The Multi-Scenario Multi-Environment BioSecure Multimodal Database (BMDB)

Javier Ortega-Garcia, *Senior Member, IEEE,* Julian Fierrez, *Member, IEEE,* Fernando Alonso-Fernandez, *Member, IEEE,* Javier Galbally, *Student Member, IEEE,* Manuel R. Freire, *Student Member, IEEE,* Joaquin Gonzalez-Rodriguez, *Member, IEEE,* Carmen Garcia-Mateo, *Member, IEEE,* Jose-Luis Alba-Castro, *Member, IEEE,* Elisardo Gonzalez-Agulla, Enrique Otero-Muras, Sonia Garcia-Salicetti, *Member, IEEE,* Lorene Allano, Bao Ly-Van, Bernadette Dorizzi, Josef Kittler, Thirimachos Bourlai, Norman Poh, *Member, IEEE,* Farzin Deravi, *Member, IEEE,* Ming W. R. Ng, Michael Fairhurst, Jean Hennebert, *Member, IEEE,* Andreas Humm, *Member, IEEE,* Massimo Tistarelli, *Senior Member, IEEE,* Linda Brodo, Jonas Richiardi, *Member, IEEE,* Andrzej Drygajlo, *Member, IEEE,* Harald Ganster, Federico M. Sukno, Sri-Kaushik Pavani, Alejandro Frangi, *Senior Member, IEEE,* Lale Akarun, *Senior Member, IEEE,* Arman Savran

***Abstract*— A new multimodal biometric database designed and acquired within the framework of the European BioSecure Network of Excellence is presented. It comprises more than 600 individuals acquired simultaneously in three scenarios: i) over the Internet, ii) in an office environment with desktop PC, and iii) in indoor/outdoor environments with mobile portable hardware. The three scenarios include a common part of audio/video data. Also, signature and fingerprint data has been acquired both with desktop PC and mobile portable hardware. Additionally, hand and iris data was acquired in the second scenario using desktop PC. Acquisition has been conducted by 11 European institutions. Additional features of the BioSecure Multimodal Database (BMDB) are: two acquisition sessions, several sensors in certain modalities, balanced gender and age distributions, multimodal realistic scenarios with simple and quick tasks per modality, cross-European diversity, availability of demographic data and compatibility with other multimodal databases. The novel acquisition conditions of the BMDB allow to perform new challenging research and evaluation of either monomodal or multimodal biometric systems, as in the recent BioSecure Multimodal Evaluation campaign[1]. A description of this campaign including baseline results of individual modalities from the new database is also given. The database is expected to be available for research purposes through the BioSecure Association[2] during 2008.**

***Index Terms*— Multimodal, biometrics, database, evaluation, performance, benchmark, face, voice, speaker, signature, fingerprint, hand, iris.**

Javier Ortega-Garcia, Julian Fierrez, Fernando Alonso-Fernandez, Javier Galbally, Manuel R. Freire and Joaquin Gonzalez-Rodriguez are with ATVS/Biometric Recognition Group, Escuela Politecnica Superior, Univ. Autonoma de Madrid, Avda. Francisco Tomas y Valiente 11, Campus de Cantoblanco, 28049 Madrid, Spain (email: javier.ortega, julian.fierrez, fernando.alonso@uam.es). Carmen Garcia-Mateo, Jose-Luis Alba-Castro, Elisardo Gonzalez-Agulla and Enrique Otero-Muras are with Departamento de Teoria de la Senal y Comunicaciones, ETSI Telecomunicacion, Universidad de Vigo, Campus Universitario, 36310 Vigo, Spain. Sonia Garcia-Salicetti, Lorene Allano, Bao Ly-Van and Bernadette Dorizzi are with Institut TELECOM, TELECOM & Management SudParis, 9 rue Charles Fourier 91011 Evry, France. Josef Kittler, Thirimachos Bourlai and Norman Poh are with Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, GU2 7XH, United Kingdom. Farzin Deravi, Ming W. R. Ng and Michael Fairhurst are with Department of Electronics, University of Kent, Canterbury, CT2 7NT, United Kingdom. Jean Hennebert and Andreas Humm are with Universite de Fribourg, de Pérolles 90, 1700 Fribourg, Switzerland. Jean Hennebert is also with Institute of Business Information Systems, University of Applied Sciences Western Switzerland, HES-SO // Wallis, TechnoArk 3, 3960 Sierre, Switzerland. Massimo Tistarelli and Linda Brodo are with Computer Vision Laboratory, University of Sassari, Alghero, Italy. Jonas Richiardi and Andrzej Drygajlo are with Speech Processing and Biometrics Group, Laboratory of IDIAP, Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland. Harald Ganster is with Institute of Digital Image Processing, Joanneum Research Forschungsgesellschaft MBH, Wastiangasse 6, A-8010 Graz, Austria. Federico M. Sukno[a,b], Sri-Kaushik Pavani[b,a] and Alejandro Frangi[b,a,c] are with Center for Computational Imaging & Simulation Technologies in Biomedicine, [a] Networking Center on Biomedical Research - CIBER - BBN, Barcelona, Spain, [b] Information & Communication Technologies Department, Universitat Pompeu Fabra, Barcelona, Spain, [c] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. Lale Akarun and Arman Savran are respectively with Department of Computer Engineering and with Department of Electrical and Electronics Engineering Bogazici Universitesig, 34342 Bebek Istanbul, Turkey.

I. INTRODUCTION

BIOMETRICS is defined as the use of anatomical and/or behavioral traits of individuals for recognition purposes [1]. Advances in the field would not be possible without suitable biometric data in order to assess systems' performance through benchmarks and evaluations. In past years, it was common to evaluate biometric products on small custom or proprietary datasets [2] and therefore, experiments were not repeatable and a comparative assessment could not be accomplished. As biometric systems are being deployed, joint efforts have been conducted to perform common experimental protocols and technology benchmarks. Several evaluation procedures [3], databases and competitions have been established in recent years, e.g. the NIST Speaker Recognition Evaluations [4], the FERET and FRVT Face Recognition Evaluations [5], the series of Fingerprint Verification Competitions (FVC) [6], the Iris Challenge Evaluation (ICE) [7] or the Signature Verification Competition (SVC) [8]. Biometric data gathered for these competitions along with experimental protocols are made publicly available and in many cases, they have become a *de facto* standard, so that data and protocols utilized in these competitions are used by the biometric community to report subsequent research work.

However, evaluation of biometric systems has been done to date in a fragmented way, modality by modality, without any

[1] http://biometrics.it-sudparis.eu/BMEC2007
[2] http://biosecure.it-sudparis.eu/AB

Digital Object Identifier 10.1109/TPAMI.2009.76 0162-8828/$25.00 © 2009 IEEE

	Year	Ref.	Users	Sessions	Traits	2Fa	3Fa	Fp	Ha	Hw	Ir	Ks	Sg	Sp
BioSecure	2008	-	971 (DS1, Internet scenario)	2	2	X								X
			667 (DS2, Desktop scenario)	2	6	X		X	X		X		X	X
			713 (DS3, Mobile scenario)	2	4	X		X					X	X
BiosecurID	2007	[13]	400	4	8	X		X	X	X	X	X	X	X
BioSec	2007	[14]	250	4	4	X		X			X			X
MyIDEA	2005	[15]	104 (approx.)	3	6	X		X	X	X			X	X
BIOMET	2003	[16]	91	3	6	X	X	X	X				X	X
MBioID	2007	[17]	120 (approx.)	2	6	X	X	X			X		X	X
M3	2006	[18]	32	3	3	X		X						X
FRGC	2006	[19]	741	variable	2	X	X							
MCYT	2003	[20]	330	1	2			X					X	
BANCA	2003	[21]	208	12	2	X								X
Smartkom	2002	[22]	96	variable	4			X	X				X	X
XM2VTS	1999	[23]	295	4	2	X								X
M2VTS	1998	[24]	37	5	2	X								X
BT-DAVID	1999	[25]	124	5	2	X								X

TABLE I

MOST RELEVANT FEATURES OF EXISTING MULTIMODAL DATABASES. THE NOMENCLATURE IS AS FOLLOWS: 2FA STANDS FOR FACE 2D, 3FA FOR FACE 3D, FP FOR FINGERPRINT, HA FOR HAND, HW FOR HANDWRITING, IR FOR IRIS, KS FOR KEYSTROKES, SG FOR HANDWRITTEN SIGNATURE AND SP FOR SPEECH.

common framework. Some efforts in the past have been made regarding multimodal research and applications. Several multimodal databases are available today, typically as a result of collaborative European or national projects, but most of them include just a few modalities. Biometric database collection is a challenging task [2]. The desirability of large databases with presence of variability (multi-session, multi-sensor, multi-scenario etc.) makes the acquisition a time- and resource-consuming process, which in the case of multimodal databases, requires important additional efforts. A number of pre-acquisition and post-acquisition tasks are also needed: training of acquisition operators, recruitment of donors, supervision of acquired data, annotation, error correction, labeling, documentation, etc. Furthermore, legal and operational issues surround the donor's consent and storage and distribution process. Usually, a license agreement has to be signed between the distributor and the licensee, and the size of new databases (from gigabytes to terabytes) complicates their distribution.

