Computer Platform	Embedded System Platform 1	Embedded System Platform 2	Embedded System Platform 3
Task 1	Image segmentation	2.810 ms	1083.219 ms	299.578 ms	227.035 ms
Task 2	Image normalization	0.470 ms	178.940 ms	46.960 ms	32.772 ms
Task 3	Image isotropic filtering	7.030 ms	5304.010 ms	719.703 ms	467.329 ms
Task 4	Field orientation	2.500 ms	987.089 ms	407.445 ms	289.661 ms
Task 5	Filtered field orientation	0.620 ms	113.959 ms	30.987 ms	20.171 ms
Task 6	Image directional filtering and binarization	15.940 ms	4460.569 ms	1014.939 ms	720.095 ms
Task 7	Image smoothing	14.220 ms	1752.322 ms	412.503 ms	261.745 ms
Task 8	Image thinning	1.410 ms	1767.383 ms	552.091 ms	402.946 ms
Task 9	Minutiae extraction and minutiae filtering	0.630 ms	93.783 ms	45.002 ms	29.487 ms
Task A	Field orientation maps alignment	3224.530 ms	279636.069 ms	210269.854 ms	138208.006 ms
Task B	Minutiae alignment, feature sets matching and authentication decision	4.220 ms	370.712 ms	161.973 ms	107.972 ms
Total Execution Time:		**3274.380 ms**	**295748.055 ms**	**213961.035 ms**	**140767.219 ms**

Table 3. Authentication process execution time performance

6. Run-time reconfigurable embedded system design

There exist in the market many automatic biometrics-based personal authentication systems implemented on high performance computer platforms –HPCs, PCs, etc.– (One Touch SDK, n.d.; Verifinger SDK, n.d.), embedded general-purpose or application-specific processors – MPUs, MCUs, GPUs, ASSPs– (FxIntegrator, n.d.; plusID, n.d.; SDA, n.d.), embedded digital signal processors –DSPs– (MV1210 and MV1250, n.d.; SFM, n.d.; TMS320, n.d.), or embedded systems based on central processing units –CPUs– plus application-specific hardware accelerators –ASICs– off-chip or on-chip (FPC2020 and FPC-AM3, n.d.; ML67Q5250, n.d.; SecurASIC, n.d.; TCD50D, n.d.). Furthermore, many research articles have been published dealing with the acceleration of some of the stages that take place in one personal recognition algorithm by means of field programmable logic –FPGAs, SoPCs– (Liu-Jimenez et al, 2006; Lopez-Ongil et al, 2004; Pavan Kumar et al, 2007; Yang et al, 2006). However, to the best of the authors' knowledge, up to date there is no work that takes

advantage and exploits the dynamic reconfigurability performance of FPGAs (Becker et al, 2007) in the physical implementation of a complete personal recognition application based on biometrics.

Time-to-market pressures and cost constraints are pushing embedded systems to new levels of flexibility and system integration. In this work, a novel embedded system architecture is proven to successfully address the demands of today's biometrics-based personal recognition systems in terms of computational complexity, real-time performance, development cycles and cost. The proposed embedded system architecture is based on five key factors to afford the challenging demands:

a.   General-purpose microprocessor system.

As in most of the embedded systems in the market today, the usage of low-cost and mid-performance microprocessors (of 16-bits or 32-bits, running at operating frequencies of up to 200-600MHz) provides certain flexibility required in any application. Software-based solutions have additional advantages such as the rapid development of the application by making use of a set of libraries with application-specific functions, which avoids writing the software application from scratch, and provides a cost-effective solution. However, in those applications demanding a high computational power and real-time performance, certain limitations exist when trying to develop the entire application with purely software platforms based on either one single processor (MPU, MCU, DSP, etc.) or multicore/multiprocessor systems due to the inherent limitations in working frequency, restricted data path, shared resources, sequential workflow execution, and reduced parallelism characteristics featured by those standard products.

b.   Programmable logic device embedded in the system.

When purely software-based systems are not enough to meet the expected real-time performances of one real-world application, the usage of hardware-based accelerator devices as complementary processing units has been proven to be an efficient solution. Programmable logic devices such as FPGAs are much more flexible than semi-custom or custom devices like ASSPs or ASICs. ASSPs and ASICs have a fixed peripheral set that limits the number of applications that they can be efficiently used in; but FPGAs allow implementing custom peripherals and made-to-measure glue logic tailored to the requirements of any application. Over recent years, FPGA devices have gained an enormous amount of processing power and functionality thanks to the continuous advances in silicon technologies. The current FPGAs are able to embed much more memory and logical resources, as well as many DSP blocks, multiple clock management units and big amounts of high-speed transceivers for fast communication purposes in one single device. The technology has evolved till the point that the size of today's FPGAs is several orders of magnitude higher than the first FPGAs, reaching values above two millions of flip-flops and LUTs. The programmability performance of FPGAs make them unique in the market and the continuous improvements in the semiconductors field permits reducing the costs of FPGA devices, making them more and more competitive. The flexibility of FPGAs eliminates the long design cycle associated with ASICs, and the usage of IP libraries written in standard hardware description languages and automated design/verification tools reduce the development cycles of those applications based on programmable logic devices.

c.   Hardware-software co-design techniques.

The usage of one general-purpose MPU and one FPGA as a companion chip offers a much greater degree of flexibility and allows the development of any application by means of

hardware-software co-design techniques. The exposed system architecture approach gives flexibility at two levels: at software level, with the MPU-based application management; and at hardware level, with the design of modular cores synthesized in the FPGA.
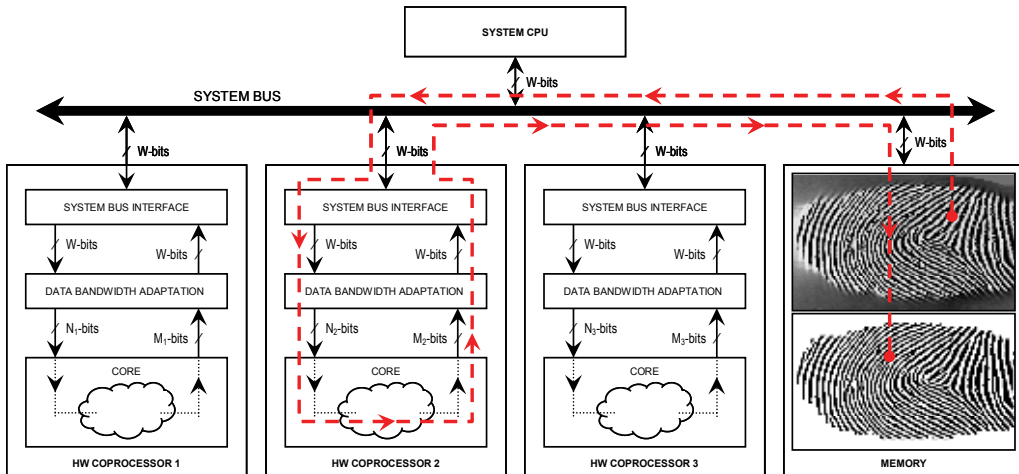

Fig. 8. Physical implementation of one computational platform based on a general-purpose MPU (system CPU), several hardware cores (HW coprocessors) and one memory block. Process execution flow example of one image processing task carried out by one of the application-specific hardware coprocessors instantiated in the system

The FPGA is introduced in the system as a general-purpose device where to instantiate those application-specific hardware coprocessors required to speed up those critical tasks of the application. It permits to design an adaptive and highly-integrated multiprocessor system oriented to the development of real-time applications. Apart from the inherent flexibility featured by the microprocessor, the programmable logic device provides additional flexibility and a high degree of parallelism in the implementation of functional circuits. In the FPGA it is possible to instantiate either additional microprocessors (e.g. VHDL instances of soft-core processors) or made-to-measure VLSI hardware accelerators in charge of specific tasks aiming at offloading those MPU algorithm-intensive operations, as shown in Fig. 8. With an improved bandwidth among the MPU –system CPU-, the FPGA, the memory resources and the rest of peripherals available in the embedded system, soft and hard real-time applications can be successfully developed through this approach.

d.   Run-time reconfigurable FPGAs.

The FPGA device embedded in the system allows exploiting the parallelism and acceleration features inherent to the programmable logic design, so it is possible to meet real-time performance by spreading the functionality across the different core resources (MPU and FPGA) available in the system. However, the resources available in the FPGA are not unlimited, and the cost of those resources increases exponentially when the size of the FGPA increases. Therefore, it is convenient to reduce the size of the FPGA in the design to reach affordable costs for the complete system. In this direction, and owing to the fact that the proposed biometrics-based personal recognition applications feature a sequential

nature (the personal recognition algorithm consists of a set of mutually exclusive image processing tasks executed one after the other), it is possible to exploit the reconfigurability performance featured by some FPGA devices in order to minimize the system hardware needs.

Dynamic partial reconfigurability performance of some existing FPGAs refers to the ability of modifying the functional content of one portion of the FPGA –reconfigurable region– on-the-fly while keeping the rest of the FPGA –static region– fully operative without interruption. The main benefit of doing so is the optimization in the functional density of the device: the same hardware resources available in the reconfigurable region of the FPGA can be time-multiplexed in order to allocate different functionalities (FPGA contexts) along the application execution time. Therefore the amount of needed resources in any application can be minimized, and the total size of the FPGA can be reduced in comparison to the static implementation of all the functionalities instantiated permanently in a bigger FPGA. The main constraint in the usage of run-time reconfigurable FPGAs is the reconfiguration overhead: the time needed in order to modify the functional content of the reconfigurable region in the different contexts. Therefore the minimization of the reconfiguration latencies plays an important role in those systems. Fig. 9 shows the comparison between static and dynamic FPGA-based design concepts.
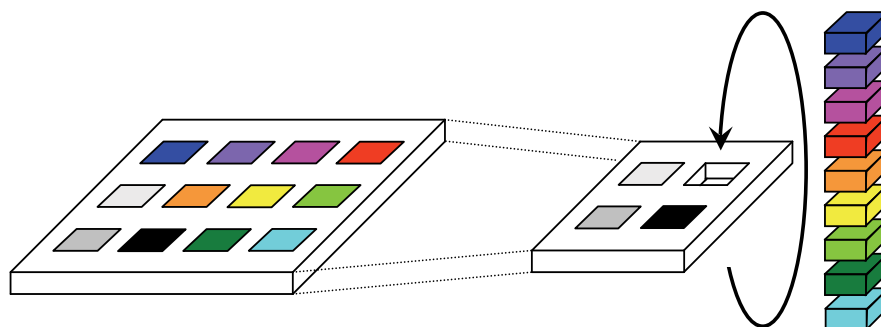

Fig. 9. Comparison between static FPGA-based design concept (left side) and run-time reconfigurable-FPGA-based design concept (right side). The coloured boxes represent each of the different functional blocks in which the application is partitioned

Any application that can be structured as a sequence of mutually exclusive tasks can be proposed to be implemented by means of run-time reconfigurable FPGAs. Fig. 10 shows the scheduling of one application into a sequence of mutually exclusive stages, and the partitioning of each of the processing stages present in the chain into either series or parallel tasks. Each of the tasks can be executed by hardware or by software. Those critical tasks are implemented by hardware to take advantage of higher processing bandwidths and acceleration data path architectures. In this direction, it is possible to make the process truly parallel and at the same time to free some master CPU resources. The rest of less expensive tasks remain as software tasks to be handled by the master CPU of the system. The final partitioning of the application into software tasks, static hardware tasks and dynamically reconfigurable hardware tasks mainly depends on the cost (resources availability, power consumption, etc.) and timing (real-time performance) constraints demanded to the system.
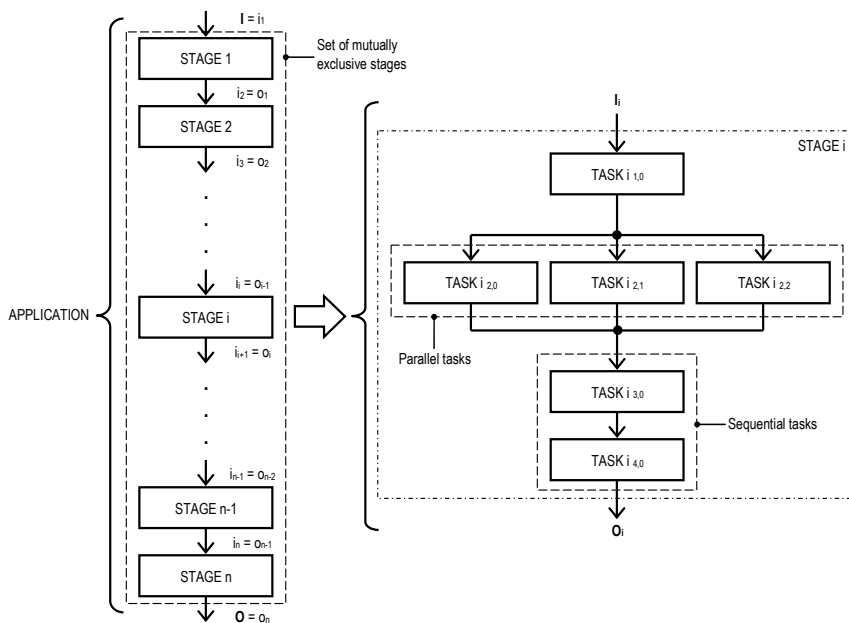
Fig. 10. Deployment of one application as a set of mutually exclusive stages that can be implemented through dynamic reconfigurable embedded systems. Partitioning of each of the stages into hardware and software tasks executed either sequentially or in parallel taking advantage of programmable logic

e.   System-on-programmable chip platform.

The usage of a general-purpose MPU together with programmable and reconfigurable logic gives a high level of flexibility to the system and provides the mechanisms to achieve real-time performance. However, higher integration means lower costs. Therefore, the integration of those main resources and other key peripherals such as memory, timers, interrupt controllers, etc. on a single chip provides an efficient way of optimizing the whole system cost. Embedded biometric recognition is therefore possible by making use of highly integrated platforms. Additional benefits of the system integration are the improvements in reliability and security. It is possible to embed most of the processing in a single SoPC device well-protected against external attacks by means of security protocols and cryptographic processors dealing with the exchange of information between the SoPC device and the external world. For this reason, the usage of SoPC or system-on-chip devices that embed one FPGA is especially encouraged in the experimental tests carried out in this work.

The suggested system architecture is depicted in Fig. 11. At least one run-time reconfigurable region is present in the programmable logic device to synthesize those flexible application-specific hardware coprocessors that can be dynamically instantiated on demand along the application execution time. One specific reconfiguration controller is in charge of the reconfiguration task, supervised by the master processing unit. The AFAS application is connected to the external world by means of a series or parallel communication link with a Host. All or some of the functional blocks depicted in Fig. 11 are embedded in the same chip.
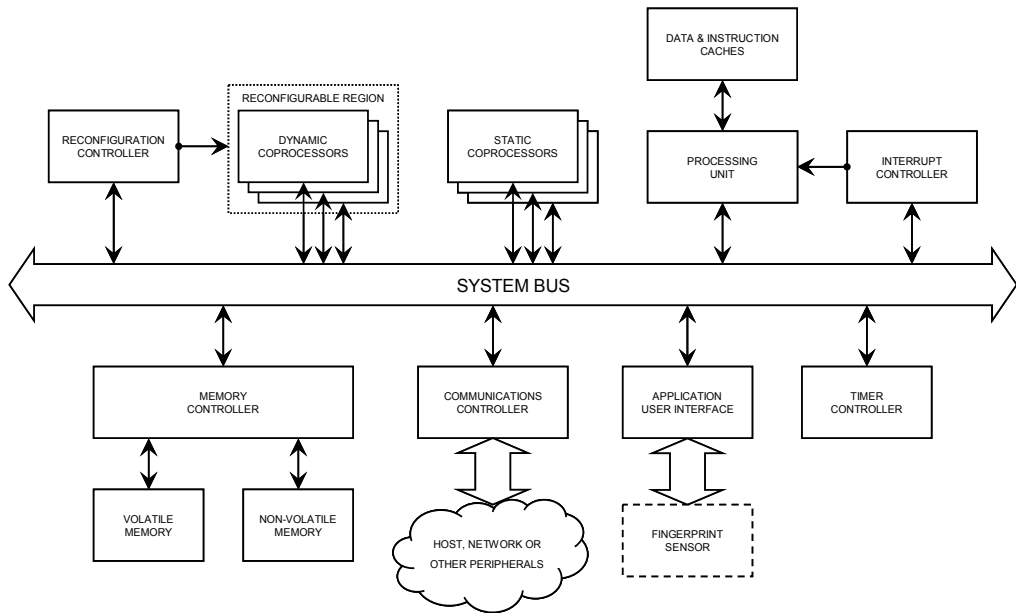
Fig. 11. Run-time reconfigurable embedded system architecture

7. Proof of concept II: hardware-software implementation

A run-time reconfigurable embedded system is presented in this section as general-purpose processing platform where to implement the AFAS application by means of hardware-software co-design techniques. A commercial development board ML401 based on the system-on-programmable-chip device Virtex-4 XC4VLX25 from Xilinx Inc. is used to verify the validity of the proposed system architecture. Additionally to the highly-integrated ML401 development platform, a fingerprint sensor has been connected to the I/O expansion ports of the evaluation board in order to make possible the acquisition of fingerprints in the application, and one RS-232 link has been established between the evaluation board and a personal computer platform in order to simulate the interface between the recognition module and the host or high-level application that makes use of the personal recognition result, as shown in Fig. 12.

