



UNIVERSITÉ DE FRIBOURG SUISSE
UNIVERSITÄT FREIBURG SCHWEIZ

MYIDEA - SENSORS SPECIFICATIONS AND ACQUISITION PROTOCOL

DIUF-RR 05-12

By alphabetical order : Bruno Dumas¹ Jean Hennebert²
Andreas Humm³ Rolf Ingold⁴ Dijana Petrovska⁵
Catherine Pugin⁶ Didier Von Rotz⁷

INITIAL: JANUARY 2005 - REVISION 1.2.5 AS OF JUNE 2005

COMPUTER SCIENCE DEPARTMENT RESEARCH REPORT

Département d'Informatique - Departement für Informatik • Université de Fribourg -
Universität Freiburg • Chemin du musée 3 • 1700 Fribourg • Switzerland

phone +41 (26) 300 84 65 fax +41 (26) 300 97 31 Diuf-secr-pe@unifr.ch <http://diuf.unifr.ch>

¹DIUF, University of Fribourg, Ch. du Musée 3, 1700 Fribourg, Switzerland, bruno.dumas@unifr.ch

²DIUF, University of Fribourg, Ch. du Musée 3, 1700 Fribourg, Switzerland, jean.hennebert@unifr.ch

³DIUF, University of Fribourg, Ch. du Musée 3, 1700 Fribourg, Switzerland, andreas.humm@unifr.ch

⁴DIUF, University of Fribourg, Ch. du Musée 3, 1700 Fribourg, Switzerland, rolf.ingold@unifr.ch

⁵INT, Dept. EPH, Intermedia, 9 rue Charles Fourier 91011 Evry, France, dijana.petrovska@int-evry.fr

⁶DIUF, University of Fribourg, Ch. du Musée 3, 1700 Fribourg, Switzerland, catherine.pugin@unifr.ch

⁷EIF, Bd de Pérolles 80 - CP 32, 1705 Fribourg, Switzerland, didier.vonrotz@eif.ch

Abstract

In this document we describe the sensor specifications and acquisition protocol of MyIdea, a new large and realistic multi-modal biometric database designed to conduct research experiments in *Identity Verification (IV)*. The key points of MyIdea are twofold.

1. it is strongly multi-modal, allowing the studies of potential correlations between synchronized/unsynchronized modalities
2. it implements realistic scenarios in an open-set framework

The combination of these two points makes MyIdea novel and pretty unique in comparison to existing databases. Furthermore, special care is put in the design of the acquisition procedures to allow MyIdea to complement existing databases, such as, for example the BANCA database. MyIdea includes face, audio, fingerprints, signature, handwriting and hand geometry. Further to the asynchronous acquisition of each modality, two synchronized recordings are performed: face-voice and writing-voice. The general specifications of MyIdea are: target of 104 subjects, different quality of sensors, various but realistic acquisition scenarios, organization of the recordings to allow for open-set experimental scenarios.

Keywords: MyIdea, multimodal biometric database

1 Overview

Multi-modal biometrics has raised a growing interest in the industrial and scientific community. Indeed, it is expected that multi-modal biometrics will improve reliability by combining multiple source of data and robustness by making forgeries much more difficult to realize. From the user point of view, these security improvements will also certainly ease the acceptance of the technology. From a scientific point of view, multi-modal biometrics raise many important issues: how to model and combine multiple sources of biometric features, are the modalities correlated, what are the impact of sensors quality in a multi-modal setting There is a clear need for multi-modal databases to allow the scientific community to analyze, develop and benchmark their multi-modal *Identity Verification (IV)* algorithms. As a matter of fact, the currently existing multi-modal databases are small with few recorded modalities, often implementing unrealistic closed-set scenarios and sometimes using even more unrealistic acquisition conditions.