This paper presents the recently acquired BioSecure Multimodal Database, which has been collected between November 2006 and June 2007. An important integrative effort has been done in the design and collection of this database, involving 11 European institutions of the BioSecure Network of Excellence [9]. The new database includes new features not present in existing databases. More than 600 individuals have been acquired simultaneously in three different scenarios (over the Internet, in an office environment with a desktop PC, and in indoor/outdoor environments using mobile devices) over two acquisition sessions and with different sensors for certain modalities. New challenging acquisition conditions that will allow the evaluation of realistic scenarios are included, such as Internet transactions using face and voice acquired with commercial webcams, or fingerprint and handwritten signature on modern mobile platforms. An example of such an evaluation is the recently conducted BioSecure Multimodal Evaluation Campaign (BMEC) [10], where several tasks were defined using data extracted from the BioSecure Multimodal Database. Being the largest publicly available multimodal database, it will allow novel research like: subpopulation effects, only attainable by having a large population; large scale multi-scenario/multi-sensor interoperability tests; etc. Furthermore, the width of the database gives us the ability to have very tight confidence intervals on results.

The contributions of the present paper can be summarized as: 1) overview of existing resources for research on multimodal biometrics, and 2) detailed description of the most comprehensive multimodal biometric database available to date, including aspects such as: relation with other existing databases, acquisition design, post-processing tasks, legal issues regarding distribution of biometric data, and a core experimental framework with baseline results that can be used as a reference when using the database. From these baseline results, some novel experimental findings can also be observed in topics of practical importance such as biometric recognition using mismatched devices, and multiple acquisition environments.

The rest of this paper is organized as follows. Section II summarizes previous work done in the domain of multimodal biometric database design and acquisition. The BioSecure Multimodal Database (BMDB) is described in Section III, and its compatibility with existing databases is given in Section IV. A brief description of the BioSecure Multimodal Evaluation Campaign, together with baseline results of individual modalities from the database, is done in Section V. Some research directions that can be followed with the new databases are pointed out in Section VI and finally, legal and distribution issues are discussed in Section VII.

INSTITUTION	COUNTRY	DS1	DS2	DS3
Universidad de Vigo	Spain	101 (△)	101	-
Bogazici University	Turkey	98	-	95
Ecole Polytechnique Federale de Lausanne	Switzerland	77	-	48
Groupe des Ecoles des Telecommunications	France	117	117	117 (△)
Joanneum Research Graz	Austria	70	70	-
University of Fribourg	Switzerland	70	-	70
University of Kent	UK	80	80	80
University of Surrey	UK	104	104	104
University of Sassari	Italy	55	55	-
Pompeu Fabra University	Spain	59	-	59
Universidad Politecnica de Madrid (*)	Spain	140	140 (△)	140
		971	**667**	**713**

TABLE II

INSTITUTIONS PARTICIPATING IN THE ACQUISITION, INCLUDING INVOLVEMENT IN THE THREE ACQUIRED DATASETS. FOR EACH DATASET, THERE WAS AN INSTITUTION - MARKED WITH (△) - IN CHARGE OF COORDINATING ITS ACQUISITION.

(*) THE OVERALL ACQUISITION PROCESS WAS COORDINATED BY UNIVERSIDAD POLITECNICA DE MADRID (UPM) THROUGH THE ATVS/BIOMETRIC RECOGNITION GROUP. THE ATVS GROUP, FORMERLY AT UNIVERSIDAD POLITECNICA DE MADRID (UPM), IS CURRENTLY AT UNIVERSIDAD AUTONOMA DE MADRID (UAM).

THE GROUPE DES ECOLES DES TELECOMMUNICATIONS (GET) HAS CHANGED ITS NAME TO INSTITUT TELECOM.

II. RELATED WORKS

Multimodal biometric research was first supported by "chimeric" multimodal databases containing synthetic subjects [11]. These synthetic subjects were constructed by combining biometric data from different databases (e.g. fingerprint images from one subject in one database, face images from another subject in another database, etc.) to produce data sets which not represent real multimodal samples. As multibiometric data may be necessarily correlated [12], the use of chimeric databases should be avoided. But collecting multimodal databases has important issues: more acquisition time is generally required, subjects may react negatively to the longer acquisition sessions needed, size of the database and acquisition cost are considerably higher, and in general, the management of such a task is exponentially more complex. Fortunately, in recent years efforts have been directed to collecting real multimodal databases involving various biometric traits and hundreds of users.

First efforts were focused on the acquisition of monomodal or bimodal databases (one or two biometric traits sensed), e.g.: the **MCYT** database [20], including signatures and fingerprint images of 330 subjects; the **M2TVS** [24], **XM2VTS** [23], and **BANCA** [21] databases, with face and voice data of 37, 295, and 208 subjects, respectively; or the **FRGC** database [19], which includes 2D and 3D face images of 741 subjects; and the **BT-DAVID** database [25], with audio-visual data from 124 individuals.

There are also several multimodal biometric databases with multiple traits available today, or in the process of completion, mainly as a result of collaborative national or international projects. Some examples include: the **SmartKom** [22], **M3** [18] and **MBioID** [17] databases, and the following ones (which we present in some more detail because of their relation to the BioSecure multimodal database):

- **BiosecurID** database [13]. The BiosecurID database was collected in 6 Spanish institutions in the framework of the BiosecurID project funded by the Spanish Ministry of Education and Science. It has been collected in an office-like uncontrolled environment (in order to simulate a realistic scenario), and was designed to comply with the following characteristics: 400 subjects, 8 different traits (speech, iris, face still and talking face, signature, handwriting, fingerprint, hand and keystroking) and 4 acquisition sessions distributed in a 4 month time span.
- **BioSec** database [14]. BioSec was an Integrated Project (IP) of the 6th European Framework Programme [26] which involved over 20 partners from 9 European countries. One of the activities within BioSec was the acquisition of a multimodal database. This database was acquired at four different European sites and includes face, speech (both with a webcam and a headset), fingerprint (with three different sensors) and iris recordings. The baseline corpus [14] comprises 200 subjects with 2 acquisition sessions per subject. The extended version of the BioSec database comprises 250 subjects with 4 sessions per subject (about 1 month between sessions). A subset of this database was used in the last International Fingerprint Verification Competition [6] held in 2006.
- **MyIDEA** database [15], which includes face, audio, fingerprint, signature, handwriting and hand geometry of 104 subjects. Synchronized face-voice and handwriting-voice were also acquired. Sensors of different quality and various scenarios with different levels of control were considered in the acquisition.
- **BIOMET** database [16], which offers 5 different modalities: audio, face images (2D and 3D), hand images, fingerprint (with an optical and a capacitive sensor) and signature. The database consists of three different acquisition sessions (with 8 months between the first and the third) and comprises 91 subjects who completed the three sessions.

In Table I the most relevant features of the existing multimodal databases are summarized. The current paper presents the recently acquired BioSecure Multimodal Database. It is designed to comply with several characteristics that, as can be observed in Table I, make it unique, namely: hundreds of users and several biometric modalities acquired under several scenarios.

III. THE BIOSECURE MULTIMODAL DATABASE (BMDB)

A. General description

The acquisition of the BioSecure Multimodal Database (BMDB) was jointly conducted by 11 European institutions participating in the BioSecure Network of Excellence [9], see Table II. The institution in charge of coordinating the acquisition process was Universidad Politecnica de Madrid (UPM), from Spain, through the ATVS/Biometric Recognition Group (marked with (*) in Table II). BMDB is comprised of three different datasets, with an institution in charge of coordinating the acquisition of each dataset (marked with (△) in Table II). The three datasets are:

- **Data Set 1 (DS1)**, acquired over the **Internet** under unsupervised conditions (i.e. connecting to an URL and following the instructions provided on the screen).

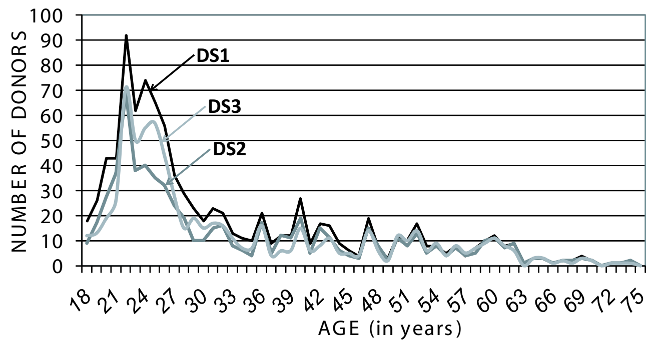

Fig. 1. Age distribution of the three datasets of the BioSecure Multimodal Database.

	DS1	DS2	DS3
Age distribution (18-25/25-35/35-50/>50)	43% / 25% / 19% / 23%	41% / 21% / 21% / 17%	42% / 25% / 17% / 16%
Gender distribution (male / female)	58% / 42%	56% / 44%	56% / 44%
Handedness (righthanded / lefthanded)	94% / 6%	94% / 6%	94% / 6%
Manual workers (yes / no)	2% / 98%	3% / 97%	2% / 98%
Vision aids (glasses, contact lenses / none)	42% / 58%	42% / 58%	43% / 57%

TABLE III

STATISTICS OF THE THREE DATASETS OF THE BIOSECURE MULTIMODAL DATABASE.

- **Data Set 2 (DS2)**, acquired in an office environment (**desktop**) using a standard PC and a number of commercial sensors under the guidance of a human acquisition supervisor.
- **Data Set 3 (DS3)**, acquired using **mobile** portable hardware under two acquisition conditions: indoor and outdoor. Indoor acquisitions were done in a quiet room, whereas outdoor acquisitions were recorded in noisy environments (office corridors, the street, etc.), allowing the donor to move and to change his/her position.

