Contract Number :	IST-2002-507634
Project Acronym :	BioSecure
Project Title :	Biometrics for Secure Authentication
Instrument :	Network of Excellence
Start Date of Project :	01 June, 2004
Duration :	36 months

Deliverable Number :	D6.3.1
Title of Deliverable :	BioSecure application oriented multimodal biometric database
Contractual Due Date :	30 June 2007
Actual Date of Completion :	30 July 2007
Workpackage contributing to the Deliverable :	WP-A6.3 – Acquisition Campaign
Nature of the Deliverable : (R/P/D/O)*	R
Lead Contractor for this Deliverable :	UPM (partner 16)
Author{s} :	Javier Ortega-Garcia, Fernando Alonso-Fernandez, Julian Fierrez-Aguilar

*Nature: R-report, P-prototype, D-demonstrator, O-other

Abstract :	This document describes the BioSecure multimodal database, including database description, statistics and problems encountered..
Keyword List :	Biometrics, multimodal database, sensor, acquisition protocol

Project co-funded by the European Commission within the sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission services)	
RE	Restricted to a group specified by the consortium (including the Commission services)	
CO	Confidential, only for members of the consortium (including the Commission services)	**X**

Commercial in Confidence

DOCUMENT EVOLUTION

Version	Date	Status	Notes
a0	30/06/2007	First draft	Submitted to Activity 6.3 for comments. With contributions from UVIGO, BU, EPFL, GET, JR, UNIFRI, UNIKENT, UNIS, UNISS, UPF
a0	30/07/2007	Second draft	Updated figures from section 2.3
a1	XX/XX/XX	final draft	Approved by Activity 6.3. Version submitted to the Steering Board.
b2	XX/XX/XX	approved version	Version to be submitted to the Commission.

EXECUTIVE SUMMARY

The objective of this document is to describe the BioSecure Multimodal Biometric Database. It has been captured during the second part of the project, from November 2006 to June 2007.

A "call for participation" was made within BioSecure, resulting in 11 partners involved in the acquisition of the BioSecure Multimodal Biometric Database. Several partners in this proposal have long experience in biometric database collection and protocol setting.

Three different datasets have been acquired. The main characteristics of these datasets are:

- Dataset DS1: PC-based, online, unsupervised conditions
- Dataset DS2: PC-based, offline, supervised conditions
- Dataset DS3: Mobile device-based, outdoor, uncontrolled conditions

Additional specifications for this database are provided in this document, including:

- Acquisition figures and statistics
- Description of the three datasets (scenario, data, acquisition protocol, acquisition software)
- Feedback and problems encountered by the partners during the acquisition

CONTENTS

Commercial in Confidence

1 INTRODUCTION

In the biometric field, there is a need for biometric databases as well as common experimental protocols in order to establish common research and technology benchmarks. Several databases, evaluation procedures and competitions have arisen in the last years (e.g. the NIST Speaker Recognition Evaluations, the FERET and FRVT Face Recognition Evaluations, the FVC Fingerprint Verification Competitions, etc). However, evaluation of biometric systems, at this moment, is done in a fragmented way, modality by modality, without any 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. Different partners of BioSecure have been already involved in the acquisition of multimodal databases that include different traits and acquisition scenarios. However, most of the existing multimodal databases include two modalities only. Few databases with hundred of users, multiple traits and several acquisition sessions across time exist, and most of them are under development.

The main integrative effort in BioSecure NoE is the design and collection of a new multimodal database. It has taken place during the second part of the project, from November 2006 to June 2007. This database will allow to create a common and repeatable benchmark of algorithms. The objective of this document is to describe the acquired database.

2   GENERAL FEATURES OF THE BIOSECURE MULTIMODAL DATABASE

2.1   Overview of the three datasets

Three different datasets have been acquired:

DS1 (Internet): Voice, face
- PC-based, on-line, unsupervised (Internet)
- 971 donors, 2 sessions, acquired in ALL acquisition sites (11)
- Responsible: UVIGO (C. Garcia-Mateo)

DS2 (Desktop): Voice, face, signature, fingerprint, hand and iris
- PC-based, off-line, supervised
- 667 donors, 2 sessions, acquired in 7 acquisition sites
- Responsible: UPM (J. Ortega-Garcia)

DS3 (Mobile): Voice, face, signature and fingerprint
- PDA-based, indoor/outdoor
- 713 donors, 2 sessions, acquired in 8 acquisition sites
- Responsible: GET (S. Garcia-Salicetti, B. Dorizzi)

2.2   Common contents for video and audio material

These materials (11 items) are common to the three Datasets (DS1, DS2, and DS3)

Corpus identifier	Item identifier	Data Type	Corpus contents
I	1-2	Image	2 still frontal face images
C	1-2	AV	2 *repetitions of a* PIN code (the same between DSs) (4 digits) (set of 100) in English
C	3-4	AV	2 *repetitions of a* PIN code (the same between DSs) (4 digits) (set of 10) in national language
D	1	AV	Digits from 0 to 9 in English
S	1-2	AV	2 *different* phonetically rich sentences in English (different between DSs)
S	3-4	AV	2 *different* phonetically rich sentences in national language (different between DSs)

I1: still image of frontal face.

I2: another still image of frontal face (*it is recommended to change the lighting conditions, if possible, under the same quality restrictions*).

C1: video with audio uttering in English a 4-digit PIN (about 4 secs). Each donor will have

the same PIN for uttering in English in all the sessions. This PIN will be assigned to the donor from a set of 100 different PINs in order to have approx. 100 groups of 10 donors uttering the same PIN (1000 donors/100 PIN's).

C2: the same as C1 with the same given PIN.

C3: video with audio uttering a 4-digit PIN in "national language" (about 4 secs). (Note: with "national language" we consider the language established for all the donors from each acquisition site). Each donor will have the same PIN for uttering in his "national language" in all the sessions. This PIN will be assigned to the donor from a set of 10 different PIN's in order to have 10 groups of about 10 donors uttering the same PIN in the same language (about 100 donors per site / 10 PIN's per site).

C4: the same as C3 with the same given PIN.

D1: video with audio uttering the "0 to 9" sequence of digits in English (about 10 secs).

S1-2: audio uttering two short sentences from a phonetically balanced corpus in English language (about 10 secs). For each donor, and each session, these sentences will be pseudo-randomly selected from a set of available sentences in the phonetically balanced corpus in English language.

S3-4: audio uttering two short sentences from a phonetically balanced corpus in "national language" (about 10 secs). For each donor, and each session, these sentences will be randomly selected from a set of available sentences in the phonetically balanced corpus in the corresponding "national language".

2.3 Numeric identifiers of the donors

A unique numeric identifier is assigned to each donor in the database. This unique identifier consists of 5 digits.

The former 2 digits identify the acquisition site:
- 01 – UVIGO
- 02 – BU
- 03 – EPFL
- 04 – GET
- 05 – JR
- 06 – UNIFRI
- 07 – UNIKENT
- 08 – UNIS
- 09 – UNISS
- 10 – UPF
- 11 - UPM

The latter 3 digits correspond to an increasing numeric identifier assigned to the donor by the supervisor of the Acquisition Site (starting in 001). The identifier has to be the same established in all the datasets acquired by each Acquisition Site in order to obtain a universal numeric identifier for every donor in the Database. It is an important task for the supervisor to

check for the consistency of all those numeric identifiers assigned to the donors into his Acquisition Site.

2.4 Acquisition figures

Sites involved in the acquisition of the three datasets and final figures of acquired users per acquisition site and per dataset.

NOTE: these figures are not definitive. They have been reported by each ac acquisition site and are pending of the integration of the three datasets.

	Partner	DS1	DS2	DS3
1	**Vigo-UVIGO**	101	101	-
2	**Istambul-BU**	98	-	95
3	**Lausanne-EPFL**	77	-	48
4	**Paris-GET**	117	117	117
5	**Graz-JR**	70	70	-
6	**Fribourg-UNIFRI**	70	-	70
7	**Kent-UNIKENT**	80	80	80
8	**Surrey-UNIS**	104	104	104
9	**Sassari-UNISS**	55	55	-
10	**Barcelona - UPF**	59	-	59
11	**Madrid-UPM**	140	140	140
	TOTAL	**971**	**667**	**713**

3   DATA SET 1 (DS1)

The main purpose of DS1 is to acquire material over the Internet and under uncontrolled situations. In this way, a broad target population can be reached and a quite realistic database can be acquired. For DS1, the acquisition protocol consists of the common part (Section 2.2)

3.1 DS1 data formats

Video: .avi files, MJPG (4:2:2), 320x240, JPG quality 75%
Audio: .wav files, 16 KHz, linear 16 bits, mono
Still frontal face image: .jpg files, 320x240, JPG quality 100%

3.2 System requirements

Finally, because of legal issues, each Acquisition Site has to install the overall acquisition framework for DS1 (J2EE application server + MySQL database). The DS1 framework can be downloaded from http://www.gts.tsc.uvigo.es/~eli/ds1. It includes a document with the instructions for the installation.