MyIdea includes **face, voice, fingerprints, signature, handwriting, palmprint and hand geometry**. Further to the unsynchronized acquisition of each modality, two synchronized recordings are performed: **face-voice** (video sequences of the talking faces of the subjects) and **writing-voice** (the subject is asked to read aloud his/her signature and what (s)he is writing). Table 1 summarizes the modalities recorded in MyIdea along with the sensors and scenarios. The general specifications of MyIdea are:

- target of 104 subjects who will participate to successive recording sessions spaced in time
- different type and quality of sensors
- various but realistic acquisition scenarios
- organization of the recordings to allow for open-set experimental scenarios
- impostor attempts scenarios for the behavioral voice and handwriting biometrics

Another objective in the design of MyIdea is to complement existing mono and multi-modal biometric databases such as, for example, BANCA [1], BIOMET [9], XM2VTSDB [13], MCYT[16], IAM [12] ¹. Sensors, scenarios and contents of MyIdea, as well as the design of the acquisition protocol are crafted to allow as much complementarities with other existing databases as possible.

The overall MyIdea project is performed in the framework of a collaboration between the University of Fribourg in Switzerland [19], the Engineering School of Fribourg in Switzerland [8] and the GET in Paris [10]. In Fribourg, MyIdea is supported by the Swiss project IM2 [14]. Feedback and interest have also been raised from the EC Network of Excellence BioSecure [15].

¹The IAM database was initially designed for off-line handwriting recognition and has been recently used to perform *Identity Verification* experiments.

Modality	Sensors	Scenarios	Complement
Voice and face	Expensive and cheap camera	Controlled, degraded, adverse	XM2VTS, BANCA, Biomet
Fingerprints	Optical and sweep sensor	Controlled and uncontrolled	MCYT, FVC2000, Biomet
Palmprints	Scanner	Various resolution	Biomet
Hand-geometry	CCD camera	Lateral and dorsal views	Biomet
Handwriting	Graphical tablet	Signature and writing	IAM, Biomet
Voice and Handw.	Graphical tablet and mic	Signature and writing	-

Table 1: Summary of sensors, scenarios and potential complementarities with existing databases for each modality acquired in MyIdea.



Figure 1: Picture of the visual audio acquisition equipments. Controlled scenario.

2 Acquisition System

2.1 Video for voice and face

- **Controlled scenario.** Two digital cameras are used²: a good quality consumer market camcorder (Sony Digital Camera DCR-HC40) and a good quality web cam (Unibrain Fire-Wire i). The integrated microphone of the expensive camera is shunted by a high-end directional microphone (Shure SM94). The response curve of this microphone is very flat according to the specifications. The web cam microphone is shunted by a good quality computer microphone (Verse 704 from Labtec). Tripods are used for the high-end camera and microphone. The digital camera, the web cam and both microphones are positioned at the same place in the room, as illustrated on Figure 2.

The settings of the high-end digital camera are the following: all automatic settings turned off, manual and unique setting of the exposition for all the recordings. Audio is recorded at 48 KHz 16 bits/sample. The high-end digital camera and microphone signals are recorded on mini-DV tapes, video signal PAL, color sampling resolution 4:2:0 and audio stream synchronized on a track of the mini DV tape. The settings of the web cam are the followings: 30 frames/sec and a resolution of 320×240 . The video sequences acquired on the web cam are directly saved on the computer using a WMV compression at 768 kbps. The audio is taken from a Shure microphone placed in front of the subject and connected to an amplifier,

²For BANCA, only one high-end camera has been used with two microphones. We use here two cameras and two microphones. The high-end camera we use is similar to the high-end camera of BIOMET.

which directs the output to the computer. The sound is thus synchronously recorded with the video signal at 44 KHz and compressed using WMA-9 at 64 kbps.

The lighting conditions are controlled with two halogen lamps directed with the help of two photograph umbrellas in order to have an uniform illumination of the face. The umbrellas and the lights are positioned at the left and right of the cameras, slightly under the level of the face in order to avoid shadows on the face of the subjects. The illuminance of the face of the subjects has been measured to an average value of 650 lux. The illuminance should not fluctuate throughout the whole acquisition since the blinds are shut and the halogen lamps are the only source of light. A blue screen is placed on the wall behind the subject, in order to ease post-processing of the video. Subjects will sit on a chair at a controlled distance of the cameras. The sound acquisition conditions are also controlled: no noise in the room, door and windows shut. Just below the camera, the text to be read by the subject is hung on a plastic baguette itself hanging to the ceiling with threads. The text is printed on A4 paper to ease the reading.