The BMDB has been designed to be representative of the population that would make possible use of biometric systems. As a result of the acquisition process, about 40% of the subjects are between 18 and 25 years old, 20-25% are between 25 and 35 years old, 20% of the subjects are between 35 and 50 years old, and the remaining 15-20% are above 50 years old, see Table III. In addition, the gender distribution was designed to be as balanced as possible, with no more than 10% difference between male and female sets. Metadata associated with each subject was also collected to allow experiments regarding specific groups. The available information includes: age, gender, handedness, manual worker (yes/no), visual aids (glasses or contact lenses/none) and English proficiency. In Table III and Figure 1, the actual statistics of the three datasets of the BioSecure Multimodal Database are shown. Other features are as follows:

- Two acquisition sessions in order to consider time variability (separated between 1 and 3 months).

Modality	Samples	# Samples
Face still	2 still frontal face images	2 image files
AV	4 digits PIN code, 2 repetitions* from a set of 100 different codes, in English	2 video files
AV	4 digits PIN code, 2 repetitions* from a set of 10 different codes, in national language	2 video files
AV	Digits from 0 to 9 in English	1 video file
AV	2 different phonetically rich sentences** in English	2 video files
AV	2 different phonetically rich sentences** in national language	2 video files
		11

* The same PIN code between datasets and sessions.
** Different sentences between datasets and sessions.

TABLE IV

COMMON CONTENTS OF AUDIO AND VIDEO IN THE THREE DATASETS.

- Multimodal data representative of realistic scenarios, with simple and quick tasks per modality (e.g. utterances of PINs or short sentences, still face images, handwritten signatures, hand images using a camera, etc.) instead of long tasks that are not appropriate in a real multimodal scenario (e.g. handwriting, hand images using a document scanner, etc.).
- Several biometric samples per session.
- Homogeneous number of samples and sessions per subject.
- Cross-European diversity (language, face, etc.) and site variability, achieved through the acquisition across different European institutions.
- Acquisition of the same trait with different biometric sensors in order to consider sensor variability.

The three datasets of BMDB include a common part of audio and video data, described in Table IV, which comprises still face images and talking face videos. Also, signature and fingerprint data has been acquired both in DS2 and DS3. Additionally, hand and iris data was acquired in DS2. All this information is summarized in Table V, which is further detailed in the subsections dedicated to each dataset.

B. Data Set 1 (DS1) - Internet

The purpose of DS1 is to acquire material over the Internet without human supervision. For DS1, the acquisition protocol consisted of the common part described in Table IV. Therefore, the modalities included in DS1 are: voice and face. The first session was acquired using a PC provided by the acquisition center, with some guidance to let donors to become familiar with the acquisition and to correctly perform it, whereas the second session was allowed to take place over an uncontrolled scenario (at donor's home or office using appropriate hardware). In most of the cases both sessions were acquired with the provided PC. Note that the speech information acquired enables both text-dependent (PINs) and text-independent speaker recognition experiments [27]. Digits are the same between sessions, enabling text-dependent speaker recognition based on digits. On the other hand, speakers utter also various sentences, being the sentences different in each session (see Table IV), enabling also text-independent speaker recognition.

The acquisition of DS1 was performed by connecting to an URL using an Internet browser and following the instructions provided on the screen. The acquisition was performed using a

General features of the database

	DATA SET 1 (DS1)	DATA SET 2 (DS2)	DATA SET 3 (DS3)
Subjects	971	667	713
Sessions	2	2	2
Supervisor	No	Yes	Yes
Conditions	Over the **Internet**	Standard office (**desktop**)	**Mobile** indoor and outdoor
Hardware	PC, webcam, microphone	PC, commercial sensors	Portable devices

Biometric data for each user and for each session in the database

MODALITY	DATA SET 1 (DS1)	DATA SET 2 (DS2)	DATA SET 3 (DS3)	# SAMPLES
Common AV - indoor	**11 samples**	**11 samples**	**11 samples**	**33**
- Audio-video	4 PIN	4 PIN	4 PIN	
	4 sentences	4 sentences	4 sentences	
	1 digits sequence	1 digits sequence	1 digits sequence	
- Face still (webcam)	2 frontal face images	2 frontal face images	2 frontal face images	
Common AV - outdoor	-	-	**11 samples**	**11**
- Audio-video			4 PIN	
			4 sentences	
			1 digits sequence	
- Face still (webcam)			2 frontal face images	
Signature	-	**25 samples**	**25 samples**	**50**
		15 genuine	15 genuine	
		10 imitations	10 imitations	
Fingerprint - thermal	-	**12 samples** $(3 \times 2 \times 2)$	**12 samples** $(3 \times 2 \times 2)$	**24**
Fingerprint - optical	-	**12 samples** $(3 \times 2 \times 2)$	-	**12**
Iris	-	**4 samples** (2×2)	-	**4**
Hand - digital camera	-	**8 samples** (2×4)	-	**8**
Face still - digital camera	-	**4 samples** (2 + 2)	-	**4**
# SAMPLES	**11**	**76**	**59**	**146**

TABLE V
GENERAL FEATURES OF THE DATABASE AND MODALITIES ACQUIRED IN EACH DATASET.

(a) (b) (c)

Fig. 2. Graphical user interface of the acquisition application for DS1.

standard webcam with microphone. In order to achieve realistic conditions, the use of no specific webcam was imposed. The appearance of the graphical user interface of the application for the acquisition is shown in Figure 2. Figure 2(a) represents the user interface prepared for the acquisition of an audiovisual sample. The left panel shows the instructions to the donor while the right panel shows the webcam stream. Figure 2(b) shows the graphical user interface just after an audiovisual sample has been acquired. The sample is presented to the donor in order to be validated before it is sent to the server. In Figure 2(c), a frontal still image has just been acquired. The donor has to adjust the position of his face to fit the overlaid "virtual glasses and chin" mask in order to normalize the pose.

C. Data Set 2 (DS2) - Desktop

The scenario considered for the acquisition of DS2 was a standard office environment. The acquisition was carried out using a desktop PC and a number of sensors connected to the

Modality	**Sensor**	**Samples**	**# Samples**
Signature	Tablet	5 genuine signatures of donor *n*	5 text files
		5 dynamic imitations of donor *n*-1 (*n*-3 in session 2)	5 text files
		5 genuine signatures of donor *n*	5 text files
		5 dynamic imitations of donor *n*-2 (*n*-4 in session 2)	5 text files
		5 genuine signatures of donor *n*	5 text files
Face Still	Webcam	Common still frontal face images (see Table IV)	2 image files
AV	Webcam+headset	Common audio-video part (see Table IV)	9 video files
Iris image	Iris camera	(Right eye, Left eye) $\times$ 2	4 image files
Fingerprint	Optical	((Thumb $\rightarrow$ index $\rightarrow$ middle) $\times$ 2 hands {right $\rightarrow$ left}) $\times$ 2	12 image files
	Thermal	((Thumb $\rightarrow$ index $\rightarrow$ middle) $\times$ 2 hands {right $\rightarrow$ left}) $\times$ 2	12 image files
Hand	Digital camera	(Right hand $\rightarrow$ Left hand) $\times$ 2 without flash	4 image files
		(Right hand $\rightarrow$ Left hand) $\times$ 2 with flash	4 image files
Face Still	Digital camera	2 photos without flash $\rightarrow$ 2 photos with flash	4 image files
			76 files

TABLE VI
ACQUISITION PROTOCOL OF DS2.

PC via the USB or Bluetooth interface. Acquisition is managed by a supervisor, who is in charge of the following activities: i) training of the contributor in case of difficulties in using a particular sensor; ii) validation of every acquired biometric sample, allowing a sample to be acquired again if it has not been properly acquired (e.g. wrong finger in the case of fingerprint acquisition or wrong utterance in the case of speech acquisition); and iii) guidance of the acquisition process by remembering the steps of the acquisition protocol and pointing out the sensor involved.

The modalities included in DS2 are: voice, face, signature, fingerprint, hand and iris. Hardware used for the acquisition included a Windows-based PC with a USB hub, and the biometric sensors specified in Figure 3. An example of the setup used by UPM is shown in Figure 4, and the data acquired in DS2 is described in Table VI.

A specific application was developed for the acquisition of DS2, aiming to provide a common and homogeneous working interface for all the sites participating in the acquisition. This software tool allowed the recapture of samples until they exhibited satisfactory quality, the correction of invalid or missing samples, and the inclusion of new users or sessions at any point of the acquisition process. In Figure 5, the screen captures of the main module and the different modules of the DS2 acquisition software are shown.

D. Data Set 3 (DS3) - Mobile

The aim of DS3 is to capture several modalities on mobile platforms. The modalities acquired in DS3 are: face, voice, fingerprint and signature. For the common audio-video recordings, each session is comprised of 2 acquisition conditions, indoor and outdoor, performed during the same day. Hardware devices used for the acquisition include the biometric sensors specified in Figure 6, and the data acquired in DS3 is described in Table VII. A supervisor was in charge of managing each acquisition session, providing appropriate training to the contributors in case of difficulties with a sensor, validating every acquired biometric sample, and allowing reacquisition in case of a sample has not been properly acquired.