3.3 Requirements (client-side)

- Minimum time established between sessions: 2 weeks.
- The first session should be guided. The second session could take place over an uncontrolled scenario, if the next requirements are fulfilled:
  - PC (*client*):
    - Minimum 900 Mhz, 512 MB RAM. In order to avoid low-performance due to memory swapping effect at least **150 MB of free physical RAM are required at the moment of capturing**.
    - For the installation of the Java application, at least *16* MB of Hard Disk will be necessary, and a Java Runtime Environment on the PC is required (at least version 1.4.2).
  - Internet connection:
    - A **broadband connection is required**. It is highly recommended a minimum of 256 Kbps for the upload link, due to the huge size of the data to be transferred to the remote server.
  - Webcam with microphone:
    - *For DS1, the microphone expected to be used for the speech acquisition is the one that is included within the webcam that is employed. Anyway, any realistic low-quality devices are suitable for the DS1 internet scenario, by opposition to the more controlled DS2 desktop conditions.*
    - The following webcams have been tested successfully:
      - Logitech® QuickCam® Chat (≈30 €)
      - Logitech® QuickCam® Express
      - Logitech® QuickCam® Zoom
      - Logitech® QuickCam® for Notebooks Pro
      - Logitech QuickCam Pro 400
      - Logitech PTZ
      - Logitech QuickCam Pro 5000
      - Philips ToUcam PRO II

- Philips SPC900NC
- Genius VideoCam Live

3.4 User guide (donors)

First of all the donor has to be registered by the supervisor into the website https://<yourservername>:8443/ds1. After logging into the website, the donor will access to a web page similar to what is shown in Figure 1. A table is presented with the evaluation status of each item for every session that has been acquired up to the moment. Depending on the supervisor decisions, each evaluated item can be marked as "valid" or "invalid". **The donor has to perform a minimum of 2 valid sessions. A session will be considered as "valid" if it includes no more than 3 items marked as "invalid"**. This information will be included in the progress table presented to the donor as showed in Figure 1.
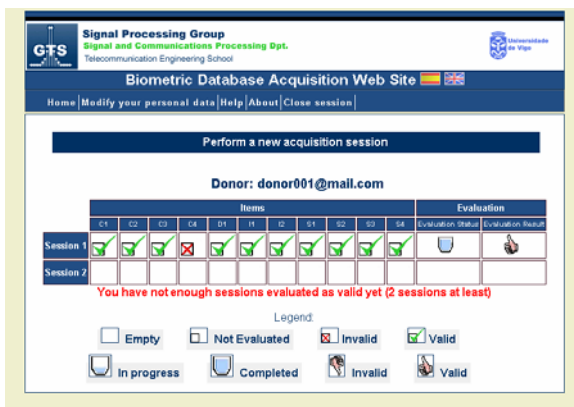
Figure 1

Figure 2 shows the web form with the metadata that the donor has to provide for each new session. For DS1, only two session dependent features for each donor are taken into account: glasses and beard. The web form asking for these metadata is automatically prompted to the donor after clicking on the "Perform a new acquisition session" hyperlink, just before the acquisition application is launched.

The appearance of the graphical user interface of the Java application for the acquisition is showed in the Figures 3, 4, and 5.

Figure 3 represents the state of the user interface prepared for the acquisition of an audiovisual sample. The left-side panel includes the instructions to the donor while the right-side panel shows the webcam stream.

Figure 4 shows the graphical user interface once an audiovisual sample has just been acquired. The sample is presented to the donor in order to be validated before send it to the server. *Note: In case of temporary high computational burden at the moment of acquisition, the audiovisual recording could fail. This fact is especially relevant when the free physical memory is not enough. As a consequence, the previewing of the acquired sample could reflect that the audio and video are not synchronized, or maybe some frames have been lost. Thus, in order to facilitate the job of the supervisor, it is important to persuade the donor to check carefully each acquired sample, and repeat those which are wrong.*

In Figure 5, a frontal still image has just been acquired. The donor has to adjust the position of his face attending the overlaid "virtual glasses and chin" in order to normalize the pose.
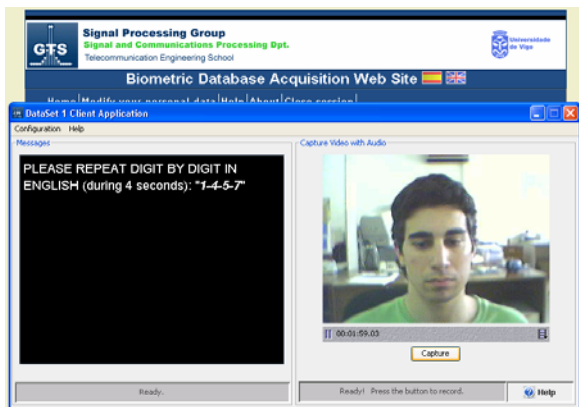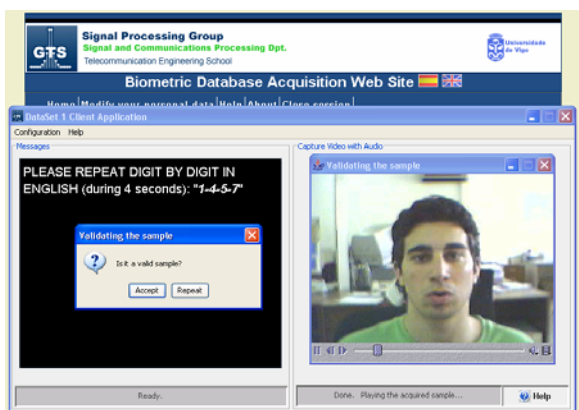

Figure 2
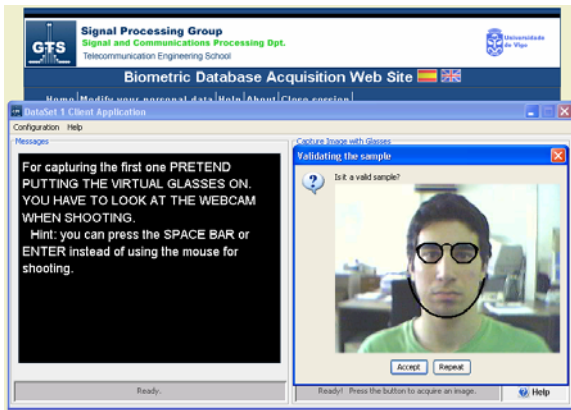

Figure 3

Figure 4

Figure 5

3.5 User guide (supervisors)

For the validation of the DataSet1 content, a web tool for the supervisors is provided. **Each supervisor has to validate the 100% of the data acquired by his donors**. Actually the validation of the first session will be immediate as this first session is expected to be acquired always in presence of the supervisor.

Previous to the start of the acquisition campaign, each supervisor has to contact with UVIGO DS1 administrators (ds1@gts.tsc.uvigo.es) and request his corresponding supervisor account. Once registered, the supervisor will have a donor account with supervisor *and administration* privileges.

The main view of the supervisor web tool ("Supervisor Home" option) is presented in Figure 6. In this case, as an example, a table with the progress for some fictitious donors is showed.
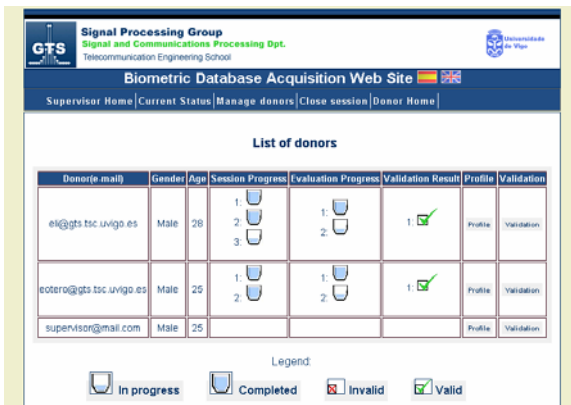

Figure 6

The first task a supervisor has to do for each donor is registering the donor into the website. The data to be introduced in the registration form (Figure 7) have to include a login name, a password, and *a unique numeric identifier of 5 digits assigned to each donor in the database (see Section 2.3).*

Figure 7

Figures *8*, *9* and *10* show some more screenshots presenting the web interface for supervisors.