The selected SoPC/FPGA device is partitioned in two regions in the biometric application: one static region and one partially reconfigurable region (PRR). In the static region, different components that will be permanently present along the application execution time are instantiated such as one 32-bit MicroBlaze soft-core processor (CPU), data and instruction caches, local memory, one memory management unit (MMU) and other memory controllers to access on-chip and off-chip memory blocks, one dedicated reconfiguration controller in charge of the dynamic reconfiguration of the device, other standard peripherals such as interrupt controller, timer, UART, general-purpose input/output ports, etc. and one specific interface between the static region and the reconfigurable region based on FIFO memories and dedicated 32-bit registers. In the reconfigurable region, application-specific hardware coprocessors will be instantiated under demand along the application execution time in order to perform those image and signal processing tasks required by the AFAS

application. Table 4 shows the amount or resources available in the proposed system-on-programmable-chip and the partitioning of the device into the static and the reconfigurable regions.
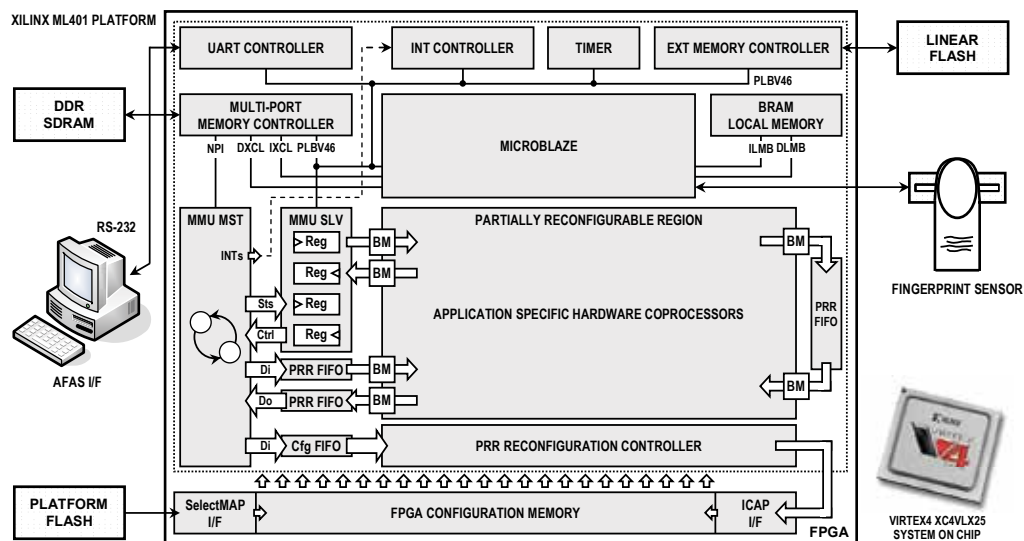

Fig. 12. Run-time reconfigurable embedded system architecture proposed in the physical implementation of an Automatic Fingerprint-based Authentication System

Resources	Xilinx XC4VLX25	Spatial Partitioning	
		Static Region	**Reconfigurable Region**
1-bit Flip Flop	21504	10240	11264
4-input LUT	21504	10240	11264
1-bit RAM	1327104	921600	405504
DSP Block	48	4	44

Table 4. Spatial partitioning of the programmable logic device into one static region and one reconfigurable region

The proposed system-on-programmable-chip is a SRAM-based device. Only volatile memory is embedded on the chip. Additionally to the on-chip volatile memory, the suggested platform is provided with off-chip volatile and non-volatile memory ICs, as it is shown in Fig. 12. Two different types of off-chip non-volatile memories are used:

- The Platform FLASH memory block (4 Mbytes) stores the initial bitstream that defines the configuration of the FPGA upon power up. This initial configuration is composed of the hardware content of the static region (master CPU, memory controllers and other peripherals), and one bootloader application which is executed by the master CPU and is in charge of initializing the system. The initial content of the reconfigurable region of the FPGA remains blank after power up. The transfer of the initial bitstream from the platform FLASH to the internal configuration memory of the

FPGA is automatically done during power up through a dedicated SelectMAP interface present in the FPGA.

- The Linear FLASH memory block (8 Mbytes) contains the definition of those reconfigurable hardware coprocessors to be instantiated in the reconfigurable region of the FPGA along the application execution time, as well as the AFAS program code to be executed by the master CPU. Moreover, the linear FLASH is used in the AFAS application as storage memory where to save the templates of those genuine users registered into the system in the enrolment stage. The reconfiguration process of the PRR is done by means of the dedicated hardware reconfiguration controller instantiated in the static region and the ICAP controller inherent to the device.

Apart from the off-chip FLASH memories, one off-chip SDRAM memory block is also present in the system. During the power up sequence, the bootloader is in charge of initializing the different controllers instantiated in the static region of the FPGA and transferring to the SDRAM memory block (64 Mbytes) the content of the linear FLASH, that is, the AFAS program code and the partial bitstreams that define each of the contexts in which the reconfigurable region is time-multiplexed along the AFAS application. In this way, the off-chip SDRAM memory acts as program and data memory in the application and can be accessed by either the CPU through the PLB bus or the MMU master controller through a dedicated NPI bus. Once all the information is properly transferred to the SDRAM memory, the bootloader gives the control to the AFAS application, and the AFAS application starts.

A multi-bus system architecture permits the interconnection between the different processing blocks. Two specific made-to-measure memory management units –MMU master and slave in Fig. 12– are instantiated in the static region, which aim at interfacing the master CPU and the rest of controllers provided in the static region with those reconfigurable coprocessors instantiated in the reconfigurable region. The interface between the static and reconfigurable regions is built through specific Bus Macros (BM) and some bidirectional FIFO memories intended for a fast exchange of big amounts of data. Moreover, some 32-bit registers are instantiated in the static region in order the master CPU to configure the static and reconfigurable hardware coprocessors, and to control and monitor the application processing flow. The interface between the MMU master and the PRR reconfiguration controller present in the static region is also implemented through a dedicated FIFO memory, as depicted in Fig. 12. The reconfiguration controller is in charge of reading the partial bitstreams previously saved in the SDRAM memory block during power up, and transferring them to the ICAP, which configures the reconfigurable region of the FPGA with the new functional content defined by each bistream. Another FIFO memory block is instantiated in the static region, which acts as a temporary buffer of that information that needs to be shared between different contexts of the PRR region. Before reconfiguring a new context in the PRR region, those parameters that have to be used in the next contexts are saved in that FIFO. After the reconfiguration process, the content of that dedicated FIFO is transferred again to the reconfigurable region in order the new reconfigurable coprocessors instantiated in the PRR to make use of such information.

The interface between the master CPU and those application-specific hardware coprocessors instantiated in the FPGA, either in the static or reconfigurable regions, is provided with some interrupt lines in order any of those hardware coprocessors to be able to notify to the

master CPU about the end of the processing task that is being executed by hardware. Furthermore, in order to reduce the reconfiguration time of the PRR, the size of the reconfigurable region has been minimized as much as possible. A specific reconfiguration controller is instantiated in the static region of the FPGA in order to allow fast reconfiguration without impacting on CPU load. The CPU is only responsible for indicating to the reconfiguration controller the specific partial bitstream that has to be downloaded in the PRR at any time, and once this is defined, the reconfiguration controller is in charge of the reconfiguration process without the need of any further action by the master CPU. Once the reconfiguration is done, the reconfiguration controller notifies the end of the task to the CPU, and the master CPU continues driving the AFAS application program flow. The soft-core processor (master CPU) has been configured to operate at a maximum frequency of 100MHz, and the hardware coprocessors instantiated in the FPGA are designed to operate at either 100MHz or 50MHz depending on the specific task.

The required skills to develop any design based on FPGAs or SoPCs are more demanding than those needed to develop purely software applications. Some background on electronic circuits and programmable logic design, as well as the knowledge of one hardware description language like Verilog or VHDL is required to develop applications based on such kind of architectures. Similarly to what happens with software programming languages and their libraries of functions, some libraries of Intellectual Property descriptions (IPs) of certain functionalities are available to speed up the development of designs based on programmable logic. Moreover, specific EDA tools dependent on the device vendor are normally available to reduce the development cycles when designing with FPGA devices, and the designer needs to get familiar with the processing flow of each automated tool.

Although commercial non-volatile FPGAs have enjoyed great success as development, rapid-prototyping and testing platforms, their use in certain embedded applications has been limited due to their relative high cost in comparison with other solutions. At this level (using the FPGA to implement a static design which keeps invariant during all its execution), the design flow and development tools have been successfully deployed by many vendors (Altera, Actel, Atmel, Lattice, Xilinx, etc.) since decades. However, if the FPGA resources become static after configuration, the device turns into an expensive, power-hungry, low-performance on-field programmable ASIC solution. For FPGAs to become more practical as end-use devices it has been promoted their dynamic reconfiguration capability, i.e., once powered up, the FPGA can be partially reconfigured at run-time, while other part of the FPGA continues operating uninterrupted and automatically maintaining state information between two consecutive reconfigured contexts. In this way, the functions processed in the FPGA can be sequentially swapped in a similar way to the program flow of a CPU-based software application. For this more flexible FPGA conception, however, the designer needs to possess some specific background in those techniques linked to the exploitation of dynamic partial reconfiguration. Moreover, the development tools that automate the new design flow for those applications based on run-time reconfigurable hardware have been an open issue since a long time ago. Recently, however, this landscape experienced a great and definitive change. Xilinx Inc. pushed a definitive impulse to that long-time open issue related to the software tools needed in the PR design flow. Just in 2006, Xilinx presented the new PR design flow fully supported in

Virtex-4 devices. The new top-down design flow eliminated the weakest points highlighted in the previous flows. While still unreleased to the general public, these tools are nowadays presented in the way of an early access version restricted to a limited number of qualified partners who deploy them and contribute feedback to their improvement. This research work is focused on Virtex-4, the first device equipped with a level of PR performance (both technological aspects and supported development tools) acceptable for commercial perspectives. Once finished all the development of our proof-of-concept application, authors think that the current PR flow is today an accepted practice for expert developers with a deep knowledge of the FPGA low-level configuration architecture and, then, it is ready for industrial use. The current Xilinx toolset available in the Xilinx Early Access Partial Reconfiguration (EAPR) lounge and used in this work made possible to automate all the PR design methodology and finish all the phases of the design flow at a reasonable time with no concerns. The toolset used in this work is composed of EDK 9.2.02i to build the PLBv46 bus processor system, PlanAhead 9.2.7 to constrain the floorplan in a friendly graphical way, ISE 9.2.04i_PR12 to generate the bitstreams, as well as ChipScope Pro 9.2i to facilitate the system debugging. In Fig. 13 it is shown all the process to generate both partial and full bitstreams to be downloaded at run-time into the FPGA.

The application is split into a set of sequential stages, and each stage is partitioned into hardware and software tasks. Only those tasks demanding a high computational power or those time-critical tasks that would take too much time if executed by the system CPU are ported to hardware. Specific hardware coprocessors are instantiated in the reconfigurable region of the device to execute such tasks meanwhile the remaining and computationally less expensive tasks are assigned to the system CPU, which furthermore acts as the master processor in charge of driving the application, scheduling the tasks, monitoring the execution flow, and handling the reconfiguration of the PRR when needed along the authentication process. Those partially or fully pipelined hardware coprocessors instantiated in the dynamically reconfigurable region of the FPGA play the role of slave processors in charge of executing those tasks commanded by the master CPU. The dynamic hardware coprocessors are present only when they are needed, thus the same hardware resources available in the reconfigurable region are reused to instantiate different circuits in the application. In Fig. 14 it is shown how the different coprocessors are downloaded into the FPGA to reach a time-multiplexing of the resources placed in the defined PR region of the FPGA. This work results one of the first contributions in the scientific literature that exploits the Xilinx Early Access Partial Reconfiguration electronic design automation tools.

In Section 5 the algorithm has been ported to the presented embedded system (referenced as Embedded System Platform 3 in Tables 1, 2 and 3) and executed purely by software by its MicroBlaze core processor alone. No dedicated hardware was implemented in that scenario, and the application was not able to meet the demanded real-time performance. However, as a result of that implementation under a purely software-based embedded platform, it has been possible to identify those time-expensive computational tasks that constrain the real-time performance of the application in the embedded system. Those time-critical tasks identified in Section 5 are now transferred to hardware to speed up the processing. Owing to the limited resources available in the programmable logic device, up to 9 different reconfigurable contexts have been needed in order to instantiate all the hardware coprocessors along the execution time. Outstanding real-time performances are achieved.

Fig. 13. PR design flow (EDA tools, source code files and resultant bitstreams)

Tables 5 and 6 provide the execution time performance of the application in both enrolment and authentication stages in two different scenarios: (i) when the application is executed purely by software under the system CPU alone, and (ii) when the application is implemented by means of hardware-software co-design techniques making use of the dynamic reconfigurability performance of the suggested FPGA. The final partitioning of the application into hardware and software tasks is also detailed. In the second scenario all tasks are ported to hardware except the fingerprint acquisition process, which is kept as software task under the action of the system CPU. The FPGA resource usage in both the static and reconfigurable regions is shown in Table 7.

Task ID	Processing Stage	Software-only Implementation [1]	Hardware-Software Implementation	
			Sw-only Task	Hw-Sw Task
Task 0	Fingerprint acquisition	500.000 ms	500.000 ms	
Task 1	Image segmentation	232.046 ms		0.672 ms
Task 2	Reconfiguration 1 → 2			0.841 ms
	Image normalization	33.087 ms		0.850 ms
Task 3	Reconfiguration 2 → 3			1.045 ms
	Image isotropic filtering	512.171 ms		2.563 ms
Task 4	Reconfiguration 3 → 4			1.025 ms
	Field orientation	285.485 ms		0.669 ms
Task 5	Reconfiguration 4 → 5			1.046 ms
	Filtered field orientation	19.143 ms		0.419 ms
Task 6	Reconfiguration 5 → 6			1.107 ms
	Image directional filtering and binarization	656.043 ms		2.465 ms
Task 7	Reconfiguration 6 → 7			1.045 ms
	Image smoothing	253.553 ms		0.447 ms
Task 8	Reconfiguration 7 → 8			0.974 ms
	Image thinning	416.316 ms		0.902 ms
Task 9	Reconfiguration 8 → 9			0.943 ms
	Minutiae extraction and minutiae filtering	25.699 ms		4.919 ms
Total Execution Time [2]:		**2433.543 ms**	**21.932 ms**	

[1] : The software-only execution times are slightly higher than in the Embedded System 3 scenario of Table 2 because of the reduction of the cache memory size in this new scenario in order to allocate additional memories in the hardware coprocessors (only 8KB of Instruction and Data caches are instantiated in MicroBlaze interface instead of the initial 32KB Instruction cache and 64KB Data cache).
[2] : Task 0 is not included in the computation of the total execution time.

Table 5. Execution time performance reached in the enrolment stage: SW-only versus HW-SW implementations

Task ID	Processing Stage	Software-only Implementation [1]	Hardware-Software Implementation	
			Sw-only Task	Hw-Sw Task
Task 0	Fingerprint acquisition	500.000 ms	500.000 ms	
Task 1	Image segmentation	232.046 ms		0.672 ms
Task 2	Reconfiguration 1 → 2			0.841 ms
	Image normalization	33.087 ms		0.850 ms
Task 3	Reconfiguration 2 → 3			1.045 ms
	Image isotropic filtering	512.171 ms		2.563 ms
Task 4	Reconfiguration 3 → 4			1.025 ms
	Field orientation	337.419 ms		0.669 ms
Task 5	Reconfiguration 4 → 5			1.046 ms
	Filtered field orientation	22.178 ms		0.419 ms
Task 6	Reconfiguration 5 → 6			1.107 ms
	Image directional filtering and binarization	774.750 ms		2.465 ms
Task 7	Reconfiguration 6 → 7			1.045 ms
	Image smoothing	287.507 ms		0.447 ms
Task 8	Reconfiguration 7 → 8			0.974 ms
	Image thinning	417.350 ms		0.820 ms
Task 9	Reconfiguration 8 → 9			0.943 ms
	Minutiae extraction and minutiae filtering	32.497 ms		7.606 ms
Task A	Reconfiguration 9 → A			1.045 ms
	Field orientation maps alignment	139935.838 ms		157.671 ms
Task B	Reconfiguration A → B			1.035 ms
	Minutiae alignment, feature sets matching and authentication decision	108.608 ms		20.737 ms
Total Execution Time [2]:		**142693.451 ms**	**205.025 ms**	

[1] : The software-only execution times are slightly higher than in the Embedded System 3 scenario of Table 3 because of the reduction of the cache memory size in this new scenario in order to allocate additional memories in the hardware coprocessors (only 8KB of Instruction and Data caches are instantiated in MicroBlaze interface instead of the initial 32KB Instruction cache and 64KB Data cache).
[2] : Task 0 is not included in the computation of the total execution time.