- **Degraded scenario.** Two cameras are used: a high-end one identical to the one of the *controlled* scenario (Sony Digital Camera DCR-HC40) and a medium-quality web cam (Logitech Quickcam Messenger)³. Figure 2 illustrates the equipments used for the degraded scenario. The integrated microphone of the expensive camera is shunted by a high-end directional microphone (Shure SM94). For the web cam, its internal microphone is used. Tripods are used for the high-end camera and microphone.

The settings of the high-end digital camera are the default ones with the default automation of the camera turned on. The video sequences are recorded on mini DV tapes for the high-end digital camera and microphone (audio stream synchronized on a track of the DV tape). The settings of the web cam are the followings: 30 frames/sec and a resolution of 320×240 pixel/frame, for a bitrate of 768 kbps. The video sequences acquired on the web cam are directly saved on the computer with identical settings as for the web cam used for the controlled scenario.

The scenario is the one of an office environment. The lighting conditions are the one that would be used in an office, i.e. external natural light during the day and ceiling lights turned on when the external light is not enough. No special lighting is used to avoid shadows on the face of the subjects. The sound acquisition is also less controlled and fits with an office environment: potential office noise from the surrounding, potential corridor and street noise with door and windows open. Potentially, people can also be passing by in the background of the picture. Subjects sit in front of a computer screen on top of which is positioned the web cam. The high-quality digital camera is placed on a tripod, taking the pictures from above the computer screen, approximately at the same location as the web cam. The high-end microphone is placed beside the digital camera, also on a tripod. The text to be read by the subject is placed on the screen, on a A4 paper. One meter behind the subject's chair, there is the office door which may be closed or open.

- **Adverse scenario.** One high-end camera and microphone are used, identical to the one of the *degraded* and *controlled* scenarios⁴. They are placed on tripods.

The settings of the high-end digital camera are the default ones with the default automation of the camera turned on. The video sequences are recorded on mini DV tapes for the high-end digital camera and microphone (audio stream synchronized on a track of the DV tape).

The scenario is the one of a highly variable environment, similar to the one of, for example, a banking machine. The environment of the recording will be uncontrolled, potentially with many people passing by and babbling around the subject, potentially with elevator doors opening and closing in the background, etc. The lighting conditions will correspond to the inner hall of a public building, with varying natural and artificial lightings. The distance

³For BANCA, only one high-end camera has been used with two microphones for the degraded scenario. We use here two cameras and two microphones.

⁴For BANCA, only one web cam has been used for the adverse scenario. Since we wanted to have a portable recording tool, we use here one high-end camera and no web cams which are not portable.

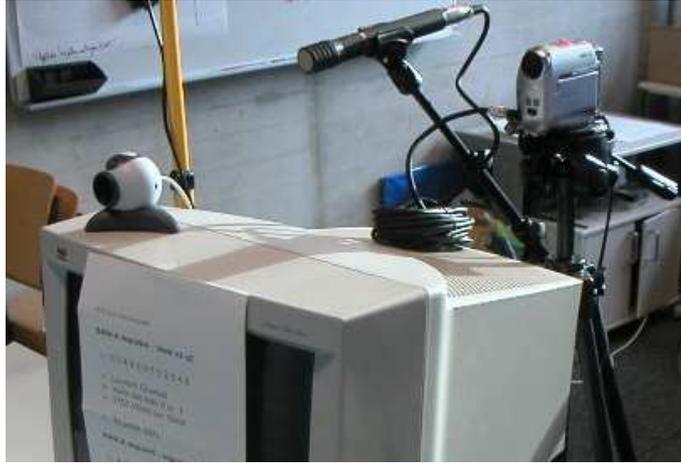


Figure 2: Picture of the visual audio acquisition equipments. Degraded scenario.



Figure 3: Pictures of the Ekey (left) and SAGEM (right) fingerprint sensors.

between the subject and the camera will not be controlled and we expect a variation of about one meter from one subject to another.

All the video hardware used for the recordings can be interfaced to a computer via a firewire or USB port.

2.2 Fingerprints

Two fingerprint sensors are used for the acquisition. Both sensors are driven by a software developed for the acquisition campaign (See section 3.2.1) but the constructor drivers are used to perform the fingerprint acquisition. Both sensors are illustrated on Figure 3.