Figure *8* shows the graphical interface for manage the validation process of a given donor.

Figure 8

Figure *9* shows the web form to be submitted in order to validate each item. The acquired item can be downloaded and reproduced. In the case that the acquired sample was not considered as valid, the "Invalid" option has to be selected and the corresponding invalidation criteria have to be checked. An additional field for possible comments is also provided. In the case of considering an item as "invalid" the content of this field will be showed to the donor as the reason for invalidating this item.

Figure 9

The **invalidation criteria** are the following:

- Uncomplete: if the item acquired does not contain all the required information, e.g., sentence to be uttered is cropped, some audio or video frames are lost, etc.
- Lack of homogeneity: if the item is not homogeneous regarding the remainder items of the session, e.g., the donor wears glasses in the majority of items in the session, but in this item he does not wear glasses.
- Bad utterance: if the actual utterance does not agree with the text asked to be uttered *or the actual language does not correspond to the language asked to be used* (audiovisual samples only).
- Bad audio quality: depends on supervisor criterion (audiovisual samples only).
- Bad image quality: depends on supervisor criterion.
  - *The face image should not be either overexposed (burnt areas) or underexposed (parts or the whole face very dark). These problems are minimized if the camera gain control or exposition control is set to automatic mode.*
  - *The image should not be too bluish or too reddish. These problems are minimized if the white balance control is set to automatic mode.*
  - *If the camera has anti-flicker control this should be set to ON (to avoid the amplitude changes due to fluorescent lamps).*
- *The face image should not be either overexposed (no burnt spots) nor underexposed (parts of the face very darkened)*
- Bad pose of the face: if the face image is cropped, or it is not a frontal face image.
- Other: any other suggestions could be taken into account.

Finally, figure *10* presents some tables that summarize the current status of the acquisition campaign for the corresponding acquisition site. These summary tables can be accessed through the "Current Status" option in the supervisor menu.
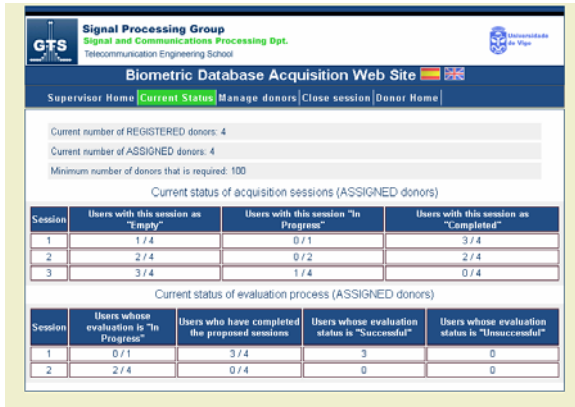
Figure 10

4   DATA SET 2 (DS2)

The scenario considered for the acquisition of DS2 is a PC-based offline supervised scenario:

- PC-based: the acquisition is carried out using a standard PC machine and a number of sensors connected to the PC via USB (or another) interface.
- Offline: acquisition of biometric data is made in different sessions, following a fixed acquisition protocol in each session for all the contributors.
- Supervised: the acquisition is carried out in an office room with a desktop for the PC and the acquisition hardware and comfortable chairs for the supervisor and the contributor. Acquisition is managed by the supervisor, who is in charge of the following activities:
  - Training of the contributor in case of difficulties in using a particular sensor.
  - 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).
  - Guidance of the acquisition process by remembering the steps of the acquisition protocol and pointing out the sensor involved.

4.1   Acquisition devices
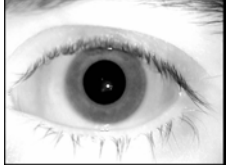
Hardware devices to be used for the acquisition include a Windows-based PC with a USB hub, and the biometric sensors next specified.

PLANTRONICS Voyager 510 (Audio Headset)			
Connect.	**Pre-Processing**	**Micro**	**Headphone**
Bluetooth	Noise cancelation	10 Hz – 10 KHz	80 Hz – 12 KHz

BIOMETRIKA FX2000 (Optical)			
Dimensions (mm)	**Resolution (dpi)**	**Image size (pixels)**	**Capture area (inches)**
122 x 73 x 61	569	400 x 560	0.52 x 0.98

YUBEE (Atmel – Sweeping thermal)			
Dimensions (mm)	**Resolution (dpi)**	**Image size (pixels)**	**Capture area (inches)**
70 x 55 x 25	500	280 x 8	0.55 x 0.02

PHILIPS SPC900NC (WebCam)				
	Dimensions (mm)	**Sensor**	**Resolution (Mpixels)**	**Illuminat.**
	142 x 74 x 74	CCD	1.3	< 1 lux	

LG Iris Access EOU3000				
	Dimensions (mm)	**Sensor**	**Image size (pixels)**	**Illuminat.**
	188 x 203 x 335	CCD	640 x 480	Infrared	

WACOM INTUOS 3 A6 (digitizing tablet)				
	Dimensions (mm)	**Resolution (lpi)**	**Pressure (number of levels)**	**Precision (mm)**
	270 x 216 x 13	5080	1024	+/- 0.25	

CANON EOS 30D (Digital camera) + 2 flashes				
	Dimensions (mm)	**Sensor**	**Resolution (Mpixels)**	
	144 x 105.5 x 73.5	CMOS	8.3		

NOTE: see Annex A for a set of photometric and physical parameters regarding face and hand acquisition for DS2 using the CANON EOS30D camera

4.2 Acquisition scenario

The scenario considered for the acquisition will be an office room, with a wide desktop for the acquisition hardware and two comfortable chairs for the supervisor and the contributor.

The acquisition hardware includes a standard PC machine and a number of data acquisition sensors (specified above) connected to the PC via USB interface by using a USB Hub. Some example images of the acquisition setup installed and used by UPM are given next (Figures 11 and 12).

Figure 11. Diagram of the acquisition scenario for Biosecure DS2
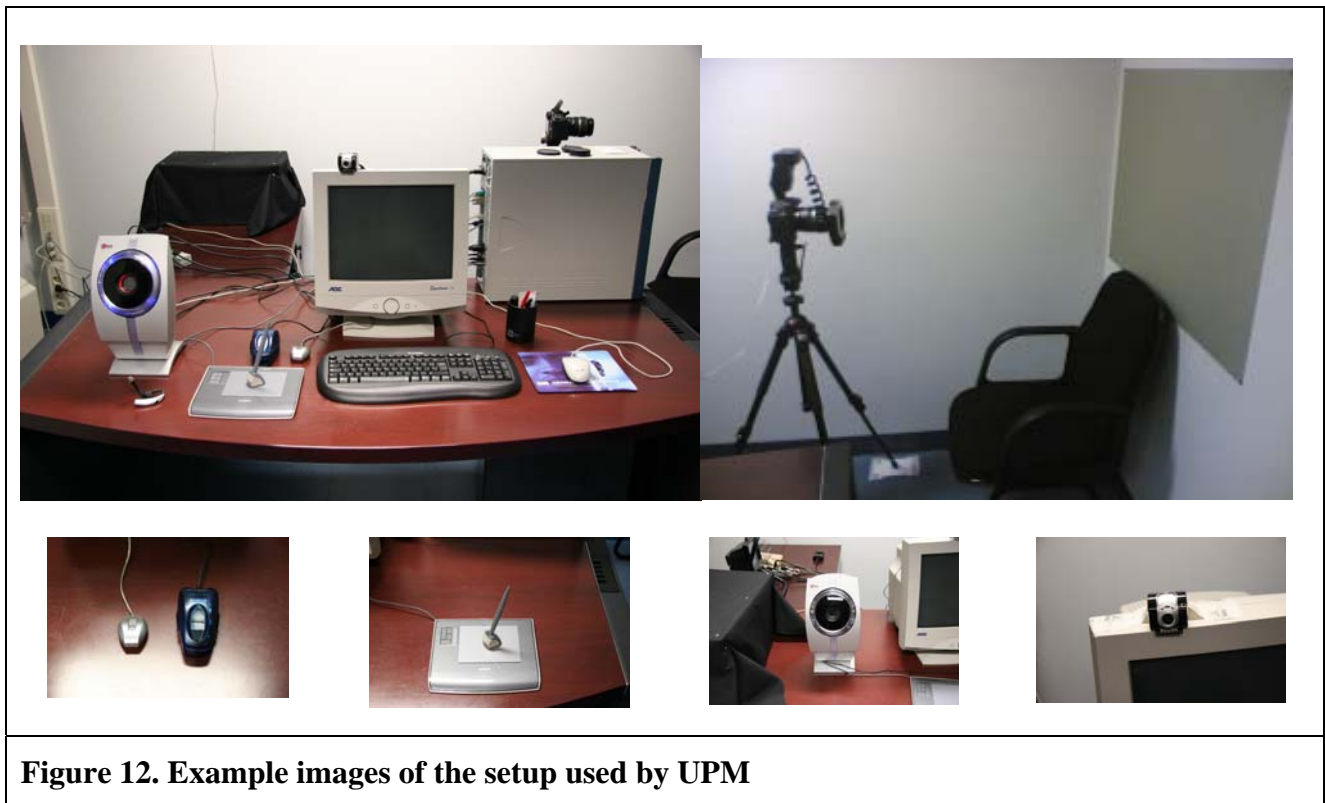
Figure 12. Example images of the setup used by UPM