The two acquisition conditions considered for audio-video recordings are intended to comprise different sources of variability. "Indoor" acquisitions were done in a quiet room, just changing the position between each audio-video sequence. "Outdoor" acquisitions were recorded in noisy environments such as building corridors, the street, etc., allowing the donor to move and to change position during and between each audio-video sequence. For signature and fingerprint, only indoor data was acquired, which can be considered as degraded with respect to DS2, as signatures and fingerprints were acquired while standing and holding the PDA.

The acquisition with Samsung Q1 was carried out using the same application used for DS2, adapted and limited to the audio-video acquisition task. For the acquisition of data with PDA, a specific software tool was developed. In Figure 7, screen captures of the main module and the different modules of the DS3 acquisition software for PDA are shown.

E. Acquisition guidelines and validation of acquired data

A data validation step was carried out to detect invalid samples within the three datasets. We should distinguish between valid low quality samples and invalid samples. Low-quality samples are acceptable, as long as the acquisition protocol is strictly followed, and the different biometric sensors are correctly used (e.g. blurred iris images, dry or wet fingerprint images, etc.). These samples were not removed from the database since they represent real-world samples that can be found in the normal use of a biometric system. On the other hand, invalid samples are those that do not comply with the specifications given for the database (e.g. wrong PIN, forgery of a wrong signature, donor's head or hand out of frame, etc.).

The first stage of the validation process was carried out during the acquisition itself. A human supervisor was in charge of validating every acquired biometric sample, being recaptured if it did not meet the specified quality standards. After completion of the acquisition, a second validation step of the three datasets was

Modality	Sensor	Condition	Samples	# Samples
Signature	PDA	Indoor	5 genuine signatures of donor n	5
			5 dynamic forgeries of donor n-1 (n-3 in session 2)	5
			5 genuine signatures of donor n	5
			5 dynamic forgeries of donor n-2 (n-4 in session 2)	5
			5 genuine signatures of donor n	5
Fingerprint	PDA (thermal)	Indoor	((Thumb $\rightarrow$ index $\rightarrow$ middle) $\times$ 2 hands $\{$right $\rightarrow$ left$\}$) $\times$ 2	12
Face Still	Q1+	Indoor	Common still frontal face images (see Table IV)	2
AV	webcam	Indoor	Common audio-video part (see Table IV)	9
Face Still	Q1+	Outdoor	Common still frontal face images (see Table IV)	2
AV	webcam	Outdoor	Common audio-video part (see Table IV)	9
				59

TABLE VII
ACQUISITION PROTOCOL OF DS3.

Fig. 3. Hardware devices used in the acquisition of DS2 together with acquisition samples.

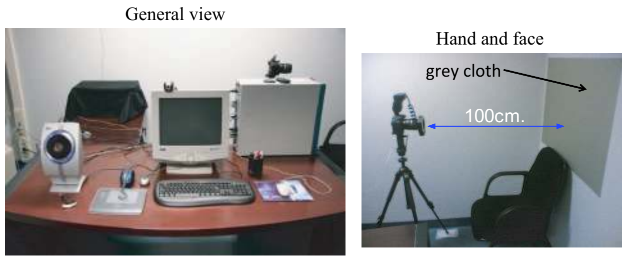

Fig. 4. Example of the setup used by UPM for the acquisition of DS2.

carried out again manually by a human expert. The following validation criteria and acquisition guidelines were given to the supervisors in charge of validation of the data:

- Face acquisition (both face still and video): donor pose should be frontal (looking straight into the camera) and with neutral expression, and donor's head should not be out of frame. In video files, audio and video fields should be synchronized. Blurred images are not discarded, unless the face is clearly non-visible.
- Iris acquisition: glasses (if any) should be removed before acquisition, but contact lenses are acceptable. A part of donor's eye falling out of frame or eye closed are not allowed.
- Hand acquisition: the hand pose is with wide open fingers. Fingers too close together or part of the hand out of frame are not allowed.
- Fingerprint acquisition with the optical sensor: the contact surface of the device should be cleaned after each donor session. For fingerprint acquisition with the thermal sensor, as it is difficult to be used correctly, the donor was allowed to try multiple times before the first acquisition. Very low quality fingerprints or very small size images due to improper use of the sensor are not allowed.
- Signature acquisition: donors were asked to sign naturally (i.e. without breaks or slowdowns). For impostor realizations, signature to be imitated could be replayed on the screen with the dynamic process and the donor was allowed to train before forging.

F. Problems encountered during the acquisition

A list of problems encountered has been gathered during the acquisition of the database thanks to the feedback provided by the donors and/or the different acquisition sites. Concerning to the usability, the most relevant are:

- The Yubee thermal fingerprint sensor was difficult to use, requiring many trials to get a reasonably good fingerprint capture. This sensor also caused annoyance in some users due to Failure to Enroll error.
- The auto-focusing function of the iris camera was not always working, needing several trials to get a reasonably sharp iris image capture. Focus on iris scanner sometimes did not always correspond to the acquisition instant.
- Lightning had influence on the quality of acquired iris images in some cases, requiring to reduce the overall illumination of the room.

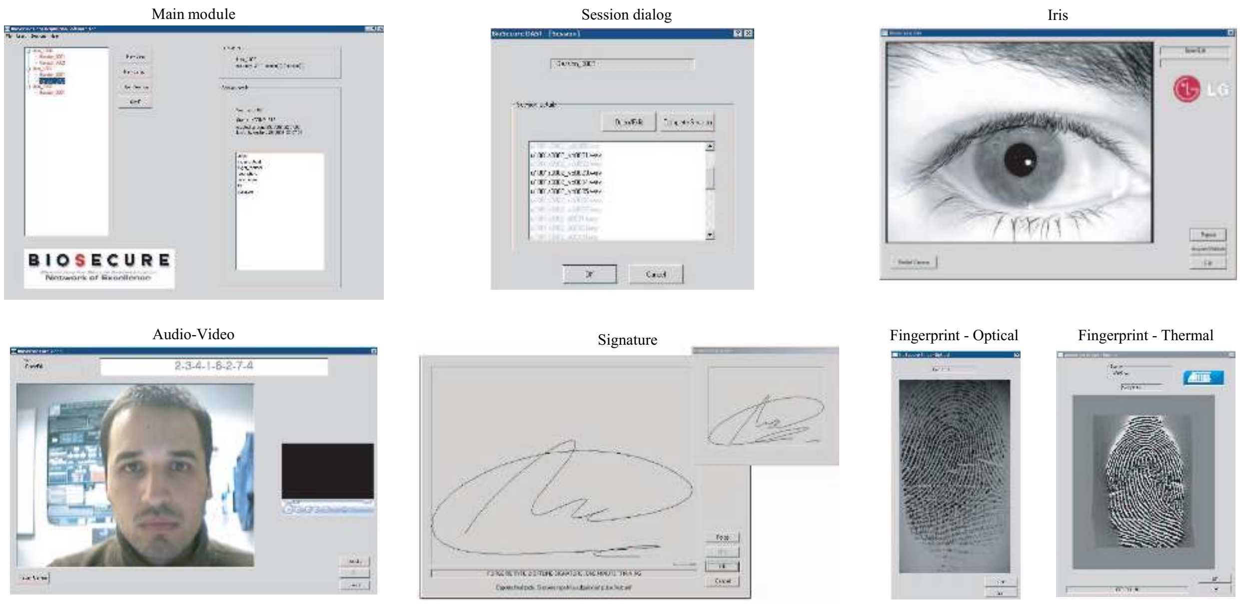

Fig. 5. Screen captures of the main module (top left) and the different modules of the DS2 acquisition software.

- For some volunteers doing the face pictures, there were too many reflections on their glasses.
- During swiping the finger on the fingerprint sensor of the PDA, there was a risk of touching the screen or the button just below the sensor. This could end the capture mode of the software, requiring a new acquisition.

IV. Compatibility of the BioSecure Multimodal Database with other databases

Several BioSecure partners have considerable experience in biometric database collection (e.g. MCYT [20], XM2VTS [23], BIOMET [16], MyIDEA [15], MBioID [17], BioSec [14] or BT-DAVID [25]). In recent years, various multimodal databases have been collected, mainly in academic environments. As a result, permanent staff members and some students have participated in the biometric data acquisition for these databases over several years. The existence of this data enables the number of sessions for a subset of such donors to be increased, therefore allowing studies of long term variability. Specifically, the BioSecure MDB shares individuals with:

- The **BioSec** database, which comprises 250 subjects. The two databases have in common the following traits: optical/thermal fingerprint, face, speech and iris. Both databases have in common 25 subjects, separated by about 2 years.
- The **BiosecurID** database, which comprises 400 subjects. The two databases have in common the following traits: optical/thermal fingerprint, face, hand, signature, speech and iris. Both databases have in common 31 subjects, separated by about 1 year.

It should also be noted that the devices and protocol of some of the traits acquired in the BioSecure MDB are compatible with other existing databases, so it can be combined with portions of them to increase the number of available subjects, specifically: the above mentioned BioSec and BiosecurID, the MCYT database (signature) and the MyIDEA database (signature and fingerprint).

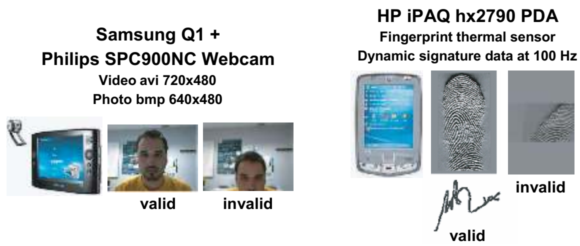

Fig. 6. Hardware devices used in the acquisition of DS3 together with acquisition samples.