The first one is the high-end optical sensor *Morphosmart MSO-100* from SAGEM [18]. Fingerprints are acquired at 500 dpi on a $21mm \times 21mm$ acquisition area, 8 bit grey scale. The glass plate of this sensor needs frequent cleaning to avoid additional noise to accumulate between acquisitions. Cleaning will be performed with alcohol-soaked napkins at a frequency defined according to the protocol (see section 4.3 of this document).



Figure 4: Picture of the EPSON scanner used for the acquisition of the hand images.

The second one is the *TocaBit* scan-thermal sensor from Ekey [3] which is based on a *FCD4B14 FingerChip* chip from Atmel [7]. A thermal fingerprint sensor based on this material measures the temperature differential between the sensor pixels that are in contact with the ridges and those under the valleys, that are not in contact. This sensor is a little more trickier to use than the first one. Provided that the fingertip has been swept across the sensor window at a reasonable rate, the overlap between successive frames enables an image of the entire fingerprint to be reconstructed [4]. This is done using software supplied by Atmel as part of the sensor deliverable. The reconstructed image is at 8 bit grey scale resolution. The reconstructed image is typically $25mm \times 14mm$, equivalent to 500×280 pixels. This sensor does not need any cleaning between acquisitions.

2.3 Palmprints

A scanner from EPSON (perfection 1660 photo) is used to scan palm images of the hand. The scanner is driven by the software developed for the acquisition campaign through the TWAIN driver (See section 3.2.1). The scanning resolution can be set to various values through this software. As for the optical fingerprint sensor, the glass plate of the scanner needs frequent cleaning to avoid additional noise to accumulate between acquisitions. Cleaning will also be performed with alcohol-soaked napkins at a frequency defined according to the protocol (see later sections of this document). Figure 4 shows a picture of the scanner.

2.4 Hand geometry

Hand geometry including the dorsum surface (top-view) and the lateral surface (side-view) of the hand is captured on a single picture with a CCD camera. For this, platform consisting of a wooden surface and two lateral mirrors has been built. Figure 5 and 6 illustrate the apparel. The dimensions of the platform are 30 cm x 18 cm, and the side mirrors are 30 cm x 7 cm each. The mirrors are placed along the wooden surface with an angle of approximately 50 (the angle was fine-tuned to get the best image possible according to the distance between the camera and the platform). The wooden surface is painted in blue, and its sides in green, in order to ease the postprocessing phase to isolate top and side views of the hand. At a height of about 30 cm up the wooden surface, an Olympus digital camera (CMEDIA digital camera C-370) is fixed on a tripod (a manfrotto 190CL/nk19) which allows to reverse the camera for shooting towards the ground. The camera itself uses macro mode setting in order to enhance the quality of the pictures ; pictures taken are of dimensions 2048x1536. On our platform, pegs are placed on the platform to guide the

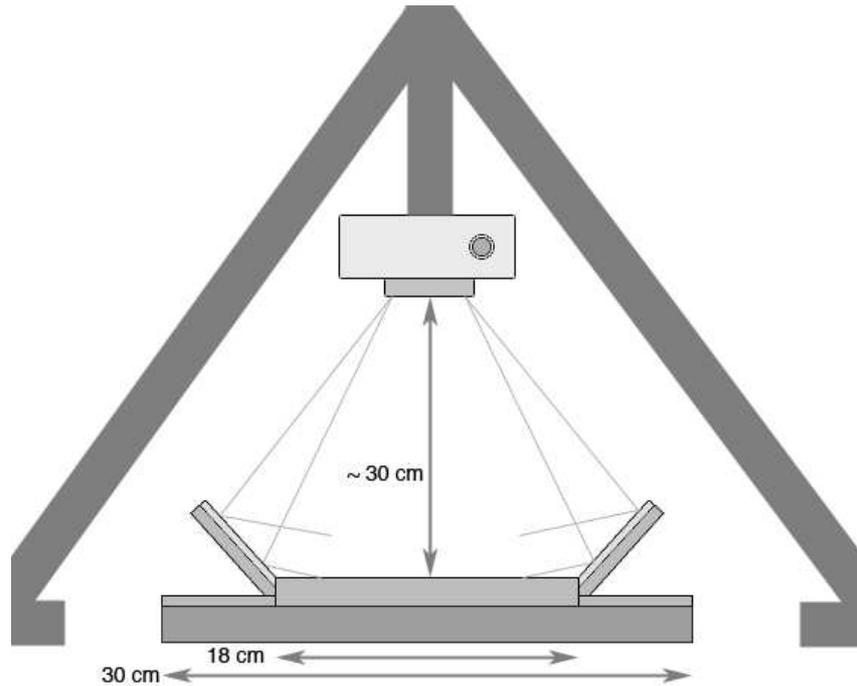


Figure 5: Schema of the hand sensor.

placement of the user's hand. In the literature, similar equipments have been built to capture the hand geometry (see for example [17]).