4.3 DS2 data formats

- Video: .avi files, 720x480
- Audio: 44 kHz, 16 bits, mono
- Webcam still face: .bmp files, 640x480
- Digital camera still face/hand: jpg files, 3504x2336
- Fingerprint optical image: .bmp files, 296x560, 96dpi
- Fingerprint thermal image: .bmp files, 400x500, 96dpi

4.4 Requirements

- Minimum time established between sessions: 2 weeks.
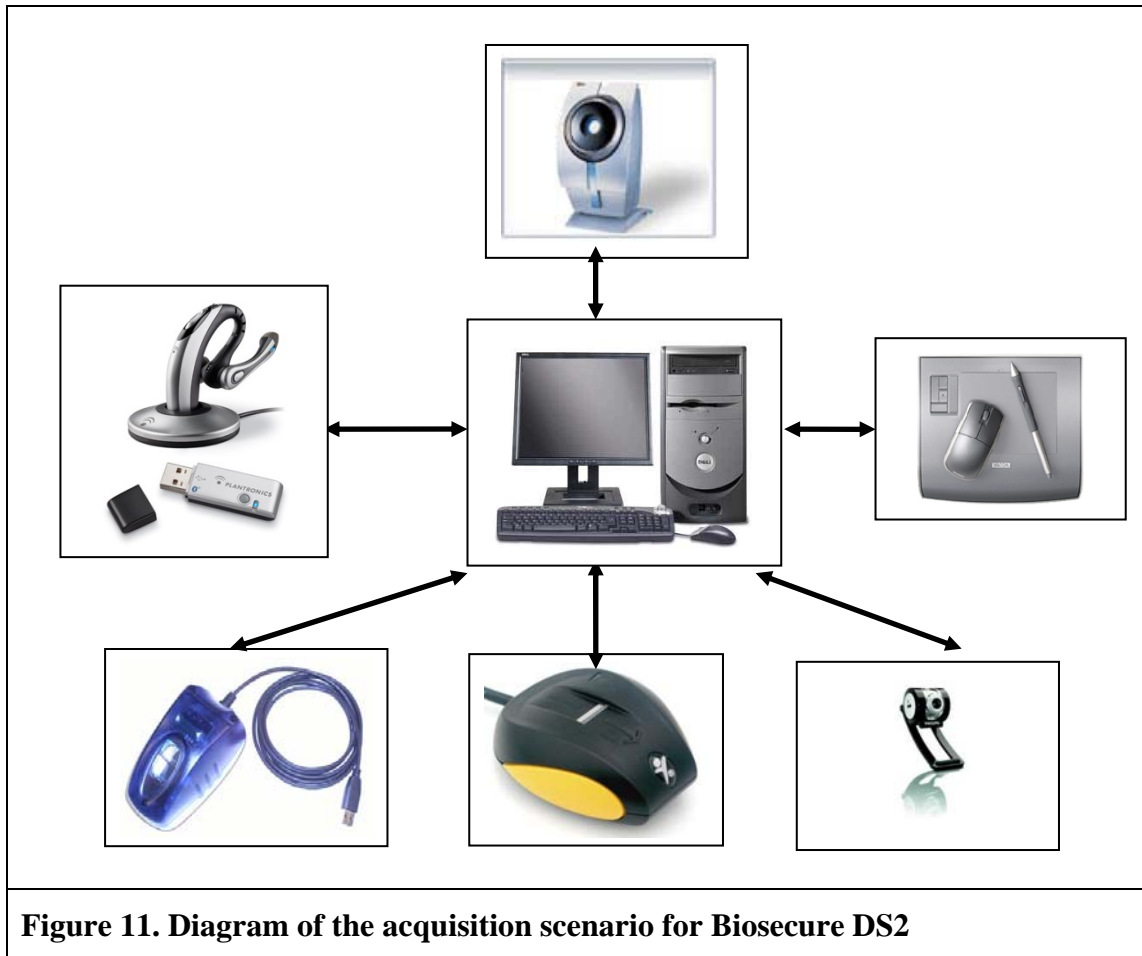- All sessions are guided by the supervisor:
  - PC:
    - Minimum 900 Mhz, 512 MB RAM.
    - About 40 MB of Hard Disk space for the application and the drivers.
  - Internet connection:
    - None.
  - Webcam: The following webcams have been tested successfully:
    - Philips SPC900NC
    - Philips ToUcam PRO II

4.5 Acquisition protocol

2 different sessions (minimum) are required per user, of approx. 20 minutes each. A session, guided by a supervisor, consists in the acquisition of the following samples:

Mode ID	Sample ID	Data Type	Sensor	Contents
SI	1-5	Signature	Tablet	5 genuine of donor n
SI	6-10	Signature	Tablet	5 dynamic imitations of donor $n - 1$ (n–3 session 2)
SI	11-15	Signature	Tablet	5 genuine of donor n
SI	16-20	Signature	Tablet	5 dynamic imitations of donor $n - 2$ (n–4 session 2)
SI	21-25	Signature	Tablet	5 genuine of donor n
Common part	Face 1-2	Face Still	Webcam	Common still frontal face images
Common part	Videos 1-9	Audio-Video	Webcam + earbud	Common Audio-video part
IR	1-4	Iris image	Iris cam	(Right eye $\rightarrow$ Left eye) x 2 times
FO	1-12	Fingerprint	Optical	(R_thumb $\rightarrow$ R_index $\rightarrow$ R_middle $\rightarrow$ L_thumb $\rightarrow$ L_index $\rightarrow$ L_middle) x 2
FT	1-12	Fingerprint	Thermal	(R_thumb $\rightarrow$ R_index $\rightarrow$ R_middle $\rightarrow$ L_thumb $\rightarrow$ L_index $\rightarrow$ L_middle) x 2

HC	1-8	Hand	Digital camera	(Right hand x 2 times → Left hand x 2 times) without flash → (THE SAME) with flash
FC	1-4	Face Still	Digital camera	2 photos without flash → 2 photos with flash (ISO-like conditions)

Phrases are selected pseudo-randomly, following the expression: PhraseID = (UserID + SessionID*50 + SampleID) % TotalNumberOfPhrases. This way, it is easy to track the phrases that each user will say, and it also warranties that no phrase is uttered twice by the same user.

4.6 Acquisition guidelines

In this section we provide some guidelines that supervisors must follow when acquiring the biometric samples. The acquisition software allows repeating the samples inconsistent with the guidelines, both in the same session or even later in other sessions if incorrect or missing samples are detected.

General observations

A spacious desk is recommended to settle the acquisition PC and all the devices. Take note also that the donor and the supervisor will be seated in front of the desk at the same time.

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.
- While invalid samples are those that do not comply with the specifications given for the database (e.g. not a neutral pose on the face acquisition, not saying the right pin in the video/voice acquisition, extremely bad fingerprint image…).

In the next sections we show some valid samples, some valid low quality samples and some invalid samples for each of the biometric traits acquired.

Although some basic guidelines are given for the supervisors, not all the possibilities can be specified, so in the end it is the supervisor himself who have to decide if a given sample is valid or not.

Face Still

- Camera placement (see Annex A for specific details)
  - Webcam: on top of the monitor (or similar position).
  - Digital camera: tripod
- Donor pose: neutral.
- Observations:
  - Ask the user to change its pose between consecutive samples, in order to avoid photos with exactly the same pose. For example: "look at the camera – take photo 1 – look at me … now look at the camera again – take photo 2".
- Invalid samples (see Figure 13):
  - Donor's head out of frame.

- Not a neutral pose.

Face Video
- Camera placement: on top of the monitor (or similar position).
- Donor pose: neutral.
- Observations:
    - Ask the donor to utter the phrase or PIN displayed on the screen.
    - PIN acquisitions must be uttered without pauses between the numbers.
- Invalid samples:
    - Section of donor's head out of frame.
    - Incorrect number.
    - Incorrect language.
    - Incorrect sentence.
    - Not a neutral pose.