Fig. 7. Screen captures of the main module (left) and the different modules of the DS3 acquisition software for PDA.

V. BioSecure Multimodal Evaluation Campaign

The BioSecure Multimodal Evaluation Campaign (BMEC) [10] has been conducted during 2007 as a continuation of the BioSecure Multimodal Database acquisition. At the time the evaluation was done, the database was not fully completed. Therefore, the BMEC was done on a selected subset of the database in order to show its consistency and value. Two different scenarios were defined for the Evaluation: an Access Control scenario on DS2 data and a degraded Mobile scenario on data from DS3. For these scenarios, different multimodal experiments were carried out using matching scores generated by several reference monomodal systems [28], in order to compare different fusion schemes both by BioSecure partners and by external researchers who responded to a public call for participation. In addition, monomodal evaluations on four modalities from DS3 were also performed. Since it is not within the purposes of this paper to describe in detail the BioSecure Multimodal Evaluation Campaign or their results, we will give only a brief outline of the experimental protocols used in these evaluations, together with some selected results, so the reader can have a reference experimental framework as well as baseline results.

A. Access Control scenario evaluation

The aim of the *DS2 Evaluation (Access Control scenario evaluation)* was to compare multimodal biometric fusion algorithms assuming that the environment is well controlled and the users are supervised, using data from BMDB DS2 (desktop dataset). Three biometric traits were employed: still face images, fingerprint and iris. Together with the scores computed using the reference systems, a set of quality measures were also extracted from the biometric samples. A LDA-based face verifier using Fisher Linear Discriminant projection [29] was used as the face reference system, with correlation as similarity measure (between two projected features). Different face quality measures were computed using the Omniperception proprietary Affinity SDK[3], including measures like contrast, brightness, etc. For the fingerprint modality, the NIST minutiae-based fingerprint system was used [30], whereas the quality measure was based on averaging local gradients [31]. Minutiae are detected using binarization and thinning operations, and for each minutiae, its position, direction, type and quality is computed. Matching is done by looking correspondences between two pairs of minutia, one pair from the template fingerprint and one pair from the test fingerprint. The fingerprint matcher is rotation and translation invariant. As iris reference system, a variant of the Libor Masek's system was used [32] and three different iris quality measures were computed: texture richness, difference between iris and pupil diameters, and proportion of iris used for matching. Masek's system employs the circular Hough transform for iris boundaries detection, and the Daugmans rubber sheet [33] model for normalization of the iris region. Feature encoding is implemented by using 1D Log-Gabor wavelets and phase quantization to four levels using the Daugman method [33]. For matching, the Hamming distance is chosen as a metric for recognition.

For this scenario, data from 333 individuals of DS2 has been used. Two sets of scores are computed, one for development and another for testing. There are 51 individuals in the development set (provided to the evaluation participants to tune their algorithms), 156 individuals in the test set (sequestered and only used by the evaluation organizers on the already tuned algorithms), and 126 individuals set as an external population of zero-effort impostors. For each person, four samples per modality are available (two per session, see Table VI). The first sample of the first session is used as template. Remaining samples are considered as query data (the other from Session 1 and the two from Session 2). The evaluated fusion algorithms were tested using only the query samples of Session 2. This testing was done using as impostors the external population of 126 individuals set as zero-effort impostors. This external population was not used elsewhere. In this way, a fusion algorithm will not have already "seen" the impostors during its training stage (for which data from Session 1 can be used), avoiding optimistic performance bias.

[3]http://www.omniperception.com

Two types of evaluations were proposed in this scenario:

- *Quality-based evaluation*, aimed to test the capability of fusion algorithms to cope with template and query biometric signals acquired with different devices, by exploiting quality information in the information fusion process [34]. Face and fingerprint modalities are considered in this case. Face images collected with the webcam (referred to as low quality data) and the digital camera (high quality data) are denoted as *fa1* and *fnf1* streams, respectively. Fingerprint data include images from the two sensors used in DS2. They are denoted as *fo* (optical) and *ft* (thermal). The case where templates and query images are acquired with different devices is also considered. For the face modality, this is denoted as *xfa1* stream, i.e. the templates acquired with *fnf1* (digital camera) and the queries with *fa1* (webcam). Similarly, for the fingerprint modality, the cross-device stream is denoted as *xft*, i.e. the templates acquired with *fo* (optical) and the queries with *ft* (thermal). It should be noted that the purpose of these cross-device experiments were just to evaluate the effect of matching biometric signals coming from different devices, without any special adjustment of the pre-processing or matching steps to deal with this issue. Therefore in this case, the order of the samples (template, query) has not relevant impact.
- *Cost-based evaluation*, aimed to achieve the best performance with a minimal cost of acquiring and processing biometric information. The use of each biometric trait is associated with a given cost. Only face images from the webcam (*fa1*) and fingerprint images from the thermal sensor (*ft*) are used, together with iris images (denoted as *ir1*). No cross-device experiments are conducted in this case. In the evaluation, for each modality used in the fusion, 1 cost unit is charged.

B. Mobile scenario evaluation

The *DS3 Evaluation (Mobile scenario evaluation)* was aimed to compare biometric recognition algorithms assuming that the data is acquired using mobile devices and the users are not supervised, using data from BMDB DS3 (mobile dataset). For multimodal experiments, 2D face video, fingerprint and signature data was used. Monomodal experiments on 2D face video, fingerprint, signature and talking face data were also carried out. 2D face video scores for the multimodal evaluation were generated using an eigenface-based approach developed by Bogazici University [35]. It uses the standard eigenface approach to represent faces in a lower dimensional subspace. All the images used by the system are firstly normalized. The face space is built using a separate training set and the dimensionality of the reduced space is selected such as the 99 per cent of the variance is explained by the Principal Component Analysis. After that, all the target and test images are projected onto the face space. Then, the L1 norm is used to measure the distance between the projected vectors of the test and target images. For the fingerprint modality, the NIST fingerprint system was used [30]. The signature reference system was developed by GET-INT (currently TELECOM & Management SudParis) and is based on Hidden Markov Models

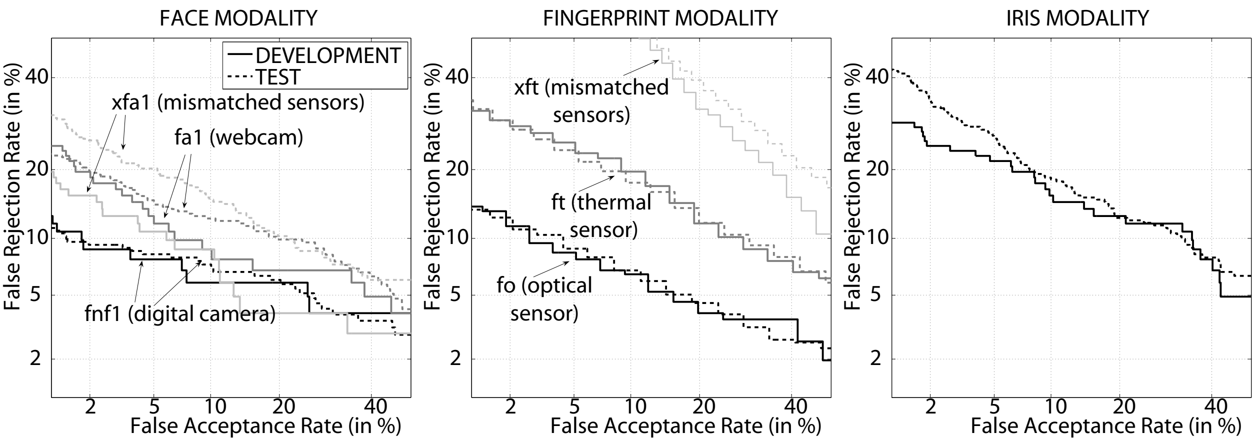

Fig. 8. **Access control scenario evaluation (BMDB DS2, desktop dataset)**. Baseline results of the face, fingerprint and iris modalities.

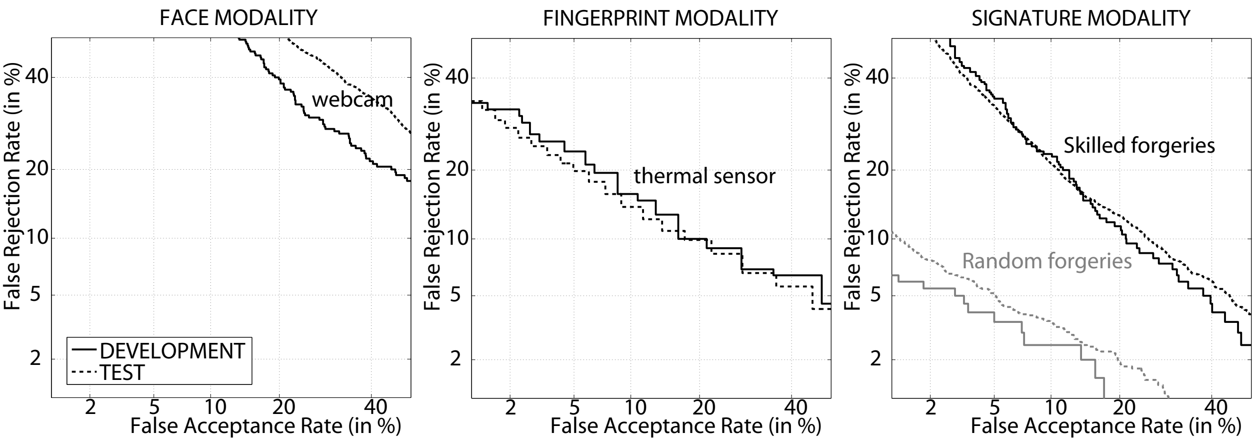

Fig. 9. **Mobile scenario evaluation (BMDB DS3, mobile dataset)**. Baseline results of the face, fingerprint and signature modalities.