2.5 Signature and Handwriting

An A4 Intuos2 graphic tablet from WACOM [5] is used with an Ink pen allowing to write on standard paper positioned on the tablet. The use of an Ink Pen allows natural signatures and handwriting. Although on-line acquisition is the main focus for MyIdea, the Ink Pen used on standard paper also allows to perform off-line experiments. The tablet records 5 parameters: x-y coordinates, pressure, azimuth and altitude at a frequency of 100 Hz. The tablet is driven by a software developed for the acquisition campaign (See section 3.2.1) using the drivers provided by the constructor. Figure 7 illustrates the graphic tablet sensor. For the synchronized acquisition of writing and voice where the subject is asked to read aloud what (s)he is writing, a computer microphone mounted on a headset is used.

2.6 Room details

The biometrics lab is located on the third floor of the EIF building, Boulevard de Pérolles 80, Fribourg, Switzerland. Orientation of the unique window of the room is south west. The window is equipped with blinds and the room can be almost fully occulted when they are down. Ceiling lights consist of 3 neons. As classrooms are located on the same floor, the surrounding environment is rather quiet and silent except during break times.

We arranged the room so that subjects don't need to pay attention to the material. That is, all cables are laying on the left side on the room and subjects are only asked to move through the right side. Figure 8 illustrates the positions of the sensors in the room. On the right wall, in front of the audio-video controlled-scenario sensors, a blue paper tissue about 80 cm large and 2.5 meters high has been fixed on the wall. In front of it a chair is facing the camera, web cam and both microphones used for the controlled audio-video capture. Distance between the chair and the

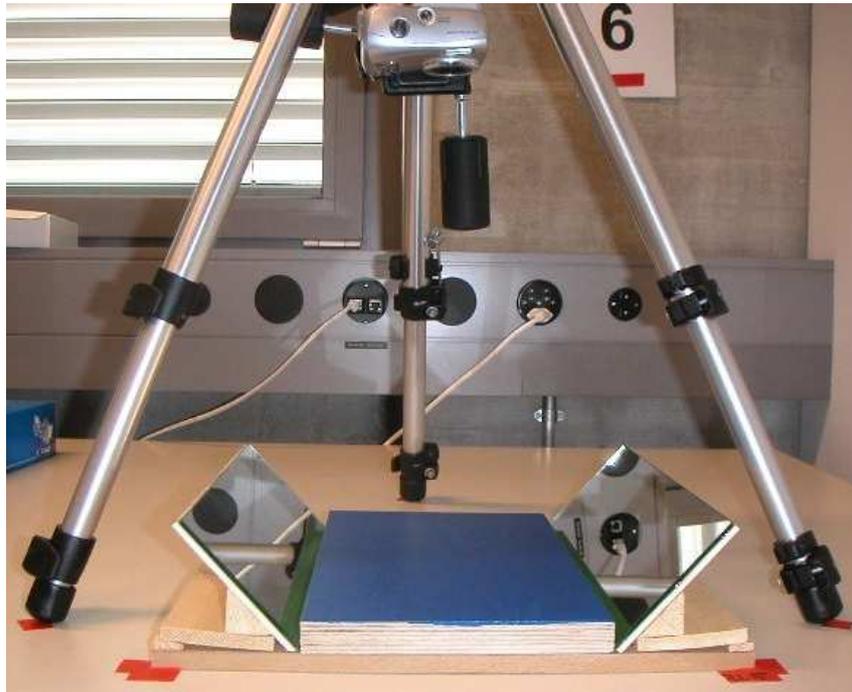


Figure 6: Picture of the hand sensor.



Figure 7: Picture of the A4 Intuos graphic tablet from WACOM used for handwriting and signature acquisition.