Figure 13

Signature
- Observations:
    - Signatures take place over a paper template (print it on the two sides of a blank page) using an inking pen, in order to keep an off-line version of the signature (see Annex B)
    - If there is an error and a sample should be repeated, the supervisor will mark it in the paper template and the donor will sign again on the back of the template page.
    - Donor will be asked to sign naturally (i.e. without breaks or slowdowns), both in genuine and impostor realizations.
- Invalid samples:
    - Moving the paper before the signature is finished (the paper can be moved between signatures).

Iris

- Observations:
    - The correct distance of the eye to the iris camera is achieved with the help of the software application messages: "Too near!", "Too far!", and "Within range".
    - If the donor wears glasses, they must be removed before the acquisition. Contact lenses are acceptable.
    - The donor will be required to put her hand over the eye not being acquired, and to try to fit the other eye in the frame drawn in the mirror over the camera objective.
- Invalid samples (see Figure 14):
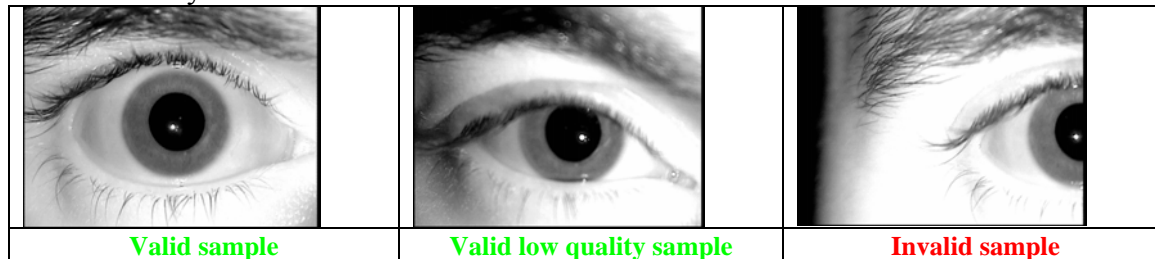   - Section of donor's eye out of frame.
    - Eye closed.

| Valid sample | Valid low quality sample | Invalid sample |

Figure 14

Fingerprint – Optical

- Observations:
    - Ask the donor to put the finger over the device, applying average pressure.
    - It's important that the donor keeps the finger on the sensor until it has been acquired, in order to avoid blank samples.
    - The contact surface of the device should be cleaned after each donor session.
- Invalid samples (see Figure 15)
    - Blank image.
    - Very low quality image.

Fingerprint – Thermal

- Observations:
    - As this sensor is difficult to use correctly, ask the donor to try it multiple times before the acquisition.
- Invalid samples (see Figure 16):
    - Blank image.
    - Very low quality image.
    - Very small size image.

Hand

- Camera placement: tripod (see Annex A)
- Hand pose: palmprint, wide open fingers.
- Observations:
    - The samples must have a dark background, which can be obtained with a grey/black cloth, paper, etc. placed on the background wall. The donor places her/his hand in front of the wall.
- Invalid samples (see Figure 17):
    - Not wide open fingers.
    - Part of the hand out of frame.

Valid sample	Valid low quality sample	Invalid sample

Figure 15

Valid sample	Valid low quality sample	Invalid sample

Figure 16

Valid sample	Invalid sample
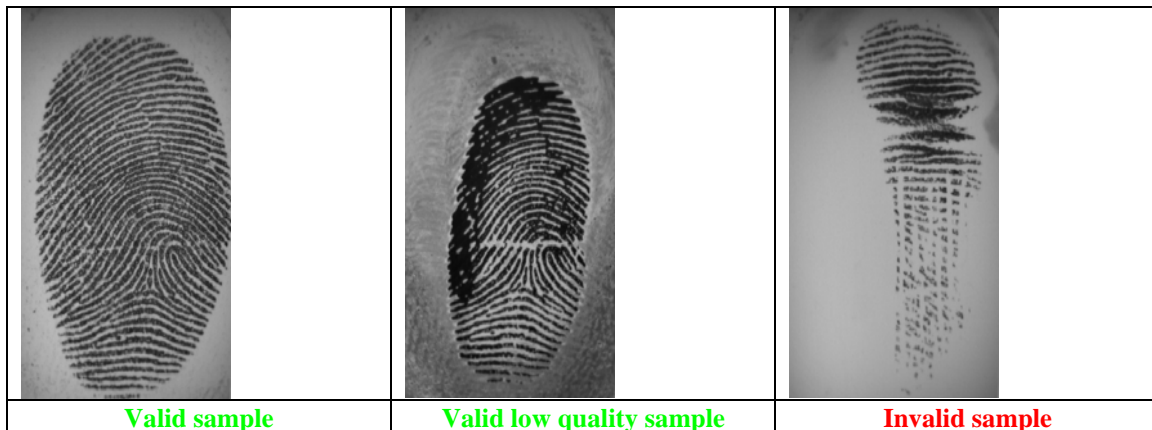
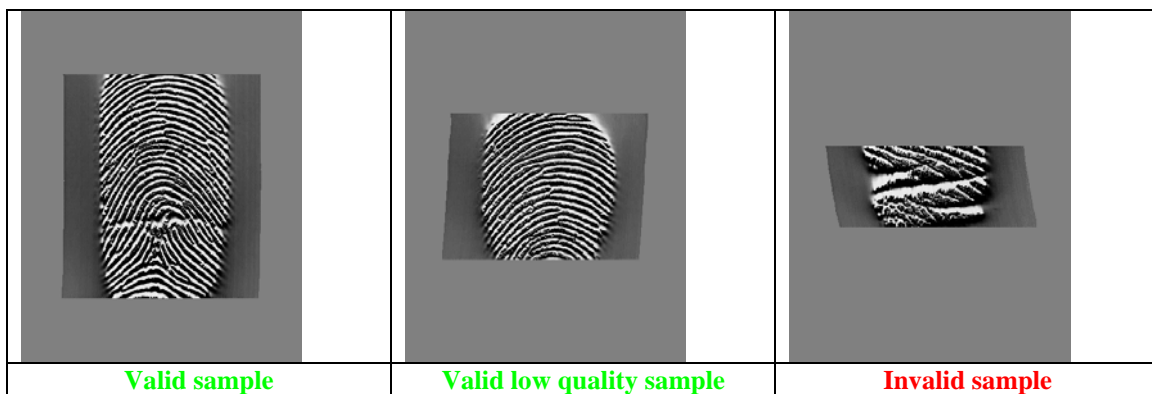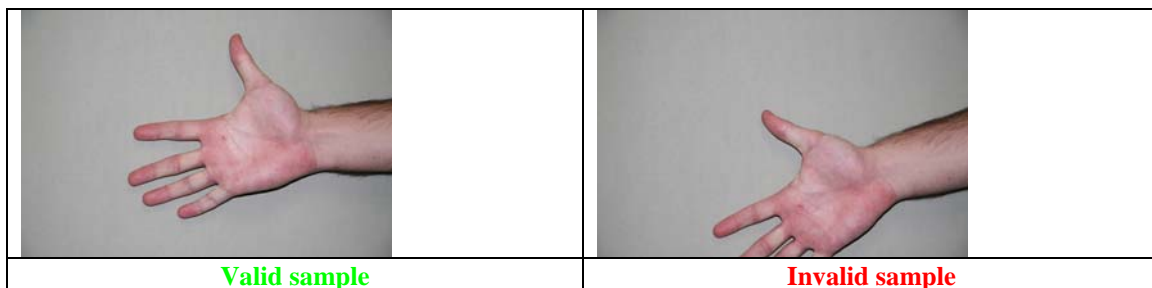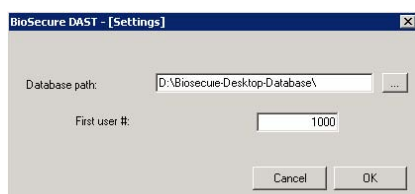**Figure 17**