Table 6. Execution time performance reached in the authentication stage: SW-only versus HW-SW implementations

Task ID	Processing Stage	Hardware Resources			
		1-bit Flip Flop	4-input LUT	1-bit RAM	DSP Block
–	Application flow (static design)	7005	8888	755712	4
Task 0	Fingerprint acquisition	–	–	–	–
Task 1	Image segmentation	4978	4612	147456	20
Task 2	Image normalization	371	334	0	8
Task 3	Image isotropic filtering	5275	5831	92160	28
Task 4	Field orientation	3339	3166	92160	8
Task 5	Filtered field orientation	2857	2983	129024	0
Task 6	Image directional filtering and binarization	5462	4166	313344	29
Task 7	Image smoothing	4892	3265	147456	0
Task 8	Image thinning	1013	2821	239616	0
Task 9	Minutiae extraction and minutiae filtering	487	3379	55296	0
Task A	Field orientation maps alignment	2632	8943	387072	0
Task B	Minutiae alignment, feature sets matching and authentication decision	642	4379	258048	5
Total Design Resources:		**38953**	**52767**	**2617344**	**102**
Total Device Resources:		**21504**	**21504**	**1327104**	**48**

Table 7. FPGA resources usage in each of the application contexts

HARDWARE COPROCESSORS
PARTIAL BITSTREAMS   SYSTEM CPU
STATIC BITSTREAM   HW/SW PROCESSING STAGES
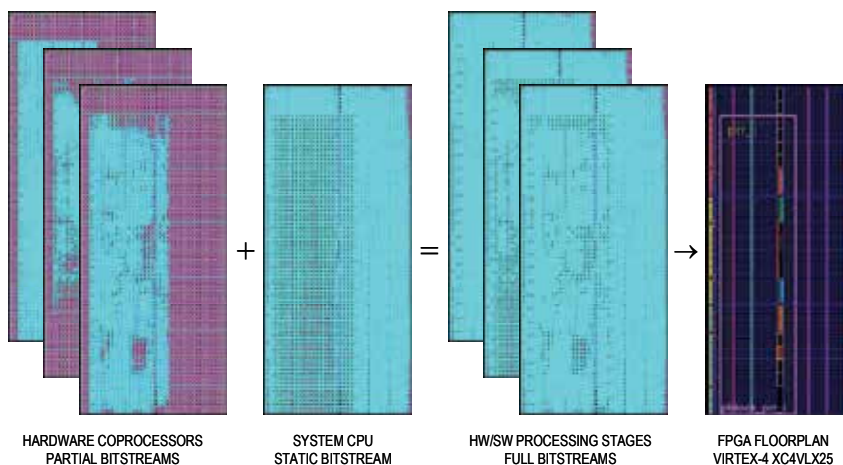FULL BITSTREAMS   FPGA FLOORPLAN
VIRTEX-4 XC4VLX25

Fig. 14. Temporal partitioning of the application in sequential tasks running in the PRR of a FPGA. The bitstream gets composed of a static region and a reconfigurable region (left). Spatial partitioning of the application floorplanned in both static and reconfigurable regions on the Xilinx Virtex-4 XC4VLX25 device (right)

The physical resources needed to implement each of the hardware coprocessors are detailed. From the resources usage shown in Table 7 it can be deduced that the reconfigurability performance of the FPGA permits a notorious reduction of the amount of resources needed in the programmable logic device in comparison with the amount of resources that would be needed in case of using a non-reconfigurable FPGA, where all coprocessors would be instantiated permanently in a static way. Thanks to the reconfigurability performance exhibited by the suggested device and the hardware-software partitioning of the application it has been possible to develop one application that demands 38953 1-bit flip-flops, 52767 4-bit LUTs, 2617344 1-bit RAM cells and 102 DSP blocks with one device that features 21504 1-bit flip-flops, 21504 4-bit LUTs, 1327104 1-bit RAM cells and 48 DSP blocks. The reuse of the hardware resources allows reducing the amount of resources at the expense of the reconfiguration overhead, which is also minimized by the design of an efficient reconfiguration controller. The amount of needed resources and the reached performances exhibited by the suggested run-time reconfigurable embedded system clearly outperform those featured by one PC platform. The total authentication execution time results in 205.025 ms, which leads to a speed up of ×686.58 (or ×695.980 depending on the used cache) when compared against the purely software implementation of the recognition algorithm under the same embedded system platform, and a speed up of ×15.97 with regard to the application execution time featured by the PC platform presented in Section 5.

8. Conclusion

The successful spread of products and services that exploit the advantages provided by fingerprint biometrics in both public and private sectors depend on several factors today. Although the universality, distinctiveness and permanence characteristics of human fingerprints are proven facts that make them reliable signs of identity, the acceptance of automated fingerprint-based personal recognition systems, focused on either identification or authentication purposes, is constrained by social and technical factors. Among the social factors, the most important ones refer to the security and privacy concerns related to the protection of the user's information integrity; and among the technical factors, the most limiting ones refer to the accuracy of the recognition system, the authentication response time and the cost of the whole application. All they are barriers to the broad adoption of that kind of systems worldwide. If fingerprint recognition technology continues to mature and efficient and reliable systems able to overcome all those barriers are designed, automated fingerprint-based recognition can have a profound influence on the way we conduct our daily business in the near future.

As far as authors know at the moment of publication of the present work, there does not exist in the market any AFAS application based on dynamically reconfigurable hardware. Flexible and dynamically reconfigurable hardware allows a more efficient usage of the system resources by having hardware present in the FPGA device only when it is in use. Thus given a fixed size for the FPGA, it is possible to instantiate specific coprocessors at a given time, and to eliminate them after they have been used in order to allow further coprocessors to be instantiated making use of the same FPGA resources in the following stages of the application. This technique allows reducing the overall hardware system size at nearly null cost –FPGA reconfiguration overhead–.

The results presented in this work prove that the suggested system architecture can be an efficient alternative to those existing AFAS based on either expensive personal computer

platforms or embedded systems that make use of MPUs, GPUs, DSPs, ASSPs or ASICs. This novel approach, focused on the exploitation of run-time reconfigurable FPGA devices and hardware-software co-design techniques, pursues two main objectives: (i) to meet the required expectations for the application, which means to fulfil the functionality demands (accurate FAR/FRR personal recognition rates) with the proper response time (real-time) and reliability levels (protection against fraudulent attacks); and (ii) to meet those requirements with the minimum possible cost for the system, and with the proper flexibility to allow future changes/improvements in the personal recognition algorithm (added-value). There are endless uses for embedded systems based on SoPC or FPGA devices in the consumer, military, aerospace, automotive, communications, and industrial markets worldwide. In this direction, the proposed embedded system architecture, based on run-time reconfigurable hardware, is proven to be a valid and cost-effective solution that encourages the reduction of system resources in the physical implementation of those complex computational applications demanding high processing power and real-time performances such as the ones resulting from the biometrics field. As computer technology continues to advance and economies of scale reduce costs, fingerprint biometric systems based on the suggested topology can become a more efficient and cost-effective means for personal verification in both public and private sectors. The proposed system architecture can thus help in paving the way for the exploitation of biometric systems all over the world.

9. References

Becker, J.; Hübner, M.; Hettich, G.; Constapel, R.; Eisenmann, J. & Luka, J. (2007). Dynamic and Partial FPGA Exploitation. *Proceedings of the IEEE*, Vol. 95, No. 2, (February 2007), pp. 438-452, ISSN 0018-9219

Fons, M.; Fons, F. & Cantó, E. (2010). Fingerprint Image Processing Acceleration through Run-Time Reconfigurable Hardware. *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 57, No. 12, (December 2010), pp. 991-995, ISSN 1549-7747

FPC2020 ASIC Fingerprint Processor and FPC-AM3 Biometric Module, (n.d.), 2011, Available from http://www.fingerprints.com

FxIntegrator Fingerprint Recognition Module, (n.d.), 2011, Available from http://www.biometrika.it

Liu-Jimenez, J.; Sanchez-Reillo, R.; Lindoso, A. & Miguel-Hurtado, O. (2006). FPGA Implementation for an Iris Biometric Processor, *Proceedings of IEEE International Conference on Field Programmable Technology*, pp. 265-268, ISBN 0-7803-9729-0, Bangkok, Thailand, December 13-15, 2006

Lopez-Ongil, C.; Sanchez-Reillo, R.; Liu-Jimenez, J.; Casado, F.; Sánchez, L. & Entrena, L. (2004). FPGA Implementation of Biometric Authentication System Based on Hand Geometry, *Proceedings of International Conference on Field Programmable Logic and Appplication*, pp. 43-53, LNCS 3203, Antwerp, Belgium, August 30 – September 1, 2004

Maio D.; Maltoni, D.; Cappelli, R.; Wayman, J.L. & Jain, A.K. (2004). FVC2004: Third Fingerprint Verification Competition, *Proceedings of International Conference on Biometric Authentication*, pp. 1-7, LNCS 3072, Hong Kong, China, July 15-17, 2004

Maltoni, D.; Maio, D.; Jain, A.K. & Prabhakar, S. (2009). *Handbook of Fingerprint Recognition, Second Edition,* Springer-Verlag, ISBN 978-1-84882-253-5, London, England

ML67Q5250 Fingerprint Authentication MCU, (n.d.), 2011, Available from http://www.okisemi.com

MV1210 and MV1250 Bioscrypt Fingerprint Modules, (n.d.), 2011, Available from http://www.l1id.com

Nanni, L. & Lumini, A. (2009). Descriptors for Image-based Fingerprint Matchers. *Experts Systems with Applications*, Vol. 36, No. 10, (December 2009), pp. 12414-12422, ISSN 0957-4174

One Touch SDK. (n.d.), 2011, Available from http://www.digitalpersona.com

Pavan Kumar, A.; Kamakoti, V. & Das, S. (2007). System-on-Programmable-Chip Implementation for On-Line Face Recognition. *Pattern Recognition Letters*, Vol. 28, No. 3, (February 2007), pp. 342-349, ISSN 0167-8655

plusID Universal Biometric Devices (n.d.), 2011, Available from http://www.privaris.com

SDA Stand-Alone Fingerprint Recognition Modules (n.d.), 2011, Available from http://www.secugen.com

SecurASIC Chip, (n.d.), 2011, Available from http://www.cogentsystems.com

SFM Series Fingerprint Modules, (n.d.), 2011, Available from http://www.supremainc.com

TCD50D Digital ID Hardware Engine, (n.d.), 2011, Available from http://www.upek.com

TMS320 Texas Instruments DSP Platforms, (n.d.), 2011, Available from http://www.ti.com

Verifinger SDK. (n.d.), 2011, Available from http://www.neurotechnology.com

Yang, S.; Sakiyama, K. & Verbauwhede, I. (2006). Efficient and Secure Fingerprint Verification for Embedded Devices. *EURASIP Journal on Applied Signal Processing*, Vol. 2006, No. 3, (January 2006), pp. 1-11

Yang, J.C. & Park, D.S. (2008). A Fingerprint Verification Algorithm Using Tessellated Invariant Moment Features. *Neurocomputing*, Vol. 71, No. 10-12, (June 2008), pp. 1939-1946, ISSN 0925-2312

BiSpectral Contactless Hand Based Biometric Identification Device

Aythami Morales and Miguel A. Ferrer
Instituto para el Desarrollo Tecnológico y la Innovación en Comunicaciones (IDeTIC),
Universidad de Las Palmas de Gran Canaria,
Spain

1. Introduction

Biometrics plays an increasingly important role in authentication and identification systems. The process of biometric recognition allows the identification of individuals based on the physical or behavioral characteristics. Among the most common biometric features used are fingerprint, iris, face, voice, signature and hand. Hand based biometric systems exhibit many desirable characteristics when working with low resolution sensors (which are most appropriate for civil and commercial applications), including low cost sensors, acceptable indentification performance, robustness to environmental conditions and individual anomalies, and high speed identification algorithms. For higher security applications such as forensics, high resolution images are more suitable (Jain et al, 2001) (Konga et al, 2009).

Most hand-based biometric schemes in the literature are based on measuring the hand silhoutte as a distinctive personal attribute for an authentication task. First it was accomplished using guiding pegs mounted on a flat surface of the imaging device (Sanchez-Reillo et al, 2000) (Jain et al, 1999). Although the guiding pegs provide consistent measuring positions, they cause some problems as well: 1) The pegs can deform the shape of a hand (Wong & Shi, 2002) and 2) The users must be well trained to cooperate with the system. Thus, peg-free hand geometry techniques were considered giving the hand some motion freedom (Bulatov et al, 2004).

There are two main approaches for geometrical features extraction; those based on measure the finger lengths and widths at various positions, palm size, etc. and another based on represent the global hand shape (Öden et al, 2003) (Yörük et al, 2006). Both approaches use the finger tip points and the finger valley points as the landmarksfor image alignment.

The palm texture can be also used as biometric trait for personal identification. It can be used both by itself (Han et al, 2003) (Ribarić & Fratric, 2005) (Sun et al 2005) (Kong & Zhang, 2004) (Badrinath & Gupta, 2009) or combined with hand shape (Ribaric et al, 2003) (Li et al, 2006) (Kumar et al, 2006) (Kumar & Zhang, 2004) at score level or at representation level. Although fusion increases accuracy, it generally increases computation costs and template sizes and reduces user acceptance.

Recently, perhaps due to hygiene consideration, contact-free hand biometric systems have been proposed. The two main issues to be dealt with in a contact-free system are hand segmentation and the projective distortions associated with the absence of the contact plane.

Previous research on contactless systems includes (Haeger, 2003), where once the centroid of a segmented hand was detected a number of concentric circles were drawn around the centroid passing through the fingers. Using these circles 124 different finger sizes were measured and used for biometric identification with limited results. (Hao et al, 2008) proposes a contactless biometric system based on a fusion of palm texture and palm vein pattern based on feature level and image level fusion. To realize the acquisition the user introduces the hand in a black box. Therefore illumination and background were controlled. The use of such black box can raise concerns or unwillingly scare the users and lower the user acceptance. Doi and Yamanaka (Doi et al, 2003) created a mesh of a hand image captured by an infrared CCD camera. The mesh was created using 20 to 30 feature points extracted from the principal creases of the fingers and palm. Root-mean-square (rms) deviation was used to measure the alignment distance between the meshes, which was was also sensitive to perspective distortion. In reference (Morales et al, 2008), the contactless hand geometry system able to obtain images in non controlled environments is investigated. The hand geometry based feature extraction methods show poor results due to projective distortion problems. Palmprint authentication based on contactless imaging was proposed in (Morales et al 2010). In this chapter was proposed a combined method based on two palm features approaches. The combination with an uncorrelated biometric as hand geometry was mentioned as future work. As result of the above experience, the aim of this chaper is get together all the previous experience and propose a contact-free biometric system based on the combination of hand geometry and palmprint using only low cost devices for medium security environments. The device uses infrared illumination and infrared camera to reduce some problems as changing lighting conditions or complex background containing surfaces and objects with skin-like colors. To acquire the the palm texture information a second camera in the visible band is added. The visible image can then be segmented using the information from the infrared camera. We propose the use of Active Shape Models to correlate the hand contourns from the infrared and visible images. The projective distortion problem is alleviated using a template guide on the video screen. The verification methodology includes 1000 hand images from a database acquired with the proposed device.

The outline of the chapter is as follows. In the next section we will introduce the proposed bispectral contactless hand-based biometric device. Section III describes the geometry parameters and Section IV presents the palmprint approaches based on OLOF and MSIFT. Section V presents our experimental results we assess the proposed device. The chapter is closed with conclusions, acknowledgements and references.

2. Acquisition device

The acquisition device used consists of two inexpensive, standard web cams that obtain images of the hand at the same time. The so called infrared (IR) webcam acquires images in the infrared band (750 to 1000nm) and the so called visible (V) camera acquires images in the visible range (400 to 700nm).

The IR webcam was created by simply taking out the webcam lens that eliminates the infrared radiation and adding a filter that eliminates the visible band. We used Kodak filter No 87 FS4-518 and No 87c FS4-519 with no transmittance below 750 nm.

The images of the IR webcam were used for hand geometry. So, we increase the image contrat by setting the IR webcam specification as follows: maximum value of contrast and low values of brightness, gain and exposure time. An example of the image acquired can be seen in Figure 1.

Fig. 1. Left, hand acquired with a standard webcam; right, hand acquired with the IR webcam.

The infrared illumination is composed of a set of 24 GaAs infrared emitting diode (CQY 99) with a peak wavelength emission of 925 nm and a spectral bandwidth of 40 nm. The diodes were placed in an inverted U shape with the IR and V webcams in the middle (see Figure 2). The open part of the U shape will coincide with the wrists of the hand image. The focus of the IR webcam lens is adjusted manually the first time the webcam is used.
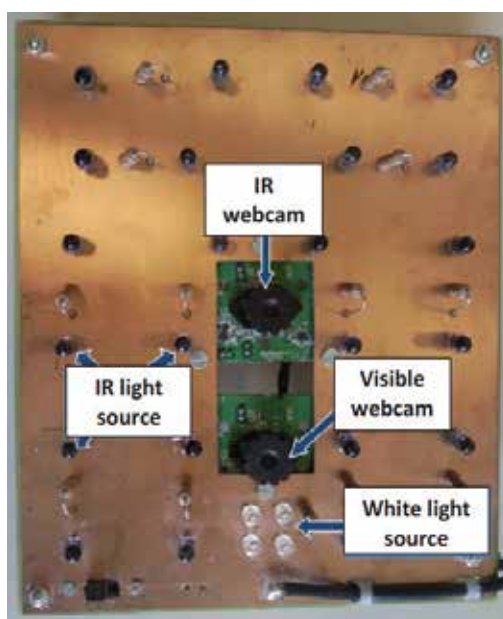
Fig. 2. Bispectral hand based biometric system

To alleviate the projective distortion of the hand image acquired we used a hand mask in the video screen of the computer: the user places his or her hand over the camera and adjusts the position and pose of the hand in order to overlap with the hand mask drawn on the device screen. When the hand and mask overlap more than 70%, the device automatically acquires both the IR and visible image. An example of this process can be seen in Figure 3. The mask used was the averaged hand silhouette from the GPDS hand database scaled to the webcam resolution and dilated with a 30 by 30 structuring element.
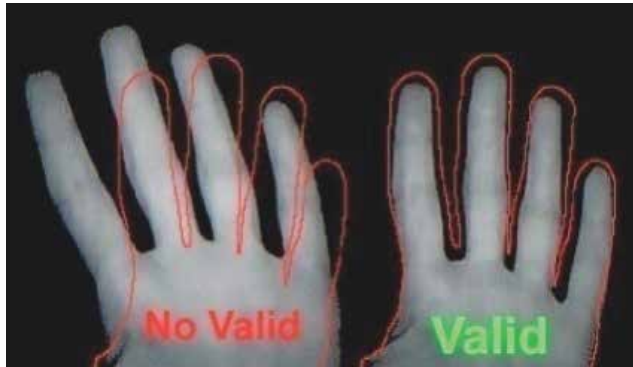
Fig. 3. Hand mask and hand overlapping. Valid stands for overlapping greater than 70%.

The V webcam used for palm print biometries is located just 2 centimeters below the IR webcam. The settings of the V webcam are configurated by default. The lens focus is adjusted manually the first time it is used. The illumination consists of a set of 4 white LEDs emitting in the 400nm to 700nm band. An example of images acquired can be seen in Figure 4.
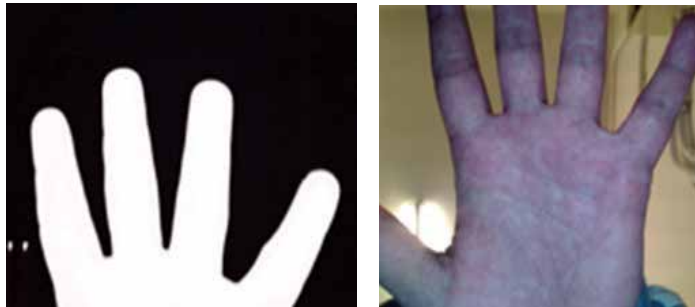

Fig. 4. Left: image acquired by the IR webcam, right: visible image of the hand palm.