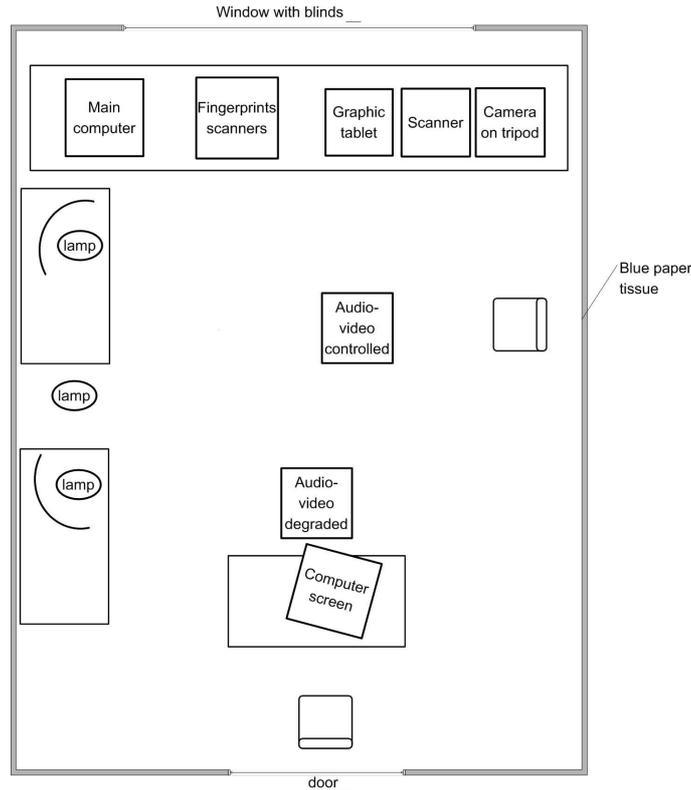


Figure 8: Map of MyIdea Biometric Lab

sensors is about 1 meter.

In front of the door a computer screen is placed on a table. On the screen is the web cam and behind the table the camera and the microphone used for the degraded capture. The client is sitting on a chair back to the door around 1 meter from the table. Text to be read is stuck on the screen.

Opposite to the door, next to the window is a large table on which are placed from left to right, the main computer, both capacitive and thermic fingerprint scanners, the graphic tablet, a scanner, and finally the camera on its tripod used for hand surface recognition as described below.

Both camera and microphone used for the adverse capture are kept on the left corner close to the door. The text to be read is placed just below the camera. Adverse captures are made in random places such as, for example the corridor, in front of the elevator, in front of the coffee machine etc.

3 Software

3.1 Protocol generation

A software has been built to generate automatically the protocols. The main idea is to limit the impact of human mistakes by generating a documented protocol for each subject and each session. The software is written in *Java* and controls *Velocity* templates to generate automatically \LaTeX files which are then compiled into pdf printable documents. The generated documents have actually two parts:

1. The first part is dedicated to the assistant. It includes a step-by-step guide through the acquisition for a given subject and a given session. The document is split into sections corresponding to the acquisitions with the different sensors. Each section summarizes different

points : (1) any preparation of hardware and software the assistant should perform prior to the acquisition, (2) the list of recommendations to give to the subject prior the acquisition, (3) the procedure to follow during the acquisition and (4) potential cleanup procedures after the acquisition. The text includes typically checkboxes in front of each operation to perform. The assistant crosses the checkbox as soon as (s)he has performed the operation. Some fields are also available for the assistant to include information that is not automatically recorded by the acquisition software such as, for example, the starting and ending time code of the video cameras for a given recording. The content recorded by the subject is also reproduced in this document, so that the assistant can easily spot invalid inputs given by the subject.

2. The second part is dedicated to the subject. It includes all the content that the subject will have to record, such as, for example, the set of phrases to record during the video acquisition or the text to write during the handwriting acquisition. The sections of this document which are used for the video acquisition are actually printed on A4 paper to ease the reading in front of the camera.

When a given subject has finished with all the acquisition, both documents are classified in a binder for later use during the validation and annotation phase.