4.7 Acquisition software: interface

Database creation (first access only)
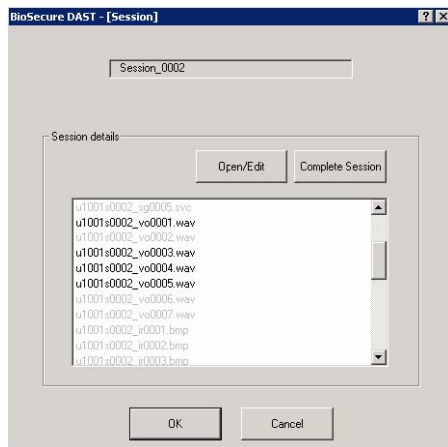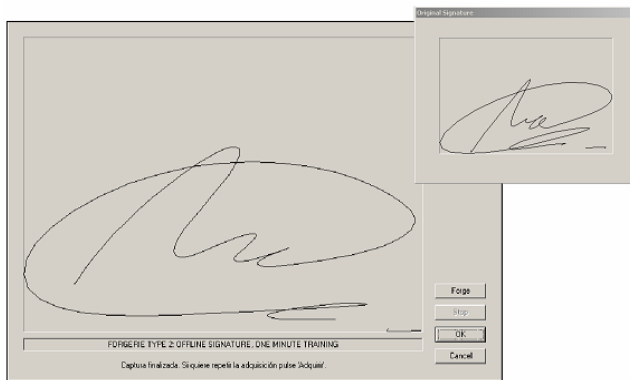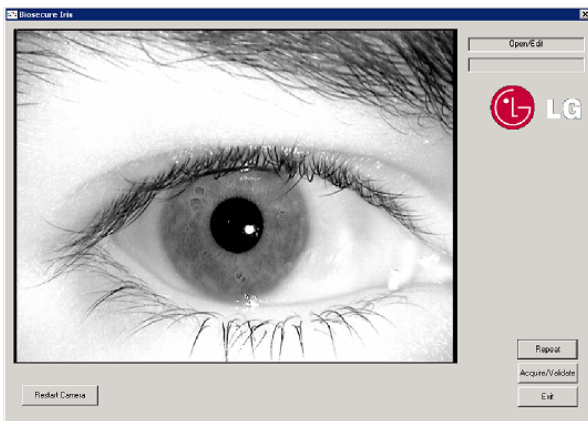
Main application

Session dialog

Video+Voice acquisition

Signature acquisition

Iris acquisition

Fingerprint acquisition

Optical　　　　　　　　　　　　　**Thermal**

5 DATA SET 3 (DS3)

The objective of this dataset is to have a multimodal database with several modalities that are acquired on mobile platforms nowadays. Four modalities will be considered: Face, Voice, Fingerprint and Signature.

5.1 DS3 description

In order to have time-variability in the database, 2 sessions between November 2006 and May 2007 have been considered. Each session comprise 2 acquisition conditions for audio video recordings done the same day: indoor and outdoor.

- Minimum time established between sessions: 2 weeks.
- All sessions are guided by the supervisor:

For audio-video recordings we will consider different acquisition conditions: indoor and outdoor, with sources of variability for each modality:

- Voice:
  ➢ indoor: quiet room
  ➢ outdoor: ambient noise (street noise, people talking …)

- Face (still images and video sequences): changing light (changes in intensity and direction of light)
  ➢ Indoor: the donor is moving in the room, just changing his or her position between each video sequence (from a corner to the other, in front of the window or not ….).
  ➢ Outdoor: the donor is moving, just changing his or her position between each video sequence.

For signature and fingerprint recordings we will consider only one degraded acquisition condition because controlled quality data will be recorded in DS2:

- Signature:
  ➢ Indoor: signing while standing and the user has to keep the PDA in his or her hand.

- Fingerprint:
  ➢ Indoor: giving fingerprints while standing and the user has to keep the PDA in his or her hand.

Subjects must be capable of reading and speaking in English

5.2 Acquisition devices

Hardware devices used for the acquisition include a PDA HP iPAQ hx2790 (for fingerprint and signature) and a Mobile PC SAMSUNG Q1 + Web Cam (for face and voice), see Figure 18.

5.3 DS3 data formats

Video: frequency ≈ 20 frames per s, resolution 240*320 pixels
Voice: frequency 22 kHz (Maximum frequency 44 kHz)
Signature: frequency ≈ 100 Hz and resolution of touch screen 1280*960 pixels
Fingerprint: width resolution 280 pixels

5.4 DS3 protocol and recommendations

For each session, all acquisitions are done during the same day. Each session is composed of multiple recordings (indoor –outdoor).

Complete protocol

Condition	Mode ID	Sample ID	Data Type	Sensor	Contents
Indoor / standing	SI	1-5	Signature	IPAQ	5 signatures of donor n
Indoor / standing	SI	6-10	Signature	IPAQ	5 dynamic imitations of donor (n-1 for session 1 and n-3 for session 2)
Indoor / standing	SI	11-15	Signature	IPAQ	5 signatures of donor n
Indoor / standing	SI	16-20	Signature	IPAQ	5 dynamic imitations of donor (n-2 for session 1 and n-4 for session 2)
Indoor / standing	SI	21-25	Signature	IPAQ	5 signatures of donor n
Indoor / standing	FT	1-12	Fingerprint	IPAQ	2x (R_thumb -> R_index -> R_middle ->L_thumb -> L_index -> L_middle)
Indoor	Common part	1-11 Face 1-2 Videos 1-9	Audio-Video	Q1+ webcam	Common Audio-video part
Outdoor	Common part	1-11 Face 3-4 Videos 10-18	Audio-Video	Q1+ webcam	Common Audio-video part

Signature (PDA)

15 genuine signatures are acquired for each session. The donor will be asked many times to sign alternatively genuine and forged signatures

Forgeries:
At each session, a donor is asked to imitate 5 times the signature of two other persons (for example client n-1 and n-2). In order to imitate the dynamics of the signature, the forger can replay the signature they have to imitate on the screen with the dynamic process. The donor has to train a little before to save the forged signatures.

| HP iPAQ hx2790 | SAMSUNG Q1 + WebCam |

Figure 18: Sensors used in DS3

**Fingerprint (PDA)**

At each session:
- 6 fingers are considered: right and left thumbs, indexes and middles.
- 2 fingerprint images (of the 6 fingers) are acquired.

The donor will have to swipe his right and fingers on the fingerprint sensor. The finger should be straight and flat. The donor's finger must touch the sensor. The acquisition must be done from the first phalanx to the end of the finger.

The supervisor has to check on the screen that the fingerprint image is good enough. It should be large and well contrasted. Be also careful that the donor does not rotate the finger but slides it straight. Here are some examples of good and bad quality images.

Good quality images   bad quality images: too small
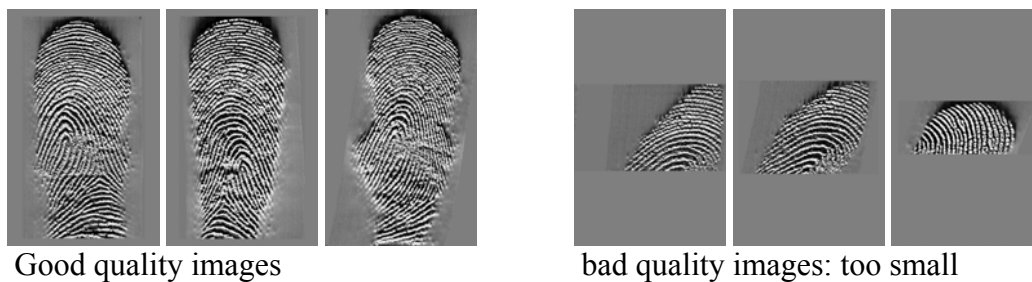
Specific conditions for the common audio-video part (Q1)

I: still frontal images

- Change of lighting conditions:
    - Indoor: I1 close to the window, I2 far from the window
    - Outdoor: I1 and I2 with different light directions
- For each image, donor has to check on the screen if he or she recognised himself or herself as a quality control.
- Camera placement: web cam on top of Q1 using the specific plastic piece distributed to each acquisition site.
- Donor pose: neutral.
- Invalidation causes: Section of donor's head out of frame (see Figure 19).