3. Geometrical hand biometries

Due to the webcam setup and IR illumination, a reliable hand contour can be obtained binarizing the IR image with its Otsu's threshold.

To work out the tips and valleys between the fingers we convert the Cartesian coordinates of the contour to polar coordinates (radius and angle) considering the center of the image base as the coordinates origin. The peaks in the radius coordinate locate the provisional position of the finger tips and the minima of the radius indicate the valleys between fingers. Let and $\varphi_c(i), 1 \leq i \leq L$ the radius and angle of the ith hand contour pixel. The index i_p^j of the ith radius peaks are obtained as:

$$i_p^j \in peak \quad if \; r_c\left(i_p^j\right) = \max\left\{r_c(i)\middle| i_p^j - 100 \leq i \leq i_p^j + 100\right\} \tag{1}$$

With $101 < i_p^1 < i_p^2 < \ldots < i_p^5 < L - 100$. If the number of radius peaks obtained is greater than 5, we suppose than the hand detector has been fault and go to hand detection module waiting for a hand. As the hand is expected, i_p^1 corresponds to the little finger tip.

The index i_v^j of the jth valley is worked out as:

$$i_v^j \in valley \quad if \; r_c\left(i_v^j\right) = \min\left\{r_c(i) \middle| i_p^j \le i \le i_p^{j+1}\right\} \tag{2}$$

The exterior base of the index and little fingers are obtained as the nearest pixel of the exterior contour to the valley between the index and middle fingers and the valley between the index and little fingers, respectively, i.e.:

$$i_{index} \in valley \quad if \; i_{index} = \arg\min_i\left\{d(\langle x_c(i), y_c(i)\rangle, \langle x_c(i_v^3), y_c(i_v^3)\rangle) \middle| i_p^5 \le i \le i_v^4\right\} \tag{3}$$

$$i_{little} \in valley \quad if \; i_{little} = \arg\min_i\left\{d(\langle x_c(i), y_c(i)\rangle, \langle x_c(i_v^1), y_c(i_v^1)\rangle) \middle| 1 \le i \le i_p^1\right\} \tag{4}$$

Being $d(\cdot,\cdot)$ the Euclidean distance. We will call $i_v^1 = i_{little}$, and $i_v^5 = i_{index}$.
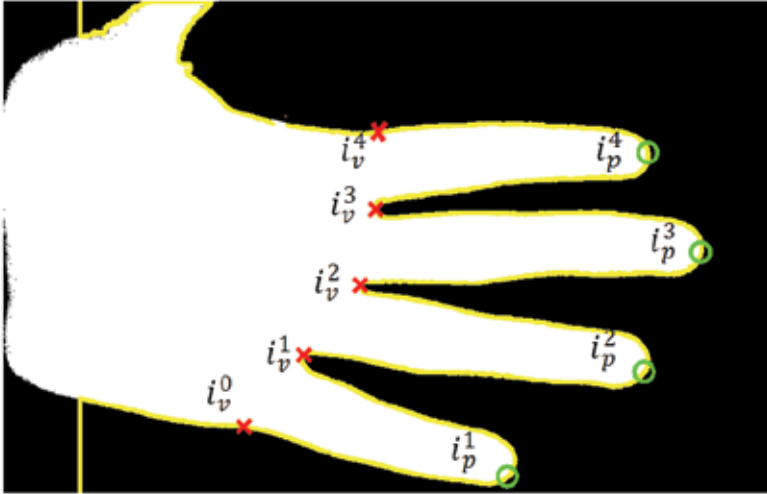
Fig. 5. Tips, valleys and exteriors of fingers localization

The position of the tip of the finger is finely adjusted as follow. The lines that minimize the square error with each side finger contour are obtained as follows:

1. Four equal spaced points are selected from the 35% to the 80% of each finger side. The 35% is selected to avoid the presence of rings, and the 80% is selected to avoid the tip curvature of the finger tip. For the right side of the finger, the four points are calculated as $i_{rfs}^j = (i_p^j - i_p^{j-1}) * C(k) + i_v^{j-1}$, being $C(k) = \{0.35, 0.50, 0.65, 0.80\}$, and for the left finger side are calculated as $i_{lfs}^j(k) = (i_v^{j+1} - i_p^j) * C(k) + i_p^j$.

2. The lines that minimize the square error with the selected point of each finger side are calculated. For the right side, the line is defined as $y = m_r^j \cdot x + b_r^j$, being b_r^j and m_r^j calculated as:

$$\begin{pmatrix} b_r^j \\ m_r^j \end{pmatrix} = pinv \begin{pmatrix} 1 & x(i_{rfs}^j(1)) \\ 1 & x(i_{rfs}^j(2)) \\ 1 & x(i_{rfs}^j(3)) \\ 1 & x(i_{rfs}^j(4)) \end{pmatrix} \cdot \begin{pmatrix} y(i_{rfs}^j(1)) \\ y(i_{rfs}^j(2)) \\ y(i_{rfs}^j(3)) \\ y(i_{rfs}^j(4)) \end{pmatrix} \tag{5}$$

being pinv the pseudoinverse. For the left side, the line $y = m_l^j \cdot x + b_l^j$ is obtained as above using $i_{rfs}^j(k)$, Fig 6.
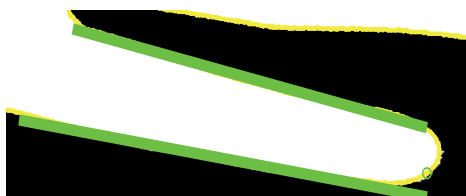

Fig. 6. Finger contour line approximation $y = m_r^j \cdot x + b_r^j$ and $y = m_l^j \cdot x + b_l^j$

3. The average of the two lines is considered the finger axis and calculated as $y = m_a^j \cdot x + b_a^j$ being $m_a^j = (m_r^j + m_l^j)/2$ and $b_a^j = (b_r^j + b_l^j)/2$, see Fig 7.
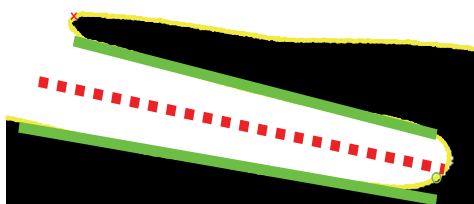

Fig. 7. Finger axis caluculation.

4. The tip of the finger is the point where the finger axis and the finger contour intersect, Fig 8.

$$i_p^j = \arg\min_i \left\{ d(\langle x_c(i), y_c(i) \rangle, y = m_a^j \cdot x + b_a^j) \Big| i_v^{j-1} \leq i \leq i_v^{j+1} \right\} \tag{6}$$
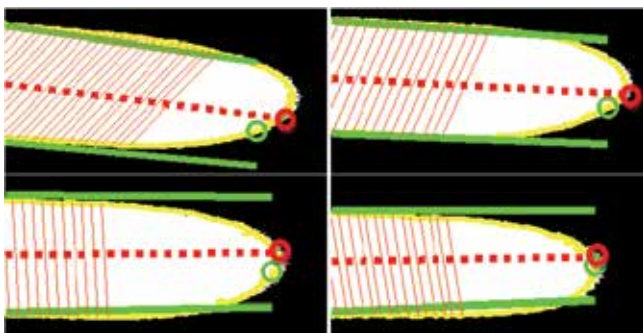

Fig. 8. In green initial tip localization; in red accurate tip localization.

The geometric features are obtained by measuring the widths of each finger. It is done as follows: The center base of each the finger $\left\langle x_{fb}^j, y_{fb}^j \right\rangle$ is defined as the point where the finger axis $y = m_a^j \cdot x + b_a^j$ intersects the finger base line

$$y = \frac{y(i_v^j) - y(i_v^{j-1})}{x(i_v^j) - x(i_v^{j-1})}(x - x(i_v^{j-1})) + y(i_v^{j-1}) \tag{7}$$

We select 12 equal space points between $\left\langle x_{fb}^j, y_{fb}^j \right\rangle$ and $\left\langle x_c(i_p^j), y_c(i_p^j) \right\rangle$ as follows:

$$x_s^j = (x_c(i_p^j) - x_{fb}^j) * C(k) + x_{fb}^j \tag{8}$$

$$y_s^j(k) = m_a^j \cdot x_s^j(k) + b_a^j \tag{9}$$

with $C(k) = \{0.20, 0.26, 0.32, \ldots, 0.80, 0.86\}$.

The perpendicular line to the finger axe is obtained in this point as

$$y = \frac{-1}{m_a^j}(x - x_s^j(k)) + y_s^j(k) = m_{pa}^j \cdot x + b_{pa}^j \tag{10}$$

The nearest contour points this line

$$i_{cr}^j(k) = \arg\min_i \left\{ d(\langle x_c(i), y_c(i) \rangle, y = m_{pa}^j \cdot x + b_{pa}^j) \Big| i_v^{j-1} \le i \le i_p^j \right\} \tag{11}$$

and

$$i_{cl}^j(k) = \arg\min_i \left\{ d(\langle x_c(i), y_c(i) \rangle, y = m_{pa}^j \cdot x + b_{pa}^j) \Big| i_v^j \le i \le i_v^j \right\} \tag{12}$$

Being the width at this point $d_w^j(k) = d(\langle x_c(i_{cr}), y_c(i_{cr}) \rangle, \langle x_c(i_{cl}), y_c(i_{cl}) \rangle)$. The geometric features are obtained by measuring 100 widths of each finger from the 15% to 85% of the finger length. An example can be seen in figure 9.

Fig. 9. Finger widths measured for the geometrical template

The width measures of the four fingers are concatenated resulting in a vector of 400 components $d_w^j(k), 1 \leq j \leq 4, 1 \leq k \leq 100$. The maximum of the vector is normalized to 1 to reduce the projection distortion and its average substracted. In order to reduce the dimensionality of the vector, the DCT transform is applied and the geometrical hand template is obtained by selecting from the 2nd to the 50th coefficients of the DCT transform.

As verifier we have used a Least Squares Support Vector Machine (LS-SVM). SVMs have been introduced within the context of statistical learning theory and structural risk minimization. Least Squares Support Vector Machines (LS-SVM) are reformulations to standard SVMs which lead to solving linear KKT systems. Robustness, sparseness, and weightings can be imposed to LS-SVMs where needed and a Bayesian framework with three levels of inference is then applied (Suykens et al, 2002).

The meta-parameters of the LS-SVM model are the width of the Gaussian kernels σ and the regularization factor γ. The regularization factor is taken as $\gamma = 20$ and is identical for all the LS-SVM models used here. The Gaussian width σ parameter is optimized as follows: the training sequence is randomly partitioned into two equal subsets $P_i, 1 \leq i \leq 2$. The LS-SVM is trained $L = 30$ times with the first subset P_1, $\gamma = 20$ and Gaussian width equal to L logarithmically equally spaced values between 10^1 and 10^4 $\sigma_l, 1 \leq l \leq L$. Each one of the L LS-SVM models is tested with the second subset P_2 obtaining L Equal Error Rate $EER_l, 1 \leq l \leq L$ measures and their associated thresholds $TEER_l, 1 \leq l \leq L$. As the positive samples are trained with target output +1 and the negative samples with target value -1, the threshold is limited to values between $-1 \leq TEER_l \leq 1$. The Gaussian width σ of the signature model and its decision threshold $TEER$ are obtained as $\sigma = \sigma_j$ and $TEER = (TEER_j + 1) \cdot 0.8 - 1$, where $j = \arg\min_{1 \leq l \leq L} \{EER_l\}$. Finally, the user hand model is obtained training the LS-SVM with all the training sequence.

To verify that an input image belongs to the claimed used, we calculated the score of the LS-SVM that models the claimed user. If the score is greater than the claimed user $TEER$, it is accepted as genuine.

4. Palm print subsystem

4.1 Hand segmentation

To extract the palm texture we use the visible image of the hand. The major problem in the visible image is the hand segmentation to obtain an invariable area of the hand palm. As the relation between the pixels of both images is variable depending of the distance from the camera to the hand, the contour obtained by the IR image is taken as initial guess of the hand contour in the visible image and the orientation, scale, position, and shape of the IR contour is adjusted to the visible image using an Active Shape Model (ASM) (Cootes et al, 1995).

ASMs are flexible models of image structures whose shape can vary. The models are able to capture the natural variability within a class of shapes, in this case hands, and can then be used for image segmentation (in addition to other applications). The ASM model was constructed from a dataset of 500 hand contours from the first 50 users of the GPDShand database (Ferrer et al, 2007).

For the point distribution models of the contours, we selected as landmark points the valley of the fingers. Between each pair of consecutive landmark points we selected 70 additional

points. In the trained model 96% of the variance could be explained by the first 9 eigenvectors or modes of variation.

Trained the ASM model, to segment the hand in the visible image, the landmarks and point between landmarks are obtained over the contour of the acquired hand in the IR image and they are displaced, rotated and distorted inside the ASM limits looking for edges in the visible image.The edges of the visible images were obtained summing the morphological gradient of the red, green and blue images. Figure 10 shows the results at the initial and final stages of the algorithm.
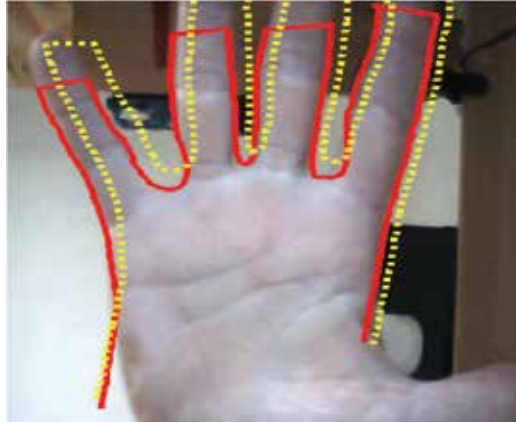

Fig. 10. Visible image, dotted line: initial contour, solid line: final contour..

4.2 Palmprint texture parameterization

The Orthogonal Line Ordinal Features (OLOF) method was originally introduced in (Sun et al, 2005) and was investigated for the palmprint feature extraction. The comparison of OLOF method with several other competing methods (Kong & Zhang, 2004) in this reference suggests the superiority of OLOF with such competitive feature extraction methods. The OLOF presented significantly improvement results but on conventional databases that are acquired from constrained imaging.

This method is based on 2D Gaussian filter to obtain the weighted average intensity of a line-like region. Its expression is as follows:

$$f(x,y,\theta) = \exp\left[-\left(\frac{x\cos\theta + y\sin\theta}{\delta_x} \right)^2 - \left(\frac{-x\sin\theta + y\cos\theta}{\delta_y} \right)^2 \right] \tag{13}$$

where θ denotes the orientation of 2D Gaussian filter, δ_x denotes the filter's horizontal scale and δy denotes the filter's vertical scale parameter. There no signicicant differences on results in the range $\delta_x, \delta_y \in [0.5-10]$. We empirically selected the parameters as δ_x = 5 and δ_y = 1.

To obtain the orthogonal filter, two Gaussian filters are used as follows:

$$OF(\theta) = f(x,y,\theta) - f(x,y,\theta + \frac{\pi}{2}) \tag{14}$$

Each palm image is filtered using three ordinal filters, $OF(\theta)$, $OF(\pi/6)$ and $OF(\pi/3)$ to obtain three binary masks based on a zero binarization threshold. In order to ensure the robustness against brightness, the discrete filters $OF(\theta)$, are turned to have zero average. Once filtered the central portion for images is cropped and binarized giving a value of 1 to 25% of the highest gray level pixels and the rest reset to 0 values. Finally, the three images are reduced to 50x50pixels. An example of this image can be seen in Figure 11.
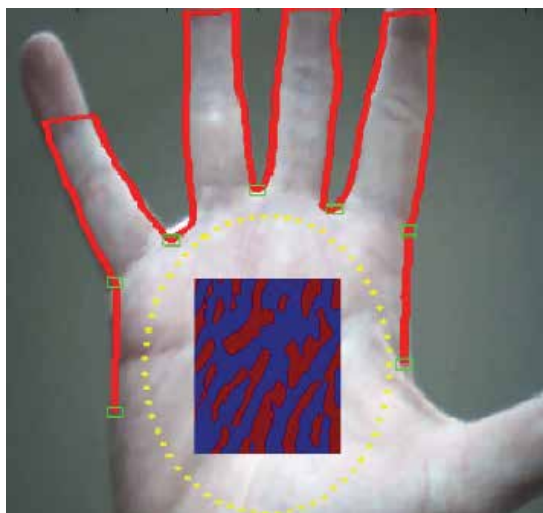

Fig. 11. Invariant area of the hand palm and the OLOF features overlapped to the palm

To verify that an input texture Q belongs to the identity with image texture (template) P we have used a normalized Hamming measure which can be described as:

$$D = 1 - \frac{\sum_{i=1}^{2n+1}\sum_{j=1}^{2n+1} P(i,j) \otimes Q(i,j)}{(2n+1)^2} \qquad (15)$$

where the boolean operator $\otimes$ is equal to zero if and only if the two bits $P(i,j)$ and $Q(i,j)$ are equals. It is noted that D is between 0 and 1 (best matching). Because of the imperfect preprocessing, we need to vertically and horizontally translate one of the features a range of 4 to 4 and match again. The maximum D value obtained is considered to be the final matching score. If the matching score is greater than a threshold, the hand is accepted.