3.2 Biometrics Acquisition

During the data acquisition phase, three softwares are mainly used. The only one of them which hasn't been designed specifically for the task is *Windows Movie Maker*, used to capture video from the different webcams. The other two softwares are *Biblios* and *SignReplay*. The first one, *Biblios*, is used to monitor the acquisition protocol and control the different devices used throughout the acquisition. *SignReplay* is used to show and replay written data acquired by *Biblios*, e.g. signatures. Those two softwares are detailed below.

3.2.1 Biblios

Biblios, a dedicated C++ software designed to perform multi-modal acquisition of biometric data with various sensors has been developed⁵. This software pilots the drivers of the different sensors and records the acquisition data in a file system database, guaranteeing that the files follow a uniform naming convention including the unique id of each subject in each session. This software allows to quickly reconfigure the acquisition procedures to match the protocol definitions such as generated by the software described in section 3.1 (see Figure 9).

Generally speaking, *Biblios* guides the assistant through all the steps involved in a biometric data acquisition session. A setup panel permits to match the settings of *Biblios* with the requirements of the protocol; then, after the user id and session id are entered, a window displays all the possible modalities (apart from video and hand geometry, which are taken independently). For all modalities, an on-screen live playback helps the assistant verify the quality of the acquired data (see Figure 10 for an example of fingerprint acquisition). *Biblios* is able to monitor the following devices : scanner by means of the TWAIN driver, graphic tablet, audio input microphone, Sagem and Atmel fingerprint sensors.

3.2.2 Signature replay

SignReplay is a Java software designed to play-back the on-line signature data of a given subject on the computer screen. The software plots the (x, y) coordinates as a function of time, while letting the user choose a play-back speed which can be equal or lower than the acquisition speed. This software is used when training impostors to imitate the signature based on the dynamics of the signal : with the help of *SignReplay*, the user can easily visualize the track followed by the pen, with accelerations and decelerations. *SignReplay* is also used to verify the integrity and the quality of acquired written data. Figure 11 shows a screenshot of *SignReplay*.

⁵Initially developed for the acquisition of BIOMET [9] and re-designed to match the specificities of MyIdea.