AV recordings (C, D and S)

- For indoor and outdoor videos acquisition, PINs and sentences will be identical. The same audio-video part is repeated 2 times, once indoor and once outdoor. Acquisitions will be done successively.
- Between each of the 9 audio-video items, donor is asked to move a little in the room (or outside). For each recording, he or she has to check on the screen if he or she recognised himself or herself as a quality control.
- Camera placement: web cam on top of Q1 using the specific plastic piece distributed to each acquisition site.
- Donor pose: neutral.
- Observations:
  - Ask the donor to utter the phrase or pin displayed on the screen. Take note that the donor will not be looking at the camera.
  - Pin acquisitions must be uttered naturally, without pauses between the numbers.
- Invalidation causes:
  - Section of donor's head out of frame (see Figure 19).
  - Incorrect number.
  - Incorrect language.

| **Invalid sample** | **Valid sample** |
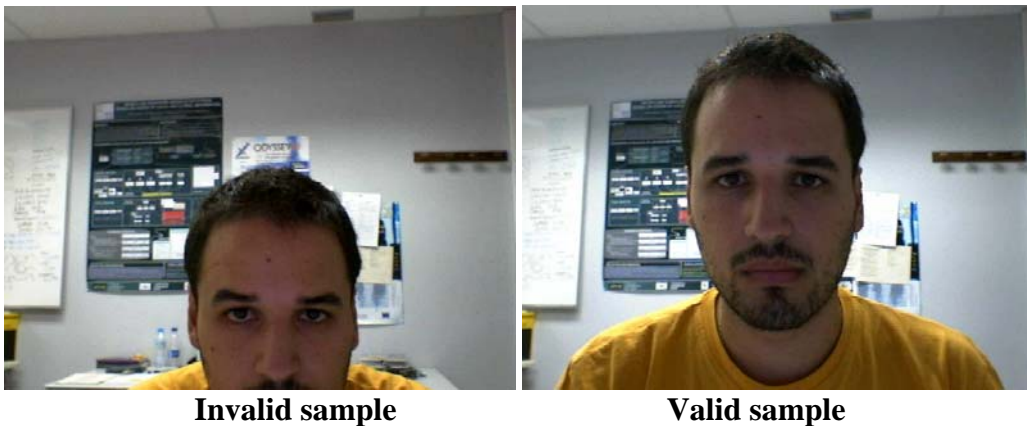
Figure 19

5.5 DS3 acquisition program for PDA and Q1

Main interface, signature acquisition screen and fingerprint acquisition screen of the acquisition program for PDA:
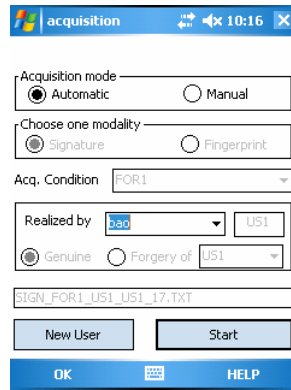
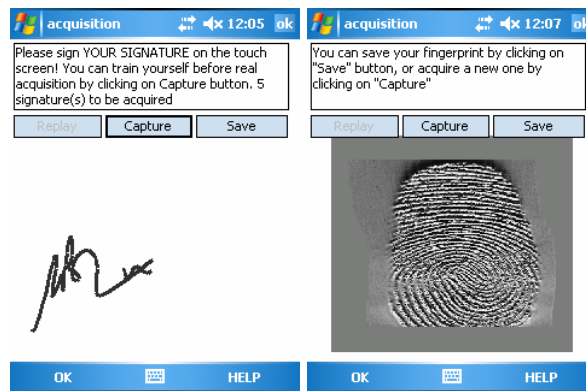**Figure 20: Main interface of the acquisition program**

Figure 21: Signature (left) and fingerprint (right) acquisition

The acquisition program for Q1 is the same software that for DS2 acquisition. It has been adapted to Q1 and restricted to the audio-video part

6 FEEDBACK AND DIFFICULTIES ENCOUNTERED DURING THE ACQUISITION OF THE DATABASE BY THE ACQUISITION SITES

6.1 UNIKENT

- We lost a lot of collection time due to absentees.

For DS1:
- Synchronization problem occurred frequently.
- Sometimes, the acquisition needs to be refreshed many times during a session.
- Once wrongly accepted, the video clip cannot be edited.

For DS2:
- The Yubee thermal fingerprint sensor was difficult to use, requiring many trials to get a reasonably good fingerprint capture.
- The auto-focussing function of the iris camera was not always working, needed many trials to get a reasonably sharp iris image capture.
- For some volunteers doing the face pictures, there were too many reflections on their glasses.
- The automatic exposure control on the Canon digital SLR was not consistent, and some collected pictures were under-exposed.

For DS3 (PDA)
- There is no viewer to see the signatures collected.
- Once wrongly saved, the signature data and the fingerprint data cannot be cancelled and edited.
- Occasionally, message appeared saying capture could not be saved for no obvious reason.
- Wireless network calls disturbed the program frequently.

For DS3 (Q1)
- It was very difficult to ask the volunteer to hold the Q1 while speaking and looking at the camera, we had to mount the Q1 on a tripod.

6.2 UNIFRI
General remarks:
- We realized that the acquisitions (including recordings, paperworks, validations, recruitments) took more time as initially planned, almost the double of the time! We have to note here that we did not choose to sacrifice the recording quality to compress the acquisition time.
- Due to this, many users simply declined to do the acquisition (we were maybe too honest to announce the real time needed for the recordings). This made our recruitment procedure even more difficult and time-consuming.
- We reduced the targeted number of users because we were instructed to finish the

acquisition phase in the end of March. This was also a question of limited budget.

For DS3:

- Our HP iPaq revealed to be pretty much unstable; we had several problems with the battery and frequent resets were necessary...
- The indoor and outdoor session (and the DS1 videos) was too much of video sequences for the donors (they were really bored and wanted to do the outdoor session quickly)
- The cancel button of the acquisition software was confusing the donors. If they did a mistake in the acquisition, they naturally clicked on cancel, making us re-record this video at the end of the session.

6.3 BU

- Our DS3 acquisition began in February due to several reasons, and therefore we will be completing the acquisitions with some delay. The main reason is our difficulty in purchasing Samsung Q1 until January. It was not sold in our country, and we could not find it in some other European countries during visits. In addition, we had some problems while trying to purchase it via internet. The hardware required for the acquisition (Samsung Q1) was not a standard tool, therefore if a hardware problem was encountered it was difficult to find the replacement parts
- After the purchase in January, we had a problem in the capture software of Q1. It was not working with the Turkish sentences. We had to wait for the update of the software, and therefore we were able to start the acquisition in the third week of February.
- Once we started acquisition, we noticed a dominant acoustic noise in the Q1 records. We realized that it was due to the camera's microphone which is very close to Q1's fan, and solved it by using the Q1's microphone instead of the camera's. But this took some time, and we had to repeat the previous acquisitions taken up to that date.
- We also had a problem while using the acquisition software in the PDA. It abruptly erased the identity file of the record database. Therefore, we had to solve the coding in this file to create a file concordant with the records, since we did not have a recent backup of this file. Without the correct identity file, it was not possible to go on with the acquisitions.
- During acquisition of the still face images with the Q1, the display screen of the camera goes black. The only way to handle this problem is to unplug the USB interface of the camera which in turn causes Q1 to flag a runtime error and to restart. This restarting issue causes an extra 5 minutes which might be disappointing for the donor. Therefore instead of restarting the operating system, we also tried to exit and reopen the program with "Complete Session" option; however this could not solve the black screen problem.
- During swiping the finger on the fingerprint sensor of the PDA, touching the screen or the button just below the sensor ends the capture mode of the software. Therefore, acquiring valid fingerprint data becomes a time consuming process.
- In the recommendations document for DS3 acquisition, there are some points which are not stated clearly. For example, it is stated that in the case of impossibility to do outdoor recordings, we can do the acquisitions indoors only if the proper lighting conditions are provided. We are confused about the proper settings such as the light parameters.
- Unable to edit a record in a session unless you perform the whole acquisition of that session.
- Erroneous native sentences were present.
- The acquisition was to be performed in a quiet room, but this was difficult to achieve.

- The data was huge in size, no compressions were applied to the videos, and therefore the backup procedure was difficult.
- If the acquisition session in the HP was interrupted, it was not possible to continue, and hence the acquisition had to be performed for the start.
- In Q1, sometimes a black screen appears and to solve the problem the system has to be restarted.
- In Q1, sometimes the acquired video was not saved.

6.4 UVIGO

- Due to borrowing the iris scanner from UAM, the first session of DS2 was performed without iris images capturing, and the second session with double number of iris samples.
- About the first 60 people from DS1 had synchronicity problems. After that we change the equipment for other more powerful and the rest of videos were perfectly synchronized and with a higher rate. To compensate from these problems we will ask some people to repeat DS1 session from home (on-line).
- Regarding the iris samples, we also detected problems having a focused image easily, and this problem delays the acquisition session. We tried turning the lights off.
- Focus on iris scanner sometimes did not corresponded to the acquisition instant and images got a little bit blurred
- Fingerprints on swept sensor caused annoyance in some users because of Failure to Enroll
- Low English proficiency caused very poor utterances in many cases.
- For DS1, with slow PCs some frames are missed or the video and audio were unsynchronized

6.5 INT

- We had some problems and particularly did not manage to use properly the Bluetooth microphone.
- Another point was very bad quality images with the iris camera. We solved the problem in turning off the light.
- We also have synchronisation problems in DS1 for some donors.