[36], [37], [38]. Twenty five dynamic features are extracted at each point of the signature. Signatures are modeled by a continuous left-to-right HMM [39], by using in each state a continuous multivariate Gaussian mixture density. The number of states in the HMM modeling the signatures of a given person is determined individually according to the total number of all the sampled points available when summing all the genuine signatures that are used to train the corresponding HMM. Matching is done using the Viterbi algorithm [39].

For this scenario, 480 individuals from DS3 were considered. A set of 50 individuals was used for development, whereas the remaining 430 were used for testing. Two different experiments were carried out in the multimodal evaluation of this scenario, one using random signature forgeries and the other using skilled signature forgeries. For the 2D face modality, two indoor video samples of the first session were used as templates, whereas two outdoor video samples of the second session were used as query data. For the fingerprint modality, two samples of the first session were used as template data and two samples of the second session as query data. Signature templates were generated using five genuine signatures of the first session. As query data, genuine signatures of the second session and forgeries acquired during both sessions are used.

C. Baseline results

We plot in Figures 8 and 9 the verification results using the scores generated for the DS2 and DS3 multimodal evaluations. For the DS2 evaluation, results are shown using only query samples of the Session 2, as testing of the algorithms is done only using query samples of this session.

By looking at Figure 8, it is observed that the performance of the face modality is degraded when using the webcam, both if we match two images from the webcam and if we mismatch one webcam and one digital camera image. This is not true for the fingerprint modality, where a significant degradation is observed in the *xft* stream. As revealed in previous studies [40], [41], matching fingerprint images coming from different sensors has severe impact on the recognition rates due to variations introduced by the sensors (e.g. see fingerprint images of Figures 3 or 5).

Regarding the results on DS3 data shown in Figure 9, it is remarkable the degradation of performance of the face modality with respect to the *fa1* stream of DS2, in which the same webcam is used for the acquisition. The more challenging environment of DS3, including outdoor acquisitions in noisy ambience, has clear impact on the performance of the face modality (e.g. see Figure 10). The face reference system is based on the PCA approach which is not adequate to cope with the huge illumination variability. Worth noting, this is not observed in the fingerprint modality, where no degradation is observed with respect to the *ft* fingerprint stream of DS2. Being the two sensors based on the same (thermal) technology, quality of fingerprint images is not affected by the differences in the acquisition conditions between DS2 and DS3.

The demographic statistics of the development and test sets used in the Mobile scenario evaluation are quite similar in terms of gender, age and handedness of the donors. The main difference is found in the visual aids. As there are more people wearing glasses in the test database, we can suppose that the 2D face test database is "more difficult" than the corresponding development database. This is mirrored in the results of Figure 9. For the other modalities (fingerprint and signature), performance results are observed to be very close on both sets.

VI. Research and Development Using the BioSecure Multimodal Database

Some of the research and development directions that can be followed using the database have been already put forward through this paper. In this section, we summarize these directions. It has to be emphasized that the BioSecure Multimodal Database includes new challenging scenarios not considered in existing biometric databases. This new database is unique in that it includes hundred of users acquired simultaneously under different environmental conditions and using multiple sensors for the same modality. It allows uses such as:

- Novel research in the available modalities or in multibiometric systems, with very tight confidence intervals on results thanks to the size of the database.
- Evaluation of multibiometric algorithms using a common framework and real multimodal data. We should remark the high error rates of the baseline results presented in Section V-C (between 5 and 10% EER for the best cases), leaving a lot of room for improvement, either with individual modalities or with multibiometric approaches.
- Evaluation of sensor interoperability on a large scale in traits acquired with several devices (face, fingerprint, speech and signature) [42], [43] thanks to the amount of available data, as done in the BioSecure Multimodal Evaluation Campaign (BMEC) [10].
- Study of differences in system performance due to environmental variations: over the Internet, in an office environment with a desktop PC, or with mobile platforms in an indoor/outdoor environment.
- Evaluation of several multibiometric realistic scenarios, e.g. bimodal passport using face and fingerprint, Internet-based access using face and voice, mobile-based access with face and fingerprint, etc.
- Study of time variability in biometric systems and template update techniques [44]. Research can be done on the short term (considering data from the same session), on the medium term (considering data from the different sessions), or on the long term (considering common data from other databases, as mentioned in Section IV).
- Evaluation of potential attacks to monomodal or multimodal systems [45].
- Effect of the different acquisition scenarios/devices on the quality of acquired samples and its impact on the recognition performance [46], [47].
- Biometric studies depending on demographic information such as age [48], gender [49], handedness, state of the hand (manual workers), or visual aids.
- Cross-European diversity and site variability studies in terms of language (speech), appearance (face), etc.

Fig. 10. Variability between face samples acquired with different sensors and in different environments. Face images are plotted for i) indoor digital camera (from DS2, left, resized at 20%), ii) indoor webcam (from DS2, medium), and iii) outdoor webcam (from DS3, right).

VII. Legal issues and distribution

The Directive 95/46/EC of the European Parliament and the Council of 24 October 1995 sets the European requirements on the protection of individuals with regard to the processing of personal data and on the free movement of such data. According to this Directive, biometric data is considered as "personal data". Based on this regulation, donors were asked to read and sign a consent agreement before starting the acquisition process, which included comprehensive information about the motivation and planed used of the biometric data, with pointers to the security measures applied to protect these data.

Other personal data were acquired and stored securely and independently of the biometric data, including: name, contact details, age, gender, handedness, manual worker, vision aids, and English proficiency. These non-biometric data are managed by the BioSecure partner involved in the acquisition of the donor at hand (also referred to as "controller" in the Directive 95/46/EC). In any subsequent use or transfer of data, only the raw biometric data plus the fields {age, gender, handedness, visual aids, manual worker and English proficiency} are considered, without any link to identities of the donors (i.e. name and contact details).

The BioSecure Multimodal Database will be distributed during 2008. The distribution will be managed by the recently created BioSecure Association[4]. This Association is concerned with the use and dissemination of the results generated within the BioSecure Network of Excellence involving Intellectual Property Rights.

VIII. Conclusion

The existence of evaluation procedures and databases is crucial for the development of biometric recognition systems. It is often the existence of datasets with new challenging conditions which drives research forward. Public biometric databases allow the creation of common and repeatable benchmarks and algorithms, so that new developments can be compared with existing ones. However, biometric database collection is a time- and resource-consuming process, specially in the case of multimodal databases. As a result, most of the existing biometric databases typically include only one or two modalities. Fortunately, in recent years important efforts have been directed to collecting real multimodal databases, mainly in the framework of collaborative national or international projects, resulting in a number of databases available or in the process of completion.

[4]More information to be found in http://biosecure.it-sudparis.eu/AB

In this contribution, the recently acquired BioSecure Multimodal Database is presented, together with a brief description of previous work in the domain of multimodal biometric database acquisition. This database is the result of an important collaborative effort of 11 European partners of the BioSecure NoE. It includes new challenging acquisition conditions and features not present in existing databases. It is comprised of three different datasets with more than 600 common individuals captured in 2 sessions: i) one dataset acquired over the Internet, ii) another one acquired in an office environment with a desktop PC, and iii) the last one acquired with mobile devices in indoor/outdoor environments. The three datasets include a common part of audio and video data which comprises still images of frontal face and talking face videos acquired with a webcam. Additionally, the second dataset includes still face (with a digital camera), signature, fingerprint (with two different sensors), hand and iris data, and the third one also includes signature and fingerprint data. Also worth noting, the BioSecure Multimodal Database shares a number of individuals with other multimodal databases acquired across several years, allowing studies of long term variability.

The new challenging acquisition conditions of this database will allow the evaluation of realistic multimodal scenarios, as done in the recently conducted BioSecure Multimodal Evaluation Campaign (BMEC). A brief description of this evaluation together with baseline results of individual modalities from the database was also provided in this paper, from which a number of experimental findings related to biometric recognition using mismatched devices and heterogeneous acquisition conditions have been highlighted.

ACKNOWLEDGMENT

This work has been supported by the European NoE BioSecure - Biometrics for Secure Authentication - and by National Projects of the Spanish Ministry of Science and Technology (TEC2006-13141-C03-03, TEC2006-03617/TCM, TIC2002-04495-C02, TEC2005-07212) and the Italian Ministry of Research. The postdoctoral research of author J. Fierrez is supported by a Marie Curie Outgoing International Fellowship. The authors F. Alonso-Fernandez and M. R. Freire are supported by FPI Fellowships from Comunidad de Madrid. The author J. Galbally is supported by a FPU Fellowship from Spanish MEC. Authors Josef Kittler and Norman Poh are supported by the advanced researcher fellowship PA0022_121477 of the Swiss National Science Foundation and by the EU-funded Mobio project grant IST-214324. The author J. Richiardi is supported by the Swiss National Science Foundation. Authors also thank the support and assistance of (in alphabetical order): Dr. Manuele Bicego (UNISS), Prof. Enrico Grosso (UNISS), Dr. Andrea Lagorio (UNISS), Aurélien Mayoue (TELECOM & Management SudParis), Dijana Petrovska (TELECOM & Management SudParis), Ms. Ajita Rattani (UNISS), Florian Verdet (UNIFRI). Authors also would like to thank to the anonymous donors that have contributed to this database.