4.3 Palmprint MSIFT parameterization

The Scale Invariant Feature Transform was originally proposed in (Lowe, 2004). The features extracted are invariant to image scaling, rotation, and partially invariant to change in illumination and projective distortion. The SIFT is a feature extraction method based on the extraction of local information. The figure 12 resume the major stages to generate the set of features proposed by Lowe and our proposal to adapt it to palmprint contactless biometric systems called MSIFT.
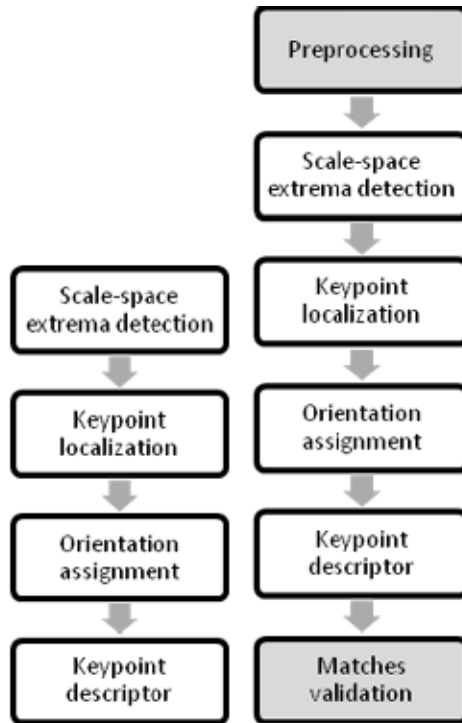
Fig. 12. On the Left the Lowe (Lowe, 2004) SIFT approach; on the right the contactless palmprint MSIFT approach proposed on this chapter.

The SIFT algorithm is based on detecting keypoints with similar propierties that are present in the reference and questioned image. In palmprint images acquired contactless from hand on movement with CMOS sensors of low quality, the images include blurring and several above mentioned distortion that reduce the ability of the SIFT algorithm to detect common keypoints. To alleviate such a problem we propose a preprocessing that highlight the interesting keypoints. The algorithm that preprocesses the image previous the application of the SIFT algorithm is called by us Modified SIFT (MSIFT) and consist of 6 steps:

Step 1. **Preprocessing:** In this chapter we propose a Gabor preprocessing to add robustness to SIFT approach. Assuming that the reference and questioned hand have similar orientation inside the image which is achieve during the segmentation stage. The real 2D Gabor filter used to process the palmprint image is defined by:

$$G(x, y, \theta, u, \varphi) = \frac{1}{2\pi\varphi^2} \exp\left\{-\frac{x^2 + y^2}{2\varphi^2}\right\} \cos\{2\pi(ux\cos\theta + uy\sin\theta)\} \tag{16}$$

where u is the frequency of the sinusoidal wave, θ defines the orientation selectivity of the function, and φ is the standard deviation of the Gaussian envelope. In this chapter we used a Gabor filter setting with $\theta = 0$, $\varphi = 2.0$ and $u = 0.1$. Greater robustness against brightness variation is assured by turning the discrete Gabor filter to average zero.

$$G'(x,y,\theta,u,\varphi) = G(x,y,\theta,u,\varphi) - \frac{\sum_{i=1}^{2n+1}\sum_{j=1}^{2n+1} G(i,j,\theta,u,\varphi)}{(2n+1)^2} \qquad (17)$$

Step 2. **Scale-space extrema detection:** It is applied over all scales and image locations. It is based on difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation. The input data is transformed to the space $L(x,y,\sigma)$ as follows:

$$L(x,y,\sigma) = g(x,y,\sigma) * I'(x,y) \qquad (18)$$

where $*$ corresponds to convolution operator, $I'(x,y)$ is the preprocessed input image and $g(x,y,\sigma)$ is a Gaussian function with bandwidth σ. The difference-of-Gaussian function is defined as:

$$D(x,y,\sigma) = (g(x,y,k\sigma) - g(x,y,\sigma)) * I'(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \qquad (19)$$

Step 3. **Keypoint localization:** A detailed model is fit to determine location and scale of each candidate location. The interpolation is done using the quadratic Taylor expansion of the Difference-of-Gaussian scale-space function $D(x,y,\sigma)$ with the candidate keypoint as the origin. This Taylor expansion is given by:

$$D(x) = D + \frac{\partial D^T}{\partial x} + \frac{1}{2} x^T \frac{\partial^2 D^T}{\partial x^2} x \qquad (20)$$

where the maxima and minima of D and its derivatives are evaluated at the candidate keypoint and $x = (x,y,\sigma)$ is the offset from this point, Fig 13.
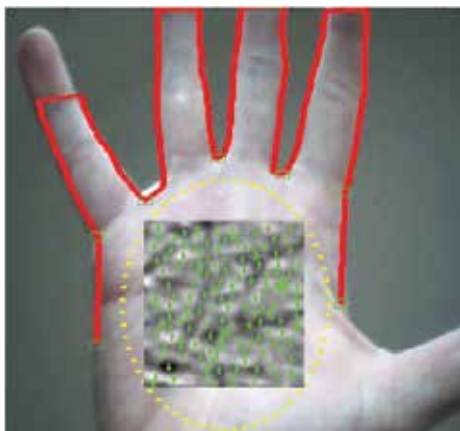

Fig. 13. Invariant area of the hand palm and the MSIFT features overlapped to the palm

Step 4. **Orientation assignment:** In our experiments we had used 16 orientations for each keypoint location based on local image gradient directions. For an image sample $L(x,y)$ at scale σ, the gradient magnitude, $m(x,y)$, and orientation, $\theta(x,y)$, are processed using pixel differences

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{21}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{22}$$

Keypoint descriptor: Around each keypoint, the local gradients are measured at the selected scale to obtain a descriptors vector $\{d_i\}_{i=1}^{M}$ with M keypoints. Once the keypoints are extracted, the query image is matched and compared with each of the features extracted with the corresponding images in the registration database (from the training feature sets). The verifier evaluates the number of matches between a questioned and the training images. Let $\{d_i^t\}_{i=1}^{M}$ and $\{d_i^q\}_{j=1}^{L}$ be the set of training and questioned keypoint descriptors respectively. The distance between keypoint descriptors is calculated from:

$$D_d(i,j) = \left\| d_i^t - d_j^q \right\|^2 \tag{23}$$

Where $\|\cdot\|$ is the Euclidean norm. We define a match between a training d_i^t and a questioned d_i^q keypoint when:

$$1.5 D_d(i,j) < \min\{D_d(i,n)\}_{n=1}^{L} \tag{24}$$

with $n \neq j$. The threshold is estimated heuristically during the training stage and it is not particularly sensitive to values in the range of 1.2 to 1.7.

Matches Validation: matches validation is common on fingerprint and other biometric feature approaches. In this chapter we propose a validation based on coordinates distance between keypoints to improve the SIFT performance on contactless palmprint biometrics. The hypothesis is that the coordinates from two keypoints matched must be similar if we correct the average displacement from all the matches. Let $c_i^t = \{x_i^t, y_i^t\}_{i=1}^{M}$ and $c_j^q = \{x_i^q, y_i^q\}_{i=1}^{L}$ be the set of training and questioned keypoint coordinates respectively. The distance between coordinates is calculated from:

$$D_c(i,j) = \left\| c_i^t - c_j^q \right\|^2 \tag{25}$$

where $\|\cdot\|$ is the Euclidean norm. We define a match between a training c_i^t and a questioned c_i^q keypoint when

$$D_c(i,j) \leq \frac{1.5}{M} \sum_{i=1}^{M} \left\| c_i^t - c_j^q \right\|^2 \tag{26}$$

Due to high pose variance in contactless imaging we used a 1.5 weight factor to allow small alignment errors between palms.

The maximum number of matches between the questioned and the training set is the similarity score. If the similarity score is grater than a threshold, the questioned image is autheticated.

5. Scores combination

Combining scores obtained from different procedures is a usual way of improving the performance of a biometric scheme. In this section we propose a method to combine the scores obtained from Geometry, MSIFT and OLOF. Figure 14 shows the distribution of genuine and imposter matching scores from the three feature extraction approaches. We can ascertain that the matching scores from the both features are widely separated. The distribution of matching scores also suggests that the matching scores from the two matchers are likely to be uncorrelated and therefore more effectively employed for combination.
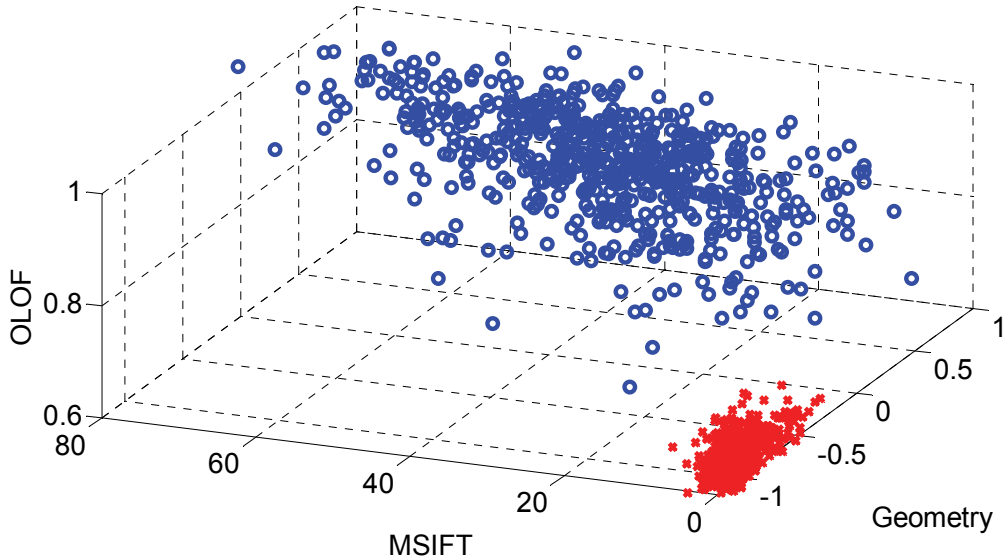

Fig. 14. Scores distributions for genuines (blue) and impostors (red) obtained for geometry, OLOF and MIFT.

Previous to combinate scores, we normalize the LSVM scores and OLOF scores based on *max/min* approach (Jain et al, 2005). The scores coming from the Hamming distance are not normalized. Now, it is possible to combine them at score level fusion based on a linear score combination functions as:

$$s_{comb} = w s_{geom} + (1-w) s_{palm} \tag{27}$$

Where s_{geom}, and $s_{palm} = (s_{OLOF} + s_{MSIFT})/2$ are the scores obtained with the image acquired in the visible and IR band respectively, w is the weighting factor and S_{comb} is the combined score which will be used for verify the input identity.

The value of w is obtained as follows. Let $s_{geom}^g(i)$ and $s_{palm}^g(i)$, $1 \le i \le N_g$ the scores of the genuine training samples in the visible and IR band respectively. A genuine score is obtained using two features vectors from the same user. Let $s_{geom}^f(i)$ and $s_{palm}^f(i)$, $1 \le i \le N_f$ the scores of the impostor training samples of in the visible and IR band respectively. An impostore score is obtained using features vectors from two different users (user x try to spoof the identity of user y). A distance measure between the distribution of genuine and impostor scores is obtained in visible band as follows:

$$\Delta_{geom} = (m_{geom}^g - m_{geom}^f)^T \sum (m_{geom}^g - m_{geom}^f) \tag{28}$$

where the means are calculated as:

$$m_{geom}^g = \sum_{i=1}^{N_g} s_{geom}^g \Big/ N_g \tag{29}$$

$$m_{geom}^f = \sum_{i=1}^{N_f} s_{geom}^f \Big/ N_f \tag{30}$$

And $\Sigma = (\Sigma_{geom}^g + \Sigma_{geom}^f) / 2$ with $\Sigma_{geom}^g = \Sigma_{i=1}^{N_g} (s_{geom}^g(i) - m_{geom}^g)^2 / N_g$ and $\Sigma_{geom}^f = \Sigma_{i=1}^{N_f} (s_{geom}^f(i) - m_{geom}^f)^2 / N_f$ the covariance matrixes.

The distance between genuine and forgeries for the palm scores Δ_{palm} is obtained in the same way. The weighting factor is obtained as: $w = \Delta_{palm} / (\Delta_{palm} + \Delta_{geom})$.

6. Experiments and analysis

The database acquired with the proposed device consists of 1000 images captured in one session of 100 users. The image was acquired automatically in a supervised experiment.
Each experiment was performed as follows. We randomly selected four hands from each user to train and left the remaining hands for testing. Table 1 list the average EER rates after repeating the procedure ten times.

Features	EER
Geometry	0.88%
Palmprint OLOF	0.98%
Palmprint MSIFT	0.31%
Score Fusion	0%

Table 1. Averaged EER Obtained by the Hand Biometric System
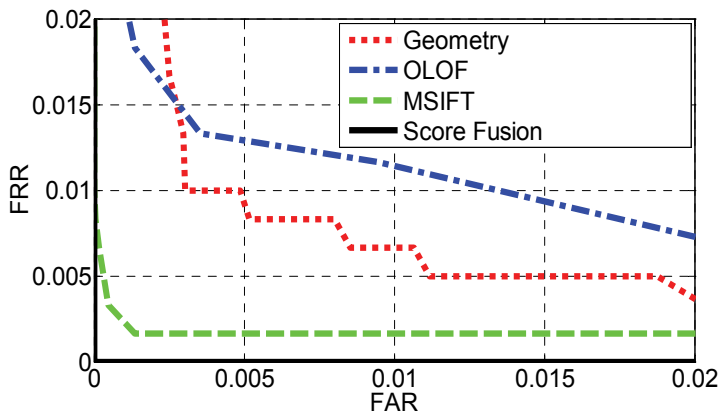

Fig. 15. DET curves of the proposed device. Dotted line: DET curve of geometrical device. Dash Dotted line: DET curve of OLOF features. Discontinuous line: DET curve of MSIFT features. Continuous line: DET curve of the combined scheme (OLOF+SIFT-Geometry).

The results show how MSIFT outperform OLOF and Geometry approaches. Geometry approach show better performance than OLOF. The fusion of all biometrics improves the performance with 0% of EER with our database.The DETs curves can be seen at Fig 15.

6.1 Computational time comparison

An analysis of the computational load of the biometric device is considered. Table 2 provides the executions times using a Dual Core processor at 1.66GHz programmed in the Matlab language for obtaining the geometric and palm texture parameters, the verification time, and (for the case of the analyzed scheme) the time that the ASM model toke to segment the hand in the visible image. Working with just the geometry, the analyzed devices answer in less than 1 second, which is appropriate for real time applications. In the case of IR plus visible images it takes less than 3 seconds, which should be speed up for real time applications.

Features	Run Time
Geometry	0.52 sec.
ASM	2.11 sec.
OLOF	0.07 sec.
MSIFT	1.51 sec.
Verification	0.38 sec.

Table 2. Time Consuming for Parameter Extraction and verification

7. Conclusions

This chapter has proposed a bispectral hand biometric system which acquires hand images in visible and IR band. The acquisition devices are two webcams, one per band, and the hand aquisition is contactless.The database used was built in an operational environment with a supervised enrollment.

The infrared images were used for geometric measures and the visible image for palmprint parameterization. An Active Shape Model was used to segment the hand in the visible image and a Least Square Support Vector Machine performed verification. An equal error rate of 0% was obtained combining biometries at score level.

Table 3 presents a summary of the most promising related work on contactless hand authentication.

Reference	Methodology	Database	Subjects	EER(%)
(Kumar 2008)	**Cohort Information**	**IITD (public)**	**235**	**1.31%**
(Hao et al, 2008)	**Multispectral Palmprint**	**Proprietary**	**165**	**0.5%**
This Chapter	**Geometry, Texture**	**Proprietary**	**100**	**0.0%**

Table 3. Releated work on contactlesshand authentication

8. Acknowledgment

9. References

Badrinath G. S. & Gupta P. (2009) "Robust Biometric System Using Palmprint for Personal Verification" *in Proceedings of International Conference on Biometrics*, Vol. 558, (2009), pp. 554-565.

Bulatov Y; Ambawalikar S. J. Kumar P. & Sethia S. (2004) "Hand Recognition using Geometric Classifiers", in *International Conference on Bioinformatics and its Applications*, Florida (December 2004).

Cootes T.F.; Taylor C.J.; Cooper D.H. & Graham J. (1995) "Active Shape Models – Their Training and Application", in *Computer Vision and Image Understanding*, Vol. 61, No. 1, (January 1995), pp. 38-59.

Doi J. and Yamanaka M. (2003) "Personal authentication using feature points on finger and palmer creases," *in the 32nd Applied Imagery Pattern Recognition Workshop*, (2003), Washington D.C.

Ferrer M. A.; Morales A.; Travieso C. M.; Alonso J. B. (2007) "Low cost multimodal biometric identification system based on hand geometry, palm and finger print texture" *in Proceedings of the 41st Annual IEEE International Carnahan Conference on Security Technolology*, (2007), pp. 52-58.

Haeger S. (2003) "Feature Based Palm Recognition", *Technical Report Univ. South Florida.* Tampa, (December 2003) Florida.

Han C. C.; Cheng H.L.; Lin C.L.; & Fan K.C. (2003) "Personal Authentication using palm print features", in *Pattern Recognition*, vol. 36, (2003), pp. 371-381.

Hao Y.; Sun Z.; Tan T. & Ren C. (2008) "Multi-spectral palm image fusion for accurate contact-free palmprint recognition," *in Proceedings of International Conference on Image Processing*, (2008), pp. 281-284.

Jain A. K.; Ross A. & Büke B. (1999) "A Prototype Hand Geometry based Verification System", in *Proceedings o 2nd International Conference on Audio and Video Based Biometric Person Authentication*, March 1999, pp. 166-171.

Jain A.K.; Bolle R. & Pankanti S. (2001) *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 2001.

Jain A. K.; Nandakumar K. & Ross A. (2005) "Score normalization in multimodal biometric systems" *Pattern Recognition*, (2005), vol. 38.

Kong W.K. & Zhang D. (2004) "Competitive Coding Scheme for Palmprint Verification," *in Proceedings of the 17th International Conference on Pattern Recognition*, vol.1, (2004), pp. 520-523.

Konga A.; Zhang D. & Kamelc M. (2009) "A survey of palmprint recognition", in *Pattern Recognition*, vol. 42, 2009, pp. 1408-1418.