6.6 JR

- On one hand we had / have some problems getting students for the acquisition (due to being a non-university research institute), but on the other hand due to the same reason the majority will be in the group of 30-40 year old donors (which will complete the probably 'younger' data of the university groups nicely).
- Concerning the hard- / software there were only minor issues (Wacom tablet driver version). We noticed some problems with the Bluetooth headset; one has to be careful to have a good enough connection to get a reasonable audio quality.

6.7 UPF

- The biggest problem we have is the lack of volunteers. There are two reasons why people say no: 1) Lack of interest/time 2) Fear that their biometric data will be misused.
- Upon suggestion, we are planning to give gifts to the people who participate in the database collection. Hopefully, this should encourage people to participate.

6.8 EPFL

For DS1:
- We have 16 first session data for users that was acquired on the remote server in Vigo, which was lost when transitioning over to local server. This is because we started the acquisitions before it was decided that each site would host their own server.

For DS3:
- SD card failed (hardware), causing data loss and confusing the acquisition software: some users were not presented with forgeries, others have missing data. GET is informed about the detailed situation.

6.9 UNISS

- Impossible to respect the age categories.
- It was not easy to capture fingerprint, both optical and thermal
- The laboratory lights were no optimal
- There are background noises during video registration.

ANNEX A

PHOTOMETRIC AND PHYSICAL PARAMETERS REGARDING FACE AND HAND ACQUISITION FOR DS2 USING THE CANON EOS30D CAMERA

- CANON EOS30D camera set-up

This section has been produced after some tests and discussions during September and October 2006. Finally we have agreed a set of photometric and physical parameters regarding face and hand acquisition for DS2 using the CANON EOS30D camera. The main objective of this collaborative work has been to end up with the easiest way to ensure a high and homogeneous picture quality among the partners involved in DS2.

The main design specifications we had in mind were:
- Targets (face/hand) evenly illuminated without shadows or burn spots
- Images without geometrical distortion
- Minimum parameters change between shots

Physical parameters: regarding distances and background
- Distance from CANON EOS 30D *RING FLASH FRONT* to the background (or to target HAND which will be against the background): **100cm**
- Chair: To have persons seated during the FACE acquisition (standing during HAND acquisition), with low back; adjustable in seat height.
- Distance from CANON EOS 30D *RING FLASH FRONT* to target FACE front: **70-80cm**. Note that a CHAIR will be against the wall, so the HEAD will be very close to the background.
- Distance from target face front to background: **20cm-30cm**
- Distance from hand to background: ~**0 cm**
- Background type and color: Uniform GREY background (suggested in the ISO-standard).
- Tripod: Distance from Tripod's head to bottom set to **119cm**.

Illumination parameters: regarding room lighting
- For shots without flash, a good, evenly illuminated target should be achieved. If room has a main source of illumination (lateral window, for example), try to compensate it with more sources from other directions. UniS has designed a good "standardization" solution described below, that can help partners to produce similar face and hand pictures. Do not worry if you have incandescent or halogen lamps instead of the ceiling OSRAM fluorescent described by UniS. Take the pictures in the Annex as a reference of good illuminated images.
- For shots with flash, just turn on the **Sigma EM-140 DG**. Set it to E-TTL II mode, to achieve high-precision and consistent flash shots. Easy set-up on the camera lenses.

CANON EOS30D parameters: regarding photographic values
- Select **Tv program (shutter speed priority)**: to **1/60 sec** (with Flash and with NO Flash: see attached images in Annex A).
- Image Quality: Select **Large/Fine**.
- Auto Focus Mode: Select **AF ON** (AUTO Focus ON) and **ONE SHOT**.
- Metering Mode: Select **Spot Metering** (Central Point Set-Up).
- DRIVE-ISO: Select Single Shooting and Set ISO to 400.
- Optical Zoom: Set the **dot ~30mm** for FACE and **~50mm** for HAND acquisition.
- White Balance: Set it to **AWD** (Auto White Balance).

Remaining ISO-like specifications (see below for a reminder on good practices for taking ISO-like face shots)

UNIS SUGGESTIONS ON ROOM LIGHTING AND CAMERA-FRAME SET-UP

The **"Camera-Frame Set-Up"** with which some test images have been taken has the following **specifications** (see Figure next):

Figure 1: UniS Camera-Frame Set-Up

CAMERA-FRAME SET-UP
- 88cm X 88cm
- Distance from the ground to bottom frame: 49cm
- 20W Fluorescent Lights **X 3**, around the frame

* **Note: the lights on the frame are on with and without the use of the flash.**

ROOM
- Dimensions: WIDTH: **4.0 m**, LENGTH: **4.65 m** and HEIGHT: **2.5 m**
- **No** windows.
- Room Lights: **8 to 10 OSRAM L 36W/23** (please see *Figure 2* below)

Figure 2: Lights: A pair of OSRAM L 36W/23.

File names of attached images
 FL/nFL = Flash On / No Flash,
8_OL= 8 Osram Lights,
10_OL= 10 Osram Lights (all ceiling lights on),
S_1_60=shutter speed 1/60
A_6.0=aperture 6.0

Remind that:
* The "Flash to Wall Distance" is set either 100cm or 70cm, and
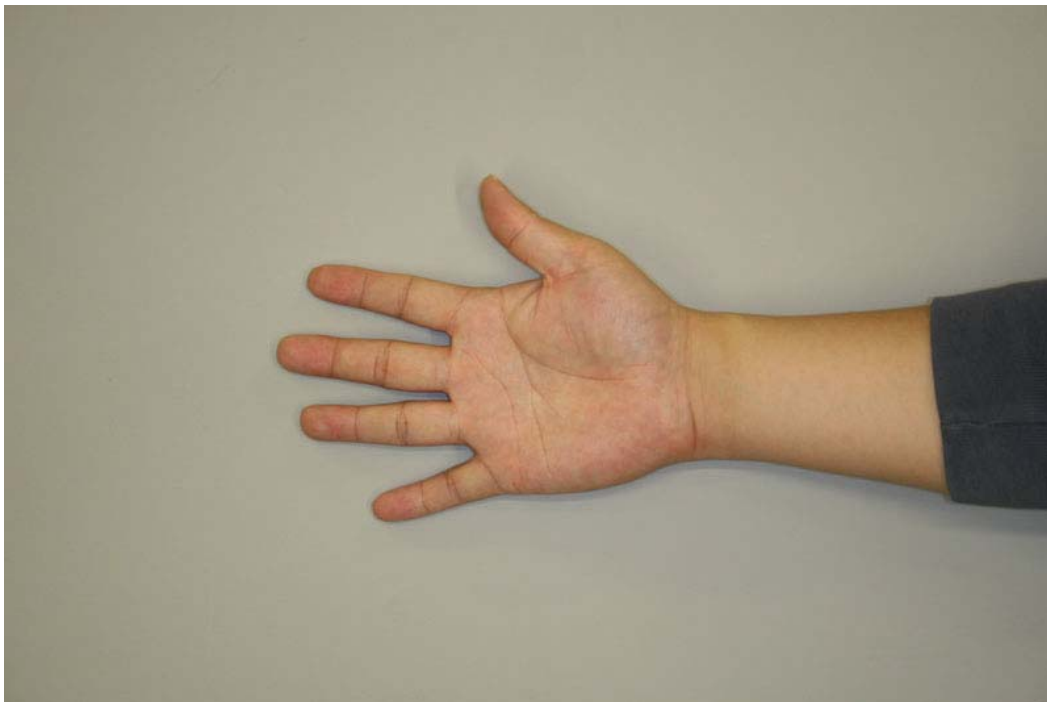* The "Optical Zoom" is set either 50/45mm (Hand) or 30mm (Face)

For example, **"HAND_FL_10_OL_S_1_60_A_6.0_100cm_50mm"** means a HAND image
with Flash ON, 10 (ALL) OSRAM Lights on the ceiling, Shutter Speed 1/60, Aperture 6.0,
"Flash to Wall Distance" set to 100cm and "Optical Zoom" set to 50mm.

FACE_FL_10_OL_Fr_on_S_1_60_A_5.6_100cm_30m.jpg

FACE_nFL_10_OL_Fr_on_S_1_60_A_5.6_100cm_30m.jpg

HAND_FL_10_OL_Fr_on_S_1_60_A_5.6_100cm_50m.jpg

HAND_nFL_10_OL_Fr_on_S_1_60_A_5.6_100cm_50m.jpg

ANNEX B

PAPER TEMPLATE FOR SIGNATURE ACQUISITION IN DS2

Name:_____ ID:_____ Session:___

GENUINE

Session: Real ID: Imitating: Real ID: Imitating:

FORGERY