REFERENCES

[1] A. Jain, A. Ross, S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. Information Forensics and Security*, vol. 1, pp. 125–143, 2006.

[2] P. Flynn, *Handbook of Biometrics*. Springer, 2008, ch. 25. Biometric Databases, pp. 529–548.

[3] A. Mansfield, J. Wayman, *Best Practices in Testing and Reporting Performance of Biometric Devices, v 2.01 - www.cesg.gov.uk/site/ast/biometrics/media/BestPractice.pdf*, 2002.

[4] NIST SRE, "NIST Speaker Recognition Evaluation - http://www.nist.gov/speech/tests/spk/index.htm," 2006.

[5] FRVT, "Face Recognition Vendor Test - http://www.frvt.org," 2006.

[6] FVC2006, "Fingerprint Verification Competition - http://bias.csr.unibo.it/fvc2006/default.asp," 2006.

[7] ICE, "Iris Challenge Evaluation - http://iris.nist.gov/ice," 2006.

[8] SVC, "Signature Verification Competition" - www.cs.ust.hk/svc2004

[9] BioSecure, "Biometrics for Secure Authentication, FP6 NoE, IST - 2002-507634 - http://www.biosecure.info," 2004.

[10] BMEC, "The Biosecure Multimodal Evaluation Campaign - http://biometrics.it-sudparis.eu/bmec2007," 2007.

[11] M. Faundez-Zanuy, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, "Multimodal biometric databases: An overview," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 8, pp. 29–37, 2006.

[12] J. Wayman, "A path forward for multi-biometrics," *IEEE Proc. Int. Conf. on Acoustics, Speech and Signal Processing ICASSP*, pp. 14–19, 2006.

[13] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. Freire, F. Alonso-Fernandez, D. Ramos, D. Toledano, J. Gonzalez-Rodriguez, J. Siguenza, J. Garrido-Salas, E. Anguiano-Rey, G. Gonzalez-de-Rivera, R. Ribalda, M. Faundez-Zanuy, J. Ortega, V. Cardenoso-Payo, A. Viloria, C. Vivaracho, Q. Moro, J. Igarza, J. Sanchez, I.Hernaez, C. Orrite-Urunuela, F. Martinez-Contreras, and J. Gracia-Roche, "Biosecurid: A multimodal biometric database," *Pattern Analysis and Applications (accepted)*, 2009.

[14] J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, no. 4, pp. 1389–1392, April 2007.

[15] B. Dumas, C. Pugin, J. Hennebert, D. Petrovska-Delacretaz, A. Humm, F. Evequoz, R. Ingold, and D. Von Rotz, "MyIdea - multimodal biometrics database, description of acquisition protocols," Proc. COST-275 Workshop on Biometrics on the Internet, Tech. Rep., 2005.

[16] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. les Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacretaz, "BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities," *Proc. Intl. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA*, pp. 845–853, 2003.

[17] D. Dessimoz, J. Richiardi, C. Champod, and A. Drygajlo, "Multimodal biometrics for identity documents (MBIOID)," *Forensic Science International*, vol. 167, p. 154159, 2007.

[18] H. Meng, P. Ching, T. Lee, M. Wai. Mak, B. Mak, Y. Moon, M. Siu, X. Tang, H. Hui, A. Lee, W. Lo, B. Ma, and E. Sio, "The Multi-biometric, Multi-device and Multilingual (M3) corpus," *Proc. of Workshop on Multimodal User Authentication - MMUA*, 2006.

[19] NIST FRGCDB, "http://face.nist.gov/frgc," 2006.

[20] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. Igarza, C. Vivaracho, D. Escudero, and Q. Moro, "MCYT baseline corpus: a bimodal biometric database," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 395–401, December 2003.

[21] E. Bailly-Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz, and J. Thiran, "The banca database and evaluation protocol," *Proc. 3rd International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA*, Springer LNCS-2688, pp. 625–638, 2003.

[22] F. Schiel, S. Steininger, and U. Türk, "The SmartKom multimodal corpus at BAS," *Proc. Intl. Conf. on Language Resources and Evaluation*, 2002.

[23] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The extended M2VTS database," *Proc. 2nd International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA*, 1999.

[24] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, March 1998.

[25] C. Chibelushi, S. Gandon, J. Mason, F. Deravi, and R. Johnston, "Design issues for a Digital Integrated Audio-Visual Database," *Proc. IEE Colloquium on IAVPRSC*, 1999.

[26] BioSec, "Biometrics and security, FP6 IP, IST - 2002-001766 - http://www.biosec.org," 2004.

[27] J. Gonzalez-Rodriguez, D. Torre-Toledano, J. Ortega-Garcia, *Handbook of Biometrics*. Springer, 2008, ch. 8. Voice biometrics, pp. 151–170.

[28] D. Petrovska-Delacretaz, G. Chollet, and B. Dorizzi, Eds., *Guide to Biometric Reference Systems and Performance Evaluation*. Springer-Verlag, London, 2008.

[29] A. Martinez and A. Kak, "PCA versus LDA," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, 2001.

[30] C. Watson, M. Garris, E. Tabassi, C. Wilson, R. McCabe, and S. Janet, *User's Guide to Fingerprint Image Software 2 - NFIS2 (http://fingerprint.nist.gov/NFIS)*, NIST, Ed. NIST, 2004.

[31] Y. Chen, S. Dass, and A. Jain, "Fingerprint quality indices for predicting authentication performance," *Proc. International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA*, Springer LNCS-3546, pp. 160–170, 2005.

[32] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, School of Computer Science and Software Engineering, University of Western Australia, 2003.

[33] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 21–30, 2004.

[34] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, J. Bigun, "Discriminative multimodal biometric authentication based on quality measures," *Pattern Recognition*, vol. 38, no. 5, pp. 777–779, 2005.

[35] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[36] B. Ly Van, S. Garcia-Salicetti, and B. Dorizzi, "On using the Viterbi path along with HMM likelihood information for on-line signature verification," *IEEE Trans. on Systems, Man and Cybernetics, part B*, vol. 37, no. 5, October 2007.

[37] S. Garcia-Salicetti, J. Fierrez-Aguilar, F. Alonso-Fernandez, C. Vielhauer, R. Guest, L. Allano, T. Doan Trung, T. Scheidat, B. Ly Van, J. Dittmann, B. Dorizzi, J. Ortega-Garcia, J. Gonzalez-Rodriguez, M. B. di Castiglione, and M. Fairhurst, "Biosecure reference systems for on-line signature verification: A study of complementarity," *Annals of Telecommunications*, vol. 62, no. 1-2, pp. 36–61, January-February 2007.

[38] J. Fierrez, J. Ortega-Garcia, D. Ramos, J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and signature modeling," *Pattern Recognition Letters*, vol. 28, pp. 2325–2334, 2007.

[39] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, 1989.

[40] A. Ross and R. Nadgir, "A thin-plate spline calibration model for fingerprint sensor interoperability," *IEEE Trans. on Knowledge and Data Engineering*, 2008.

[41] F. Alonso-Fernandez, R. Veldhuis, A. Bazen, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in fingerprint verification: A case study using minutiae- and ridge-based matchers," *Proc. IEEE Intl. Conf. on Control, Automation, Robotics and Vision, ICARCV*, pp. 422–427, 2006.

[42] P. Grother, M. McCabe, C. Watson, M. Indovina, W. Salamon, P. Flanagan, E. Tabassi, E. Newton, and C. Wilson, "MINEX - Performance and interoperability of the INCITS 378 fingerprint template," *NISTIR 7296 - http://fingerprint.nist.gov/minex*, 2005.

[43] F. Alonso-Fernandez, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in signature verification: a case study using tablet pc," *Proc. Intl. Workshop on Biometric Recognition Systems, IWBRS*, vol. Springer LNCS-3781, pp. 180–187, 2005.

[44] U. Uludag, A. Ross, and A. Jain, "Biometric template selection and update: a case study in fingerprints," *Pattern Recognition*, vol. 37, pp. 1533–1542, 2004.

[45] J. Galbally-Herrero, J. Fierrez-Aguilar, J. Rodriguez-Gonzalez, F. Alonso-Fernandez, J. Ortega-Garcia, and M. Tapiador, "On the vulnerability of fingerprint verification systems to fake fingerprint attacks," *Proc. IEEE Intl. Carnahan Conf. on Security Technology, ICCST*, 2006.

[46] P. Grother and E. Tabassi, "Performance of biometric quality measures," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 531–543, 2007.

[47] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun, "A comparative study of fingerprint image quality estimation methods," *IEEE Trans. on Information Forensics and Security*, vol. 2, no. 4, pp. 734–743, 2007.

[48] N. Sickler, S. Elliott, "An evaluation of fingerprint image quality across an elderly population vis-a-vis an 18-25 year old population," *Proc. IEEE Intl. Carnahan Conf. Security Technology, ICCST*, pp. 68–73, 2005.

[49] B. Moghaddam and M. Yang, "Learning gender with support faces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 707–711, 2002.