Kumar A.; Wong D.C.M; Shen H.C. & Jain A.K. (2003) "Personal Verification using Palmprint and Hand Geometry Biometrics", *in Proceedings of the 4th International Conference on Audio and Video Based Biometric Person Authentication*, (June 2003).

Kumar A. & Zhang D. (2004) "Integrating shape and texture for hand verification", in *Proceedings of Third International Conference on Image and Graphics*, (2004), pp.222-225.

Kumar A.; Wong D.C.M.; Shen H.C. & Jain A.K. (2006) "Personal authentication using hand images", in *Pattern Recognition Letters*, vol. 27, no.13, (2006), pp. 1478-1486.

Kumar A. (2008) "Incorporating Cohort Information for Reliable Palmprint Authentication," *in Proceedings of Indian Conference on Computer Vision, Graphics and Image Processing*, Bhubaneswar (India), (December 2008), pp 112-119.

Li Q.; Qiu Z. & Sun D. (2006) "Feature-level fusion of hand biometrics for personal verification based on Kernel PCA", in *Proceedings of International Conference on Biometrics*, (2006), pp 744-750.

Lowe D. G. (2004) "Distinctive image features from scale-invariant keypoints" on International Journal of Computer Vision, vol. 2, no. 60, (2004), pp. 91 110.

Morales A.; Ferrer M. A.; Alonso J.B.; Travieso C. M. (2008) "Comparing infrared and visible illumination for contactless hand based biometric scheme," *in Proceedings of the 42nd Annual IEEE International Carnahan Conference, on Security Technology*, (2008). pp. 191 – 197.

Morales A.; Ferrer M. A. & Kumar A. (2010) "Improved Palmprint Authentication Using Contactless Imaging"*in Proceedings of the Fourth IEEE International Conference on Biometrics Theory, Applications and Systems*, (September 2010), Washington.

Öden C.; Erçil A.; & Büke B. (2003) "Combining implicit polynomials and geometric features for hand recognition", *in Pattern Recognition letters*, vol. 24, (2003), pp. 2145-2152.

Ribaric S.; Ribaric D. & Pavesic N. (2003) "Multimodal biometric user-identification system for network-based applications", in *IEEE Proceedings, Vision, Image and Signal Processing*, vol. 150, no.6, (2003), pp. 409-416.

Ribarić S. & Fratric I. (2005) "A Biometric Identification System Based on EigenPalm and EigenFinger Features", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*", vol. 27, no. 11, (November 2005), pp. 1698-1709.

Sanchez-Reillo R.; Sanchez-Avila C. & Gonzalez-Marcos A. (2000) "Biometric identification through hand geometry measurements," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, (2000), pp. 1168–1171.

Sun Z., Tan T.; Wang Y. & Li S. Z. (2005) "Ordinal palmprint representation for personal identification,"*in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, (2005), pp. 279- 284.

Suykens J. A. K.; Gestel T. V.; Brabanter J. D.; Moor B. D. & Vandewalle J. (2002) "Least Squares Support Vector Machines" *World Scientific PublishingCo.*, Pte, Ltd (2002).

Travieso Carlos M; Alonso J. B., David S. & Ferrer Miguel A. (2004) "Optimization of a biometric system identification by hand geometry" *in Proceedings of Complex systems intelligence and modern technological applications*, Cherbourg, France, (September 2004), pp. 581-586.

Wong A. & Shi P. (2002) "Peg-free hand geometry recognition using hierarchical geometry and shape matching," in *Proceedings of IAPR Workshop on Machine Vision Applications*, Nara, Japan, (December 2002), pp. 281–284.

Yörük E.; Konukoglu E.; Sankur B. & Darbon J. (2006) "Shape-Based Hand Recognition", in *IEEE Transactions on Image Processing*, vol. 15, no. 7, (July 2006), pp. 1803-1805.

Zheng G.; Wang C.-J.; and Boult, T. E. (2007) "Application of Projective Invariants in Hand Geometry Biometrics", *IEEE Transactions on nformation Forensics and Security*, vol. 2, no. 4, (December 2007), pp. 758-768.

Biometric Application in Fuel Cells and Micro-Mixers

Chin-Tsan Wang
National I Lan University
Taiwan

1. Introduction

1.1 Fuel cells related

Over the past few decades, there has been a growing interest in the fuel cell system of power generation and micro-mixers because it has been widely applied in mobile and microfluidic systems, respectively. Regarding the fuel cell systens, Proton Exchange Membrane Fuel Cells (PEMFC) seem to be one of the better solutions for a vehicular power source in the future due to their high power density, solid electrolytes, low corrosion and low-temperature operation. However, some issues are still of concern, in particular the cost, the size, the weight and the complexity of peripheral devices (Marsala et al., 2009). Bipolar plates are one of the most important and expensive components of PEM fuel cells. This is because they account for more than 60% of the total weight and 30% of the total cost of the system. Therefore, improving or addressing a novel flow slab design seems to be workable for improving these issues with respect to weight, volume and cost.

From past studies, it is known that the uniformity of oxygen distribution in cathode channels significantly affects the performance of PEM fuel cells. Different types of flow-fields have been addressed and studied to improve power performance (Mei et al., 2006; Weng et al., 2005; Yan et al., 2008). Results show that increasing the fuel rate (Mei et al., 2006) and higher flow field uniformity (Weng et al., 2005; Yan et al., 2008) are useful to the performance of fuel cells. In addition, pressure drop would be one of the important factors for flow-field design because it can simultaneously cause an excess of motor power dissipation (Yan et al., 2007). In other aspects, the aspect ratio of the channel would also simultaneously affect the pressure drop and even the cell performance (Perng et al., 2009). Also of importance, the new flow slab designs originating from bionic features, addressed by Kloess et al., (2009) and Wang et al., (2009), are of significance to fuel cells but have rarely been noted. A biophysical flow slab design, created to mimic features of vascular flow networks, was employed by Wang et al., (2009) due to its excellent performance in the uniformity of flow distribution and lower pressure drop. Latterly, two new types of biometric flow slab, namely BFF1 and BFF2, originating from the prototype type were addressed by Wang et al., (2009) and confirmed by the performance of the cells (Wang et al., 2010).

Furthermore, it has been also extended to the study of microbial fuel cells to enable the improvement of power performance, (Chen, 2010), because of the desire for clean energy. The development of processes by which to generate biofuels and bioenergy have been of

special interest of late. Among these, microbial fuel cells have received increased attention. This process, which collects the electricity generated by microbes when they metabolize substrates, is considered to be one of the most efficient energy sources because no burning is required to produce energy (Watanabe, 2008). Also, the only raw materials needed to power fuel cells are simple organic compounds or even waste materials from other reactions (Watanabe, 2008; Lovely, 2008). There are still many obstacles that need to be overcome before this technology can be put to use. Currently the voltage and amperage generated by microbial fuel cells is so low that they have no useful applications (Watanabe, 2008; Lovely, 2008). In order to develop solutions to these problems, research is being done to engineer more efficient hardware for fuel cells in addition to understanding how different microbes interact with the anodes/cathodes when transporting electrons (Lovely, 2008; Bergel et al., 2008). As for the flow slab design of MFCs, there is an absence of sufficient discussion and research regarding the design of the flow channel and flow field (Hameler et al., 2006; Logan et al., 2004), and even less discussion as to why and how they could be applied to MFCs. However, a biometric flow channel applied to rumen microbial fuel cells (RMFCs) was first addressed by Chen (2010).

1.2 Passive micro-mixer related

A biometric concept could also be applied to the design of a passive micro-mixer because it is simple to operate and provides an excellent mixing performance under the condition of lower pressure (Wang et al., 2009). Recently, microfluidics have received a lot of attention in the development of automated miniaturized analytical devices in (bio)analytical chemistry. Microfluidics deals with microscale, physical phenomena of fluid and particle flows in microchannels that connect various functional sites on a miniaturized analytical device. Among the various functionalities, rapid mixing is crucial because biological analyses, like enzyme reactions, protein folding, and cell activation, require a rapid reaction process that can be controlled by the mixing of reactants. Unfortunately, mixing at a microscale mainly depends on molecular diffusion, resulting in an extremely slow process and long microchannel for complete mixing. This is because almost all microchannel flows are laminary, and the Reynolds number is so slow that turbulent mixing is hard to be achieved (Song et al., 2006).

As for micro-mixers, application fields of microchannel-based mixers encompass both modern, especialised issues such as sample preparation for chemical analysis in addition to the traditional, widespread usable mixing tasks, such as reaction, gas absorption, emulsification, foaming and blending (Bayer et al., 2003; Ehrfeld et al., 2000; Hessel et al., 2004; Jensen, 1998; Lowe et al., 2000). Moreover, they are suitable for integration with other devices.

Many passive micro-mixers have been developed in order to enhance and control mixing in a microchannel (Nquyen and Wu, 2005). A passive mixer uses special geometries embedded in a microchannel, such as grooves, rivets or posts, to increase the vorticity and, subsequently, to cause a chaotic advection (Johnes and Aref, 1998). Another type of passive mixer is the lamination mixer, which decreases the diffusion length and increases the contact area of fluids by splitting incoming streams into multiple substreams, and then laminating them into one stream again (Kamholz and Yager, 2002).

Concerning the most traditional passive micro-mixers, they have been constructed with straight fluid channels and designed with a combination of fillisters and/or fold paths to

enhance the mixing effect (Wong et al., 2003). However, the design of a straight channel requires a longer length to achieve the goal of uniform mixing. Hence, it is always associated with the problems of mixer size and full-field inspection. In addition, fluid mixing at the microscopic scale is far more difficult than that in macroscopic fluid devices. In a typical microfluidic device, viscosity dominates the flow and the fluid streams prefer to adopt laminar flow patterns. Thus, fluid mixing that depends primarily on molecular diffusion is very slow. To achieve optimal mixing, an efficient passive micro-mixer usually involves complex 3-dimensional geometries, which are utilized to enhance the fluid lamination, stretching and folding. As mentioned above, mixing in the passive micro-mixer occurs with the diffusion of molecules in the microsystem and the process is very slow. Therefore, the complex geometry, or long microchannel, should be utilized for efficient mixing, but would cause a large pressure drop and difficulties in the design and fabrication process. In order to overcome this, a biophysical micro-mixer that originates from the biometric concept with a higher flow uniformity and lower pressure drop would be utilized by Wang et al., (2009) to provide better flow mixing within a limited device.

2. Biometric concept applied to fuel cells

2.1 Fuel cell bionic flow slab design (Wang et al., 2009)

Fuel cells possessing high potency and low pollution are well-known and considered the new generation of power technology. However, fuel cell performance and efficiency must be improved. The cost, reliability, and safety issues must be considered in the realization of commercial fuel cells. To enhance fuel cell performance and reliability, it is necessary to learn more about the mechanisms that cause performance losses. These include non-uniform concentrations, current density distributions, high ionic resistance due to dry membranes, and high diffusive resistance due to cathode flooding. The flow field and water/thermal management fuel cells require optimal designs to achieve high performance and reliability. Flow-field plate design is one of the most significant factors that affects fuel cell efficiency.

This work presents a novel bionic concept flow slab design, which is shown in Figure 1, to improve fuel cell performance and compare it with other known flow slabs.

The variations in velocity and pressure drop uniformity are influenced by the wall effect and alter fuel cell performance. An index of the aspect ratio defined as AR=D/L was employed for 3D simulations at Re=10 and 100. Here, D is the flow channel depth and L is the flow channel width. Although flow field plate design is one of the most significant factors in fuel cell efficiency, the simultaneous effect of velocity uniformity and pressure drop on the performance of the system has rarely been examined. In this work, an index χ was used to quantitatively address the coupling effect between the velocity uniformity and pressure drop in the flow slab, as defined in (1):

$$\chi = (\frac{SD}{SD_p} + \frac{PD}{PD_p})^{-1}$$

(1)

where SD and PD indicate the standard deviation and pressure drop, respectively. The subscript p denotes that it is the value for parallel flow design, which is used as the basis when taking the ratio of the standard deviation and pressure drop. Therefore, when the value of χ is larger, fuel cell performance is better.

Numerical results obtained show that this novel biometric flow slab design will exhibit a better performance than traditional flow slabs, regardless of Reynolds numbers and aspect ratios, as it possesses a more uniform velocity and a lower pressure drop. Furthermore, the performance in the biometric flow slab's reaction area was determined to be superior (shown in Tables 1 and 2). Hence, increasing the bionic flow slab performance is worth investigation in order to obtain an optimal design, and the required numbers of inlet and outlet channels need to be studied. Different inlet and outlet numbers influence the velocity distribution and pressure drop variations, resulting in an integral performance. Two inlet channels and three outlet channels are suggested for the optimal bionic flow slab design. These findings show that the bionic concept and flow slab design addressed in this paper will be useful in enhancing fuel cell performance (Wang et al., 2009).
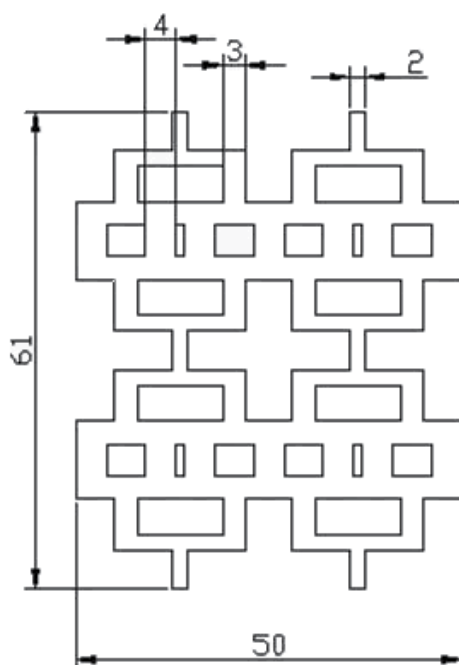

Fig. 1. Prototype of Biometric Flow Slab (unit: mm). Source: Wang et al., (2009)

	Parallel	Bionic	Net	Serpentine
χ	0.5	0.479	0.325	0.027
A_{Rea}	0.001438	0.001324	0.001	0.001512
χ/A_{Rea}	347.7	361.78	325	17.857

Table 1. Performance Index Versus Four Kinds of Flow Slabs at Re=10. Source: Wang et al., (2009)

	Parallel	Bionic	Net	Serpentine
χ	0.5	0.535	0.372	0.046
A_{Rea}	0.001438	0.001324	0.001	0.001512
χ / A_{Rea}	347.7	404.08	372	30.423

Table 2. Performance Index Versus Four Kinds of Flow Slabs at Re=100. Source: Wang et al., (2009)

2.2 Biometric flow slab applied to PEMFC (Wang et al., 2010)

As for the bipolar plates, they are one of the most important and expensive components of PEM fuel cells because they account for more than 60% of the total weight and 30% of the total cost of the system. Therefore, improving or addressing a novel flow slab design seems to be workable to improve these issues with respect to the weight, volume and cost. In this work, two kinds of novel biophysical flow slabs, namely BFF1 and BFF2, originating from the prototype of the biophysical flow slab shown in Figure 1, and due to their possession of a lower pressure drop and excellent flow uniformity, (Wang et al., 2009) shown in Figures 2 and 3,would be utilized in PEMFCs (Wang et al., 2010). They would then be compared with the two convectional flow slabs, the serpentine and parallel, which would be used for the investigation of cell performance.

The I–V cell polarization curves and I–W cell power density curves of the parallel, serpentine and two new biometric flow slabs were the first investigated and are shown in Figure 4. The results in Figure 4 show that serpentine and two biometric flow slabs (BFF1 and BFF2) have the appearance of a better performance than that of the parallel flow slab. The limited current densities at V_{cell}=0.27 for the serpentine, BFF1, and BFF2 compared with the parallel flow slab are increased by the amount of 58.19%, 58.48%, and 57.13%, respectively. When the operating voltage is lower than 0.57V, the performance of the parallel flow field seems to increase much more slowly than other flow slabs. This is because of its strong dependence on the distribution of the oxygen mass flow rate at the cathode GDL–CL interface, and a high oxygen mass flow rate will cause more oxygen to enter the CL for the electrochemical reaction.

Figure 5 shows the distribution and relation between the oxygen and liquid water at the three segments C-C1, C2-C3 and C4-C5 for BFF1. As the oxygen mass flow rate increases, the amount of liquid water from inlet to outlet decreases. The amounts of oxygen at the cross-section C-C1 and C4-C5 are less than C2-C3, resulting in lower current densities. Some baffles could be used and applied to promote the mass transport of C-C1 and C4-C5 in future studies (Perng et al., 2009).

Figure 6 indicates clearly that the BFF1 flow slab will produce a higher uniform distribution of current densities at the section of C-C1, C2-C3 and C4-C5. Hence, a higher performance for BFF1 would be expected because a higher uniform distribution of current density is one of the important factors for promoting the cell performance. Generally speaking, the lower the pressure loss is, the higher the net performance of the cell will be (Perng et al., 2009). To design a flow slab with a lower pressure drop, new flow slabs, named BFF1 and BFF2 respectively, were designed by the biophysical conception in this study.

Figure 7 displays the distribution and relation between oxygen and liquid water at the cross-sections of D-D1, D2-D3, D4-D5 and D6-D7 for BFF2. The trend of oxygen and liquid water distribution of D-D1, D2-D3, D4-D5 referred to are similar to C-C1, C2-C3 and C4-C5 of BFF1. The average oxygen distribution at the section of D6-D7, resulting from the shear stress, would be found to be the highest. Figure 8 shows that BFF2 would possess a better uniformity of flow distribution than BFF1. In addition, the shear stress would push more oxygen into CL for an electrochemical reaction, thus a greater current at the cross-section of D6-D7 could then be produced.

In this study, a pressure drop loss with respect to power density (Perng et al., 2009), defined in Equation (2), would be used to acquire a superior flow slab.

$$W_p = \frac{\Delta P A_{cha} V}{A_{total}} \qquad (2)$$

In this equation, W_p represents the cathode pressure drop loss, ΔP is the total cathode pressure drop of the fuel cell, A_{cha} is the cross-sectional inlet flow area of cathode, V is the fuel velocity at the inlet of cathode, and A_{total} is the reaction area. The pressure drop losses of the cathode and output power of the cell, with respect to a parallel, serpentine, BFF1 and BFF2 flow slab, would be calculated and listed in Table 3. Due to a high pressure drop in channels, the W_{net} of serpentine is lower than that of the BFF2 in spite of the fact that the W_{cell} of serpentine is higher than that of BFF2. In addition, the pressure drop of BFF2 is lower than that of BFF1. Hence, the net power of the four kinds of flow slab would be obtained and shown in Table 3. It shows that the novel biometric flow slab of BFF1 and BFF2 would have a better performance than that of the serpentine and parallel flow slabs (Wang et al., 2010).

To sum up, the total pressure drop and the uniformity of flow distribution are two important factors for flow slab design because of their significant influence on the performance of the PEMFC. In this study, the two biometric flow slabs, BFF1 and BFF2, addressed in this study would have a better cell performance than the serpentine and parallel flow slabs because they possess a higher uniformity of flow distribution and a stronger ability to remove the liquid water. The novel biometric flow slab would have an enhanced cell power performance compared to the serpentine and parallel flow slabs. These findings, with respect to biometric flow slabs, would be useful to improve the PEMFC and could even be expanded to other cell types. (Wang et al., 2010).

Flow field type	$\triangle P$ (Pa)	W_{cell} (w/m²)	W_p (w/m²)	W_{net} (w/m²)
Parallel	248	3529	4.4	3524.6
Serpentine	5137	5583	91	5492
BFF1	2073	5593	37	5557
BFF2	730	5546	13	5533

Table 3. Estimation of Pressure Drop Losses at an Operating Voltage of 0.27V.
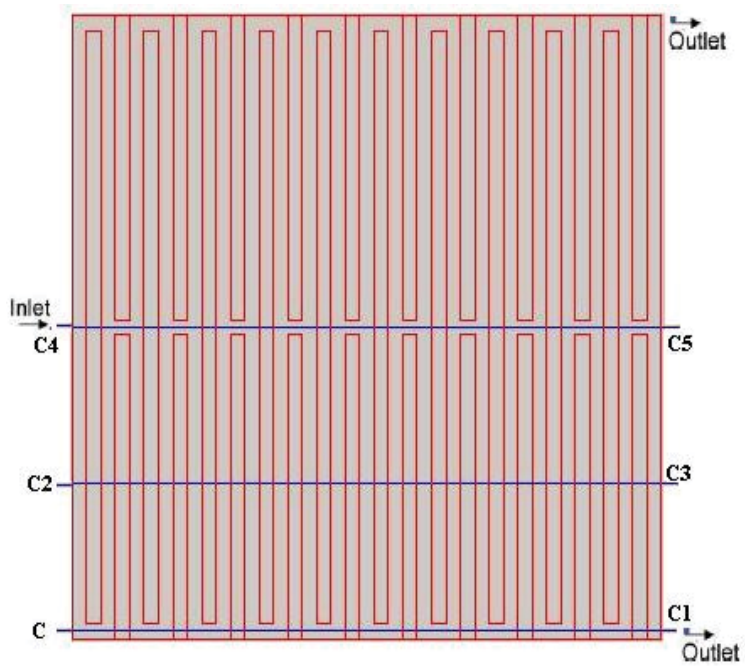Source: Wang et al., (2010)

Fig. 2. Biophysical Flow Slab (BFF1). Source: Wang et al., (2010)
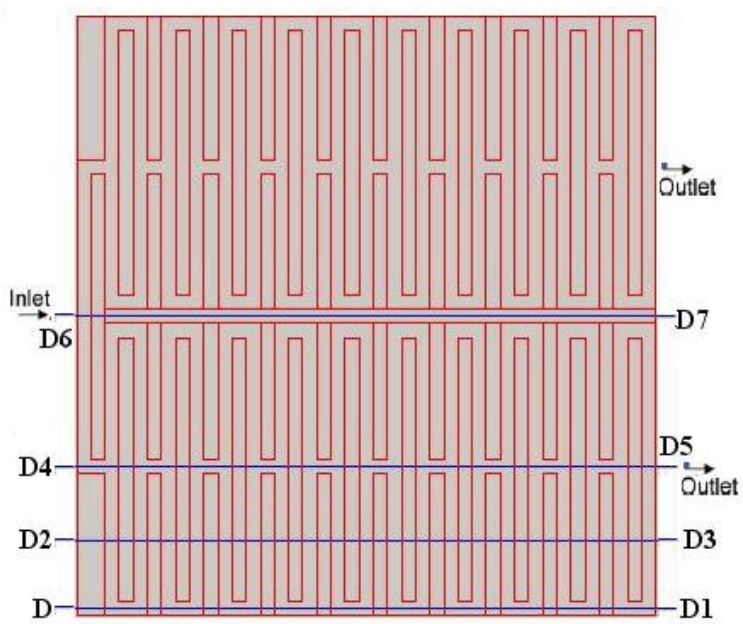

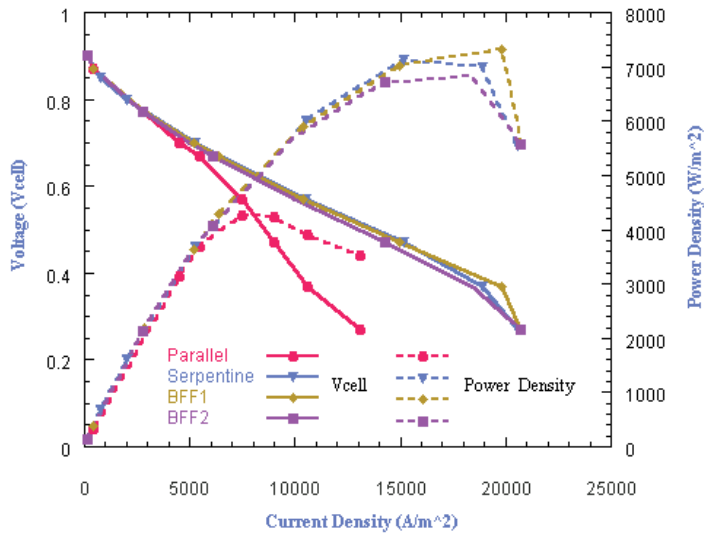Fig. 3. Biophysical Flow Slab (BFF2). Source: Wang et al., (2010)

Fig. 4. The *I–V*cell and *I–W*cell Curves for Types of Parallel, Serpentine, BFF1 and BFF2, respectively.
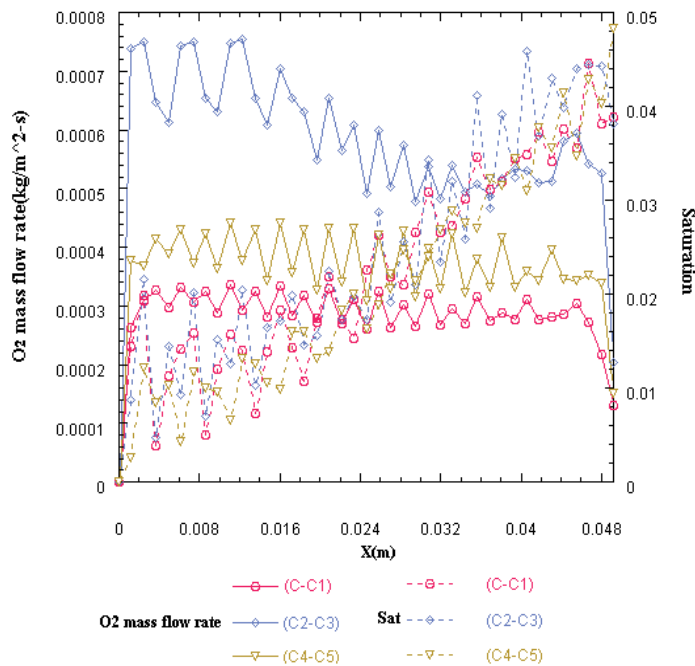Source: Wang et al. (2010)

Fig. 5. Oxygen Mass Flow Rate (kgm⁻²s) and Liquid Water Distributions at the sections of C-C1, C2-C3, C4-C5 Related to 0.7V for BFF1.
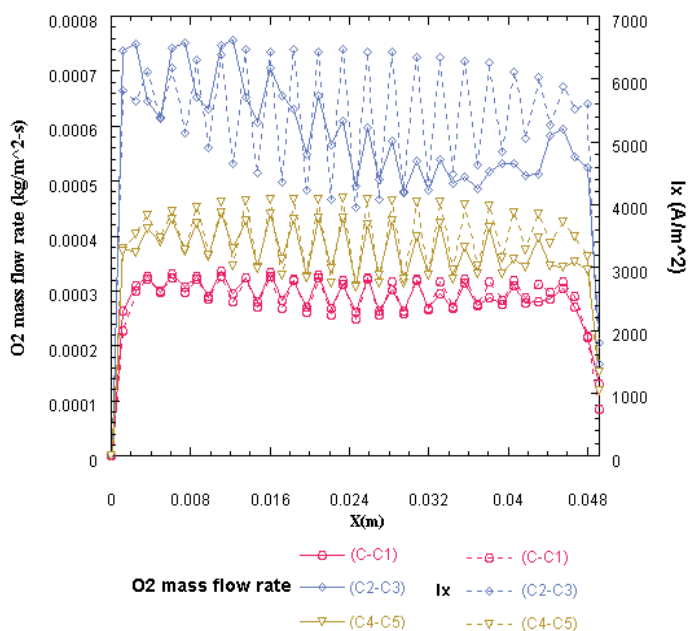Source: Wang et al., (2010)

Fig. 6. Oxygen Mass Flow Rate (kgm⁻² s) and Current Density Distributions (Am⁻²) at the Sections of C-C1, C2-C3, C4-C5 Related to 0.7V for BFF1.
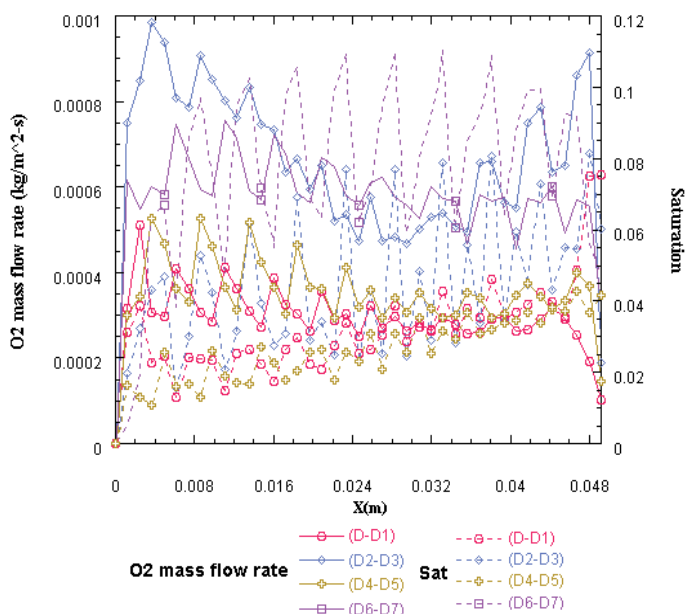Source: Wang et al., (2010)

Fig. 7. Oxygen Mass Flow Rate (kgm⁻² s) and Liquid Water Distributions at the Sections of D-D1, D2-D3, D4-D5, D6-D7 Related to 0.7V for BFF2.
Source: Wang et al., (2010)

Fig. 8. Oxygen Mass Flow Rate (kgm⁻² s) and Liquid Water Distributions at the sections of D-D1, D2-D3, D4-D5, D6-D7 Related to 0.7V for BFF2.
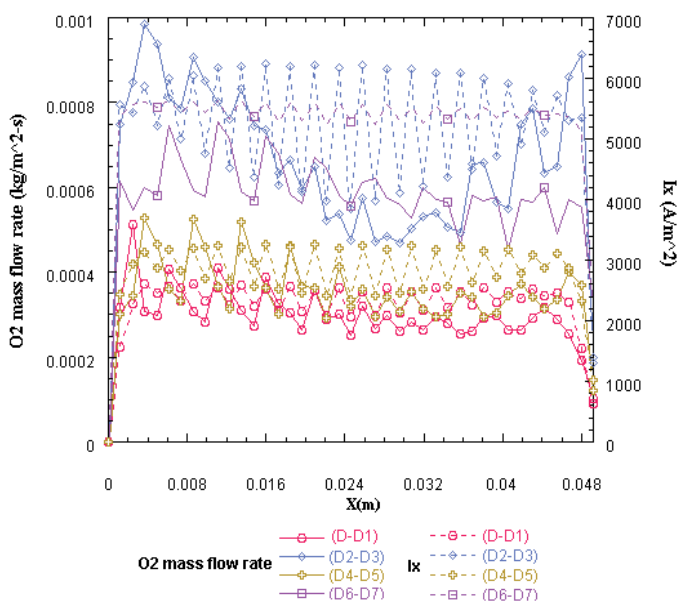Source: Wang et al., (2010)

2.3 Biometric flow slab applied to Microbial Fuel Cells (MFCs) (Wang et al., 2011)

In the academic studies of microbial fuel cells (MFCs), there is a significant absence of sufficient discussion and research regarding to the design of flowchannels and flow-fields (Hameler et al., 2006; Logan et al., 2004), and even less discussion as to why and how they could be applied to MFCs. However, the research of flow channels being applied in fuel cells have been proven to have a significant contribution to power performances (Wang et al., 2009; Sabir et al., 2005), especially with regards to fuel efficiency and power density (Sabir et al., 2005). A new biometric flow channel, shown in Figure 9 and applied in rumen microbial fuel cells (RMFCs), was first addressed by (Wang et al., 2011). Looking at Figure 10 and Table 4, the obstacle groups of No.A and No.C have a higher flow mixing efficiency inside the chamber of RMFCs. The obstacles can cause flow to split and recombine to enhance flow mixing. Since the Reynolds number is higher in the case of No.C, the flow convection at the inlet entrance of RMFCs is more intensive than in the case of No.A and also has a higher flow mixing. In Case No.D, without obstacles, flow separation is created due to a high Reynolds number (Lashkov et al., 1992; Jadhav et al., 2009) and the Coanda effect.

Therefore, proton exchange seems to be unevenly mixed because the main flow and the separation flow are almost without interaction. In addition, the electron and proton from the reactants will continue to be exhausted from the charged reaction of RMFCs, and finally creates a concentration loss in some regions of the chamber. Conversely, Case No.B does not experience that problem because of the lower Reynolds number, thus creating a smoother flow and more even reaction. Even though the flow obstacles do not exhibit noticeable benefits in the flow mixing, the effect on the flow field is overall beneficial to the electricity

system (Wang et al., 2009). This creates a better interaction on the surfaces of electrodes and proton exchange membranes, thus avoiding concentration loss and proving more efficient than flow fields that are uneven. In this study rumen microbes and plant fibers that acted as substrates were utilized in single chambers in the cases of both using obstacles and different Reynolds numbers respectively to investigate the power performance of RMFCs. The RMFC system with a biometric flow channel (with obstacles), and at a higher Reynolds number (Re= 496.18), will produce a higher power performance with a voltage potential and power density of 0.716 V and 0.022mW/m^2 respectively. This is much better than in the cases without obstacles showing a positive effect of a biometric flow channel on the power performance of RMFCs.
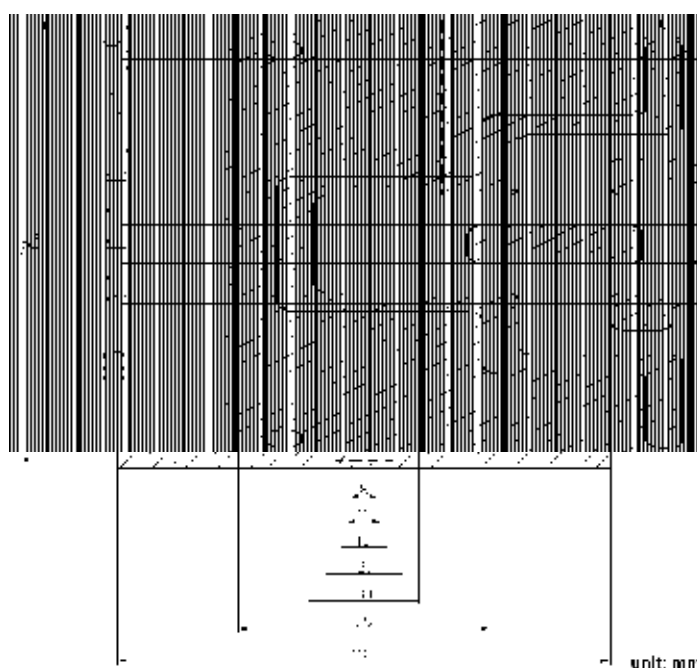

Fig. 9. Geometrical Dimensions of a Biometric Flow Channel for RMFCs. Source: Wang et al., 2011

MFCs	Re No.	Flow mixing efficiency at analyzed positions (%)				
		a	b	c	d	e
No.A	19.85	99.6	99.8	99.8	99.7	99.8
No.B		99.0	99.3	99.0	99.3	99.4
No.C	496.18	99.8	99.8	99.9	99.7	99.9
No.D		100.0	100.0	100.0	100.0	100.0

Table 4. Flow Mixing Efficiency Versus Different Flow Conditions and Cross-sections analyzed.
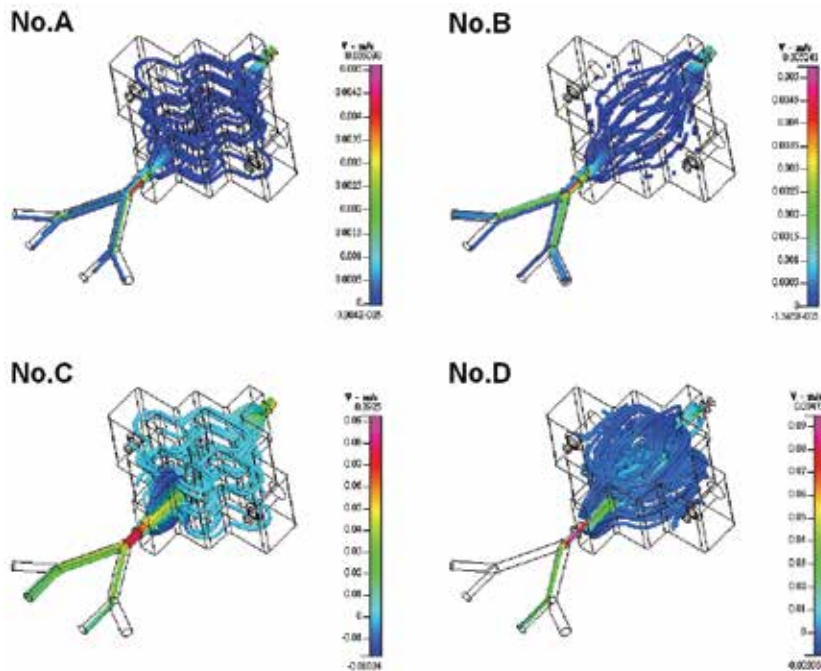Source: Wang et al., 2011

Fig. 10. 3D Flow Velocity Images Versus Different Flow Conditions and Cross-sections
Analyzed (shown in Table 4).
Source: Wang et al., 2011

3. Biophysical micro-mixer (Wang et al., 2009)

In this work, a biophysical concept was applied to passive micro-mixers, named as a
biophysical micro-mixer and shown in Figure 11, to promote mixing efficiency. The vertical
width of a channel would be gradually increased from 20 μm at the inlet to 40 μm at the
middle section of the device, and then gradually decreased along the flow downstream to
the outlet. During the flow transmission process, the flow would be split first and then
recombined with a flow motion. When the flow passes through the middle section of the
system, increasing interfaces were created exponentially by laminating the interfaces
continuously along the channel. In addition, the convection flow in the biophysical channels
had a high flow uniformity and low pressure drop to enhance the flow mixing (shown in
Figure 12). The mixing coefficient will approach 0.95 when the Reynolds number of the inlet
mid-channel is larger than 160. This result shows that the Reynolds number positively
affects mixing although it induces an increase in pressure drop (Wang et al., 2009).
Therefore, the prototype of a biophysical micro-mixer is simple and possesses a better

uniformity and lower pressure drop, so it can be expected to be useful to promote the mixing performance of passive micro-mixers when the mixing distance is restricted. These findings will be useful in the design of an optimal biophysical passive micro-mixer in further research. Parameters, such as the Reynolds number ratio and aspect ratio and their effect on mixing and pressure drop, required investigation because finding the optimal Reynolds number ratio, Rer, and aspect ratio, AR, is important for the operation of the micro-mixer.