Javier Ortega-Garcia received the Ph.D. degree "cum laude" in electrical engineering in 1996 from Universidad Politecnica de Madrid, Spain. Dr. Ortega-Garcia is founder and co-director of the Biometric Recognition Group - ATVS. He is currently a Full Professor at the Escuela Politecnica Superior, Universidad Autonoma de Madrid. http://atvs.ii.uam.es

Julian Fierrez received the PhD degree in telecommunications engineering from the Universidad Politecnica de Madrid, Spain, in 2006. Since 2002 he is affiliated with Universidad Autonoma de Madrid in Spain, where he currently holds a Marie Curie Postdoctoral Fellowship, part of which has been spent as a visiting researcher at Michigan State University in USA. http://atvs.ii.uam.es

Fernando Alonso-Fernandez received the PhD degree (Hons.) in Electrical Engineering from the Universidad Politecnica de Madrid, Spain, in 2008. Since 2004, he is with Universidad Autonoma de Madrid in Spain, where he currently holds a Juan de la Cierva Postdoctoral Fellowship of the Spanish MCINN. http://atvs.ii.uam.es

Javier Galbally received the M.Sc degree in electrical engineering in 2005 from Universidad de Cantabria, Spain. Since 2005 he is with the Biometrics Recognition Group - ATVS at the Universidad Autonoma de Madrid, where he is currently working as an assistant researcher pursuing the PhD degree. http://atvs.ii.uam.es

Manuel R. Freire received his M.Sc. degree in Computer Science from Universidad Autonoma de Madrid in 2008. Since 2006 he is a research assistant at the Biometric Recognition Group - ATVS at Universidad Autonoma de Madrid, where he is pursuing the PhD degree. http://atvs.ii.uam.es

Joaquin Gonzalez-Rodriguez received the Ph.D. degree "cum laude" in 1999 in electrical engineering, from Univ. Politecnica de Madrid (UPM), Spain. Dr. Gonzalez-Rodriguez is founder and co-director of the Biometric Recognition Group - ATVS. Since May 2006, he is an Associate Professor at the Computer Science Department at Univ. Autonoma de Madrid, Spain. http://atvs.ii.uam.es

Carmen Garcia-Mateo received the Ph.D. degree (Hons.) in telecommunications engineering from the Universidad Politécnica de Madrid, Spain, in 1993. She is Full Professor in the field of Discrete Signal processing at the Universidad de Vigo. http://www.gts.tsc.uvigo.es

Jose-Luis Alba-Castro received the Ph.D.degree (Hons.) in telecommunications engineering from the Universidad de Vigo, in 1997. He is an Associate Professor of Discrete Signal processing, Pattern Recognition, Image Processing and Biometrics at the Universidad de Vigo. http://www.gts.tsc.uvigo.es

Elisardo Gonzalez-Agulla received the Telecommunications Engineer degree from the Universidad de Vigo, Spain, in 2002, where he is currently pursuing the Ph.D. degree in the field of biometrics for web applications. http://www.gts.tsc.uvigo.es

Enrique Otero-Muras received the Telecommunications Engineer degree from the Universidad de Vigo, Spain, in 2005. He is currently working in a company related to biometrics and video-based security.

Sonia Garcia-Salicetti got her Ph. D. in Computer Science in December 1996 from the University of Paris 6 (Université Pierre et Marie Curie, France). Since 2001, she is an Associate Professor at the Intermedia (Interaction for Multimedia) research team at TELECOM & Management SudParis (Institut TELECOM). http://biometrics.it-sudparis.eu

Lorene Allano got her Ph. D. in Computer Science in January 2009 from TELECOM & Management SudParis (Institut TELECOM). She is currently working as a post-doctoral researcher at CEA (Commissariat à l'Energie Atomique).

Bao Ly-Van received the Ph.D. degree in system optimization and security from the TELECOM & Management SudParis (Institut TELECOM), Evry, France, in 2005. He is currently working as Freelance Consultant in Software Engineering.

Bernadette Dorizzi got her PhD (Thèse d'état) in Theoretical Physics at the University of Orsay (Paris XI-France) in 1983. She is Professor at TELECOM & Management SudParis (Institut TELECOM) since September 1989, at the head of the Electronics and Physics department since 1995 where she is in charge of the Intermedia (Interaction for Multimedia) research team. She is coordinator of the BioSecure Association (http://biosecure.info). http://biometrics.it-sudparis.eu

Josef Kittler received the B.A., Ph.D., and D.Sc. degrees from the University of Cambridge, Cambridge, U.K., in 1971, 1974, and 1991, respectively. He heads the Centre for Vision, Speech, and Signal Processing, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, U.K. http://www.ee.surrey.ac.uk/CVSSP

Thirimachos Bourlai received his M.Sc. in Medical Imaging with Distinction at the University of Surrey (UniS), U.K., in 2002. At UniS he completed his Ph.D. in Aug. 2006 and his Post-Doc in multi-modal biometrics in August 2007. Since Sept. 2007 he is working as a Post-Doc at the University of Houston in the fields of Cardiac Pulse and Stress Detection.

Norman Poh received a PhD (2006) in Computer Science from the Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland. He is a research fellow at CVSSP, University of Surrey since 2006. http://www.ee.surrey.ac.uk/CVSSP

Farzin Deravi received the Ph.D. degree from the University of Wales, Swansea, U.K., in 1987. In 1998, he joined the Department of Electronics at the University of Kent, U.K., where he is the Reader in Information Engineering. http://www.ee.kent.ac.uk

Ming W. R. Ng received the B.Sc. (Hons) degree in electronic and electrical engineering from the Loughborough University of Technology and the Ph.D. degree from the University of Kent in 2004. He is currently a research associate in the Department of Electronics at the University of Kent. http://www.ee.kent.ac.uk

Michael Fairhurst received the Ph.D. degree from the University of Kent, Canterbury, U.K. He has been the Head of the Department of Electronics, University of Kent, Canterbury, U.K., until 2008. He is a member of the Biometrics Working Group (BWG) and the UK Government's Biometrics Assurance Group. http://www.ee.kent.ac.uk

Jean Hennebert received the Ph.D. degree in computer science from the Swiss Federal Institute of Technology, Lausanne, Switzerland, in 1998. Since 2007, he is Professor with the Institute of Business Information Systems, University of Applied Sciences Western Switzerland, HES-SO // Wallis, Sierre, Switzerland. He is also affiliated as Lecturer with the University of Fribourg. http://iig.hevs.ch, http://diuf.unifr.ch/diva

Andreas Humm received the Ph.D. degree in computer science from the University of Fribourg, Fribourg, Switzerland, in 2008. Since 2009, he has worked as a Postdoc in the multimedia engineering DIVA group of the Computer Science Department, University of Fribourg. http://diuf.unifr.ch/diva

Massimo Tistarelli received the Phd in Computer Science and Robotics in 1991 from the University of Genoa, Italy. He is currently Full Professor in Computer Science and director of the Computer Vision Laboratory at the University of Sassari, Italy. http://biometrics.uniss.it/lecturers/tistarelli/bio_tistarelli.php, http://www.architettura.uniss.it/article/articleview/455/1/62/

Linda Brodo is a Researcher at the Faculty of Foreign Languages and Literatures of the University of Sassari (Italy) since 2004. She holds a PhD in Computer Science at University of Verona since 2003. He also collaborates with the Computer Vision Laboratory at University of Sassari on the biometric face recognition field.

Jonas Richiardi received the Ph.D. degree at Signal Processing Institute, Swiss Federal Institute of Technology, Lausanne, Switzerland, in 2008, where he is currently a postdoctoral researcher. He is also the co-founder of the research and development company PatternLab. http://scgwww.epfl.ch

Andrzej Drygajlo received his Ph.D. (summa cum laude) in Electronics Engineering, from the Silesian Technical University, Gliwice, Poland, in 1983. He is the head of the Speech Processing and Biometrics Group at the Swiss Federal Institute of Technology at Lausanne. http://scgwww.epfl.ch

Harald Ganster received his PhD degree with distinction in Technical Mathematics in 1999 from the University of Technology Graz, Austria. He is currently with the Institute of Digital Image Processing at Joanneum Research in Graz and is responsible for the field of biometrics. http://www.joanneum.at

Federico M. Sukno obtained his PhD in Biomedical Engineering at University of Zaragoza (2008). He is currently working on statistical shape models at Pompeu Fabra University, where he is also Assistant Professor in coding theory and analog electronics. http://www.cilab.upf.edu

Sri-Kaushik Pavani is a PhD candidate in Computer Science and Visual Communication in Universitat Pompeu Fabra, Barcelona. He received the M.S. degree in Electrical Engineering from the Texas Tech University in 2003. http://www.cilab.upf.edu

Alejandro Frangi holds PhD in Biomedical Image Analysis from the Image Sciences Institute of the University Medical Center Utrecht. As of September 2004, he joined the Department of Information & Communication Technologies of the Universitat Pompeu Fabra (UPF) in Barcelona, Spain, where he is currently Associate Professor. He leads the Center for Computational Imaging & Simulation Technologies in Biomedicine. http://www.cilab.upf.edu

Lale Akarun got her Ph.D. in Electronic Engineering in 1992 at the Polytechnic University, Brooklyn, New York, USA. Since 2002, she is Professor at the Department of Computer Engineering, Bogazici University, Turkey. http://www.cmpe.boun.edu.tr/pilab

Arman Savran is a Ph.D candidate at the Electrical and Electronics Eng. Dept., Bogazici University, Turkey. He recieved the M.Sc degree in Electrical & Electronics engineering from Bogazici University in 2004. http://www.busim.ee.boun.edu.tr