To address the effect of the different inlet flow conditions on the mixing performance, a parameter denoted as Rer defined in (3) was set for the Reynolds number ratio:

$$\mathrm{Re}_r = \frac{\mathrm{Re}_1 + \mathrm{Re}_3}{\mathrm{Re}_2} \tag{3}$$

Here, the operational Reynolds number defined in (4) was set in the range of Re=0.5 to 10:

$$\mathrm{Re} = \frac{\rho U_{ave} W}{\mu} \tag{4}$$

Where ρ is the density of the fluid, U_{ave} is the average velocity of the inlet channel; W, whose value is 20 μm, represents the width of inlet channel I_2 and the outlet channel. μ is the dynamic viscosity of the working fluid.

In addition, the aspect ratio, AR, ranging from 0.5 to 10 is defined in (5) and was investigated in order to study the sidewall effects on mixing performance:

$$AR = \frac{D}{W} \tag{5}$$

where D is the depth of the channel and W is fixed at 20μm for the inlet at mid-channel.

Hence, the Reynolds number ratio was decided and based on the variations of inlet Reynolds numbers from Re = 0.5 to 10 for the inlet channels. In addition, variations of aspect ratio were set as 0.5, 1, 2 and 10 for determining the sidewall effect on mixing and pressure. The results, shown in the Table 5, are addressed as follows:

First, the optimal Reynolds number ratio was Rer = 0.85, because of its outstanding mixing performance at different aspect ratios. Second, the sidewall effect will influence the variations in pressure drop and mixing performance, and increasing the AR will also decrease the pressure. An optimal aspect ratio with the highest mixing effect was found at AR = 2, which exhibited a good mixing for studied cases. In addition, the inlet angle of the side-channels and its effect on mixing and pressure was considered in the design of the micro-mixer. Hence, a variety of inlet angles of the side-channels, represented by θ, were executed with Reynolds number ratios ranging from Rer = 0.5 to 2 in the case of Re2 = 1 and its relationship to mixing performance and pressure drop are shown in Table 5.

The results of Table 6 show that a side-channel inlet angle of 30° was a better choice because it possesses a better mixing effect and has a lower pressure drop. These findings will be useful in the optimal design of a passive micro-mixer based on biophysical concepts in further experimental studies.
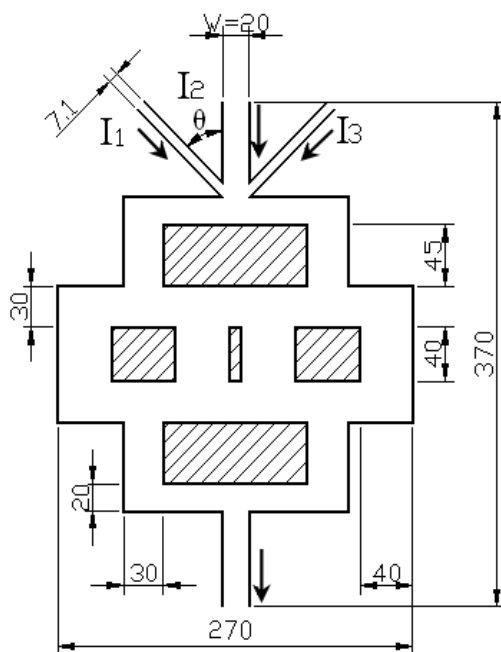
Fig. 11. Prototype of a Biophysical Micro-mixer (unit: μm)
(Arrows indicate the inlet and outlet flow direction).
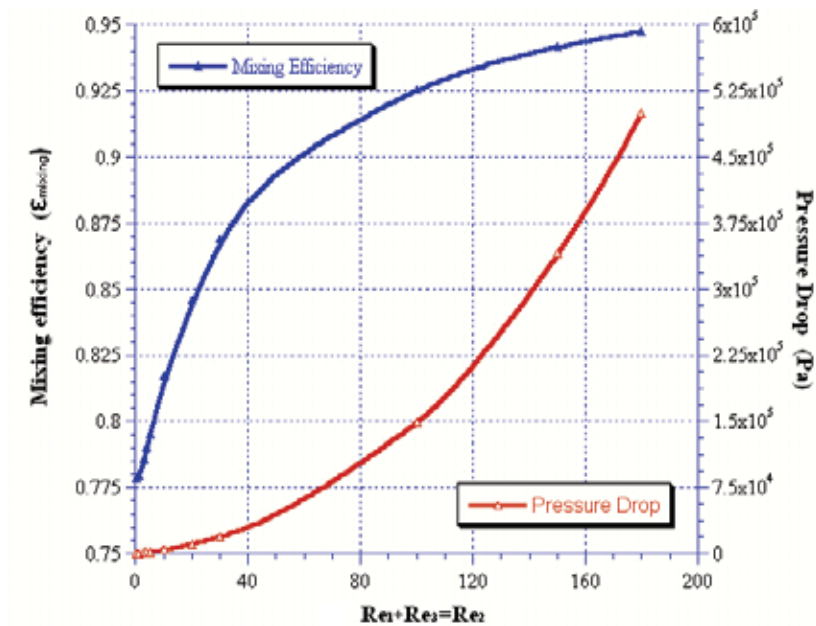Source: Wang et al., (2009)

Fig. 12. Reynolds Number Effect Versus the Mixing and Pressure Drop at Rer = 1
Source: Wang et al., (2009)

	Rer	0.5	0.85	1	2	1
	Re1	0.5	0.85	1	2	10
	Re2	1	1	1	1	10
ΔP	AR = 0.5	1890.52	2333.07	2522.87	3790.44	25999.58
ΔP	AR = 1	746.72	922.50	1891.34	1503.66	10908.40
ΔP	AR = 2	449.06	555.37	601.07	908.16	6915.66
ΔP	AR = 10	320.69	396.79	429.52	649.51	4956.02
ε_{mixing}	AR = 0.5	0.72	0.79	0.78	0.61	0.76
ε_{mixing}	AR = 1	0.73	0.79	0.78	0.59	0.80
ε_{mixing}	AR = 2	0.69	0.80	0.80	0.67	0.85
ε_{mixing}	AR = 10	0.73	0.79	0.79	0.62	0.83

Table 5. Variations of Mixing Coefficient (ε_{mixing}) and Pressure Drop (ΔP; unit: Pa) Versus the Reynolds Number Ratio (Rer) and Aspect Ratio (AR).
Source: Wang et al., (2009)

Inlet angle of side channel, θ	Rer = 0.5		Rer = 0.85		Rer = 1		Rer = 2	
	ε_{mixing}	ΔP	ε_{mixing}	ΔP	ε_{mixing}	ΔP	ε_{mixing}	ΔP
90°	0.737	310.901	0.786	384.537	0.771	417.240	0.581	632.116
60°	0.738	300.758	0.791	371.628	0.776	403.076	0.580	609.910
45°	0.739	300.025	0.796	371.122	0.779	401.692	0.584	607.082
30°	0.738	300.781	0.803	371.240	0.790	402.526	0.589	607.998
0°	0.729	289.187	0.764	356.386	0.750	385.829	0.575	586.702

Table 6. Inlet Angle of Side-channel Versus the Mixing (ε_{mixing}) and Pressure Drop (ΔP; unit: Pa) with Variations of Reynolds Number Ratios ranging from Rer = 0.5 to 2 in the case of $Re_2 = 1$.
Source: Wang et al., (2009)

4. Conclusion

In this chapter, the biometric concept was applied to the fuel cells, microbial fuel cells and micro-mixer. To sum up, some results are addressed as follows:

In the study of a biometric fuel cell, the novel flow slab provides a better performance when using the bionic concept because it can control the velocity distribution, pressure drop, coupling effect and has a stronger ability to remove the liquid water and so providing a better power performance of cells. These findings, with respect to a biometric flow slab, would be useful to improve PEMFCs and could even be expanded to other types of cells. In addition, a new design of biometric flow channel was also applied to RMFCs. Experiments in a circulation system were executed by using a double two-inlet Y-type inlet channel and connecting it with the RMFCs. The biometric flow channel would create a more uniform flow field with obstacles than one without, regardless of the Reynolds number. An extra voltage output of 0.2 V, based on the example without obstacles, was provided as in the case of the one with. This further explains the design of biometric flow channels having a greater, more positive effect on power performance.

In the study of the biometric micro-mixer, a novel micro-mixer design based on the biophysical concept was addressed. The prototype was simple and possessed a better flow uniformity and lower pressure drop, so it could be expected to be useful to promote the mixing performance of passive micro-mixers when the mixing distance was restricted. The highest mixing coefficient with $\varepsilon_{mixing} = 0.876$ occurred at a Reynolds number ratio, Rer = 0.85. These findings will be useful in the design of an optimal biophysical passive micro-mixer in future research and even show the feasibility and potential of the biometric concept widely applied in biochemical, biological andchemical analysis, along with fuel cells and bioenergy.

5. References

A. Ter Heijne, H.V.M. Hamelers, V. De Wilde, R.A. Rozendal and C.J.N. Buisman. (2006) A bipolar membrane combined with ferric iron reduction as an efficient cathode system in microbial fuel cells. *Environ. Sci. Technol,* Vol. 40：5200-5205.

A. Y. Lashkov, I. N. Sokolova and E. A. Shumilkina. (1992) Jet flow over ribbed curved surfaces. *Fluid Dynamics*, Vol.27 (1) ：135-137.

Bayer, T., Himmler, K., Hessel, V. (2003) Don't be baffled by static mixers. *Chemical Engineerinf*, Vol. 5：2-9.

B. Min and B.E. Logan. (2004) Continuous electricity generation from domestic wastewater and organic substrates in a flat plate microbial fuel cell. *Environ. Sci. Technol,* Vol. 38：5809-5814.

Dumas, C., Basseguy, R., Bergel, A. (2008) DSA to grow electrochemically active biofilms of Geobacter sulfurreducens. *Electrochimica Acta*, Vol. 53：3200-3209.

Ehrfeld, W., Hessel, V., Lowe, H. (2000) Microreactors. *Wiley-VCH*, Weinheim.

G.S. Jadhav and M.M. Ghangrekar. (2009) Performance of microbial fuel cell subjected to variation in pH, temperature, external load and substrate concentration. *Bioresour. Technol*, Vol. 100 ： 717-723.

Hessel, V., Hardt, S., Lowe, H. (2004) Chemical Micro Process Engineering-Fundamentals, Modelling and Reactions. *Wiley-VCH*, Weinheim.

Jensen, K. F. (1998) Smaller, faster chemistry, *Nature.* Vol. 393 (6) ：735-736.

Johnson, S. W., Aref, H. (1998) Chaotic advection in pulsed source-sink systems. *Physics of Fluids*, Vol. 31：469-485.

Kamholz, A. E., Yager, P. (2002) Molecular diffusive scaling laws in pressure-driving nicrofluidic channels: deviation from one-dimensional Einstein approximations. *Sensors and Actuators B*. Vol. 82(1) ：117-121.

Kloess J. P., Wang X., Liu J., Shi Z.& Guessous L. (2009) Investigation of bio-inspired flow channel designs for bipolar plates in PEM fuel cells. *J Power Sources*, Vol.188 (1)：132-140.

Lovely, D. R. (2008) The microbe electric: conversion of organis matter to electricity. *Curr Opin Biotech-nol*, Vol. 19 ：564-571.

Lowe, H., et al., (2000) Micromixing technology. *In: Fourth International Conference on Microreaction Technology*, IMRET 4, Atlanta, USA. Ai.I. Ch. E. Topic Conference Proceedings：31-47.

Maeng, J. S., Yoo, K., Song, S. (2006) Modeling for fluid mixing in passive micromixers using the vortex index. Journal of the Korean Physical Soceity, Vol. 48 (5) ：902-907.

Marsala G., Pucci M., Vitale G., Cirrincione M.& Miraoui A.(2009) A prototype of a fuelcell PEM emulator based on a buck converter. *Appl Energ*, Vol. 86：2192-2203.

Mei S. C., Yan W. M. , Yang C. H. , Soong C.Y. & Chen F.(2006) Experimental studies on optimal operating conditions for different flow field designs of PEM fuel cells. *J Power Sources*, Vol.160：284–292.

Nguyen, N. T., Wu, Z. (2005) Micromixers-a review. *J. Micromech. Microeng*. Vol. 15, Vol. 15：R1-R6.

Perng, S. W., Wu, H. W., Jue, T. C., Cheng, K. C. (2009) Numerical predictions of a PEM fuel cell performance enhancement by a rectangular cylinder installed transversely in the flow channel. *Apply Energy*. Vol. 86(10) ：2192-2203.

X. Li and I. Sabir. (2005) Review of bipolar plates in PEM fuel cells: Flow-field designs. *Int. J. Hydrogen Energy*, Vol. 30：359-371.

Wang X.D., Duan Y. Y. & Yan W. M.(2007) Novel serpentine-baffle flow field design for proton exchange membrane fuel cells. *J Power Sources*, Vol. 173：210–221.

Perng S.W., Wu H.W., Jue T.C.& Kuo-Chih Cheng. (2009) Numerical predictions of a PEM fuel cell performance enhancement by a rectangular cylinder installed transversely in the flow channel. *Appl Energ*, Vol. 86 (10) ：2192-2203.

Wang C. T., Chang P. C., Shaw C. K. & Cheng J. Y. (2009) Fuel cell bionic flow slab design. *J. Fuel Cell Sci. Technol*, Vol.6：1-5.

Wang C. T., Hu Y. C. & Zheng P. L. (2010) Novel Biometric Flow Slab Design for Improvement of PEMFC Performance, *Applied Energy*, Vol.87：1366-1375.

Wang C. T., Hu Y. C. and Hu T.Y. (2009) Biophysical Micromixer. *Sensors*, Vol. 9：5379-5389.

Wang, C. T., Yang, C. M., Chen, Z. S., Tseng S. (2011) Effect of Biometric Flow Channel on the Power Generation at Different Reynolds Numbers in the Single Chamber of Rumen Microbial Fuel Cells (RMFCs). *International Journal of Hydrogen Energy*, Vol.36 : 9242-9251

Watanabe, K. (2008) Recent developments in microbial fuel cell technologies for sustainable bioenergy, *J Biosci Bioeng*, Vol. 106：528-536.

Wong, S.H.; Bryant, P.; Ward, M.& Wharton, C. (2003) Investigation of mixing in a cross-shaped micro-mixer with static mixing elements for reaction kinetics studies sens. *Sens. Actuat. B - Chem.*, Vol. 95：.414-424.

Edited by Jucheng Yang and Norman Poh

In the recent years, a number of recognition and authentication systems based on biometric measurements have been proposed. Algorithms and sensors have been developed to acquire and process many different biometric traits. Moreover, the biometric technology is being used in novel ways, with potential commercial and practical implications to our daily activities. The key objective of the book is to provide a collection of comprehensive references on some recent theoretical development as well as novel applications in biometrics. The topics covered in this book reflect well both aspects of development. They include biometric sample quality, privacy preserving and cancellable biometrics, contactless biometrics, novel and unconventional biometrics, and the technical challenges in implementing the technology in portable devices. The book consists of 15 chapters. It is divided into four sections, namely, biometric applications on mobile platforms, cancelable biometrics, biometric encryption, and other applications. The book was reviewed by editors Dr. Jucheng Yang and Dr. Norman Poh. We deeply appreciate the efforts of our guest editors: Dr. Girija Chetty, Dr. Loris Nanni, Dr. Jianjiang Feng, Dr. Dongsun Park and Dr. Sook Yoon, as well as a number of anonymous reviewers.

IntechOpen