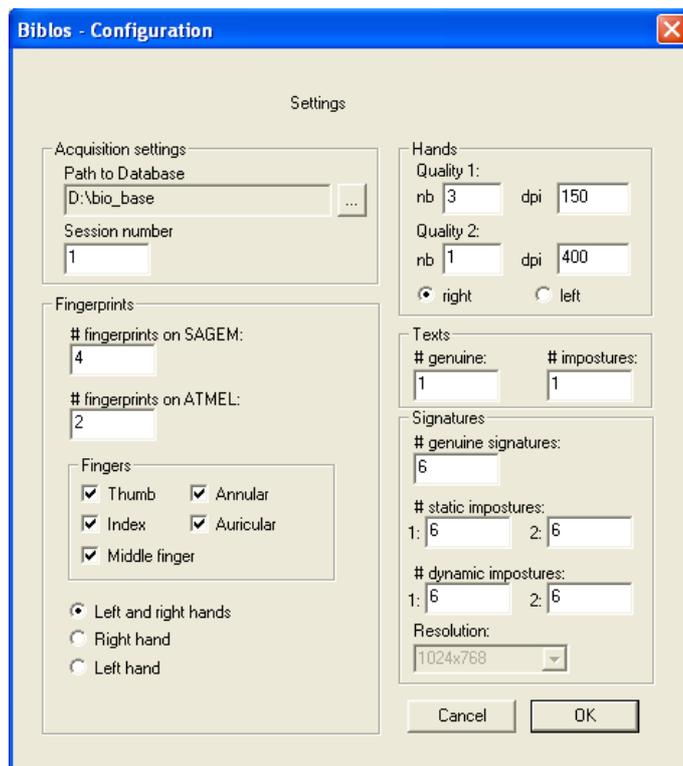


Figure 9: Setup screen of Biblos



Figure 10: Exemple of acquisition in Biblos : fingerprint

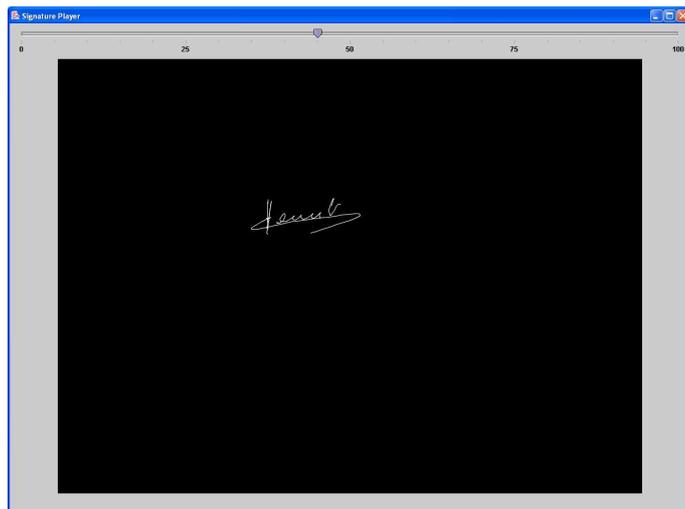


Figure 11: SignReplay in action

4 Acquisition protocol

4.1 Group constitution

The strategy for group constitution is similar to the one of the BANCA database [1] [2]. Each group has 13 subjects and is gender specific. The target is to reach 8 groups of 13 subjects, i.e. a total of 104 subjects who record their biometric data in 3 sessions⁶. Protocols have actually been generated for 208 subjects (16 groups of 13 subjects), allowing for more data to be recorded if time is left. Sessions are spaced in time of several weeks. However, no special procedure has been put in place to control a minimum period of time between two successive sessions. Groups are denoted in the following $mg1, mg2, \dots, mg8$ for the male population and $fg1, fg2, \dots, fg8$ for the female population.

We use the following notations:

- X_i^g : subject i in group g , $g \in \{mg1, \dots, mg8, fg1, \dots, fg8\}$, $i \in [1, 13]$
- $y_s^m(X_i^g)$: true client record from session s using condition/modality m by subject X_i , $s \in [1, 3]$, $i \in [1, 13]$
- $z_t^n(X_i^g, X_j^g)$: impostor record from subject X_j claiming identity of subject X_i during a session t using condition/modality n , $t \in [1, 3]$, $i \in [1, 13]$, $j \in [1, 13]$, $i \neq j$

Impostor attempts are recorded for behavioral modalities **voice** and **handwriting** only. Impostor attempts are implemented by using subjects from the same group as the true client. In this way, incomplete groups which are due to subjects who have not taken part to the whole set of recording sessions can simply be removed or used to populate a separate set of *world* data.

4.2 Video for voice and face

Voice and face are acquired using different cameras in three conditions: controlled, degraded and adverse such as described in section 2.1. Two contents are acquired in order to complement BANCA and BIOMET databases. Table 2 summarizes the content and conditions for these databases. Our protocol also includes acquisition of head rotation shots, allowing to complement with XM2VTS database [13].

⁶The acquisition which is similar to the one of the BANCA database is actually performed in 4 sessions

	BANCA	BIOMET
Controlled	X	X
Degraded	X	-
Adverse	X	-

Table 2: Conditions of video acquisition for the contents similar to BANCA and BIOMET

Before each video shot is recorded, a sequence equivalent to a clipperboard is taken in order to uniquely identify that shot. The clipperboard contains the subject unique identification number, the sequence type and session number. The clipperboard also contains a color test chart and a resolution/distorsion test chart ISO 12233. The clipperboard sequence allow to resolve any potential errors and to check that the recordings are consistent across the whole database.

Subjects received the following recommendations prior each recording :

- **Controlled.** Sit in front of the camera. Read one time the text before the recording. Do not remove glasses or jewels of any kind. Recording is stopped and performed again in case of important fluctuation of the recording conditions or in case of missed content.
- **Degraded.** Sit in front of the computer screen. Read one time the text before the recording. Dress as if at work (take off coat for example). Use (or do not use) glasses as if at work. Do not stop recording in case of fluctuation of the conditions (for example, uncontrolled background noise or passage).
- **Adverse.** Stand in front of the camera which is located at a random place (corridor, hall, cafeteria, ...). Dress as if in the street or in a hall, according to weather conditions, i.e. dressing can show large variabilities. Do not stop recording in case of fluctuation of the conditions, even important.

4.2.1 Content 1 - similar to BANCA database

Remarks: According to [1], BANCA has been recorded in 12 sessions with sessions 1 to 4 for the controlled scenario (C), sessions 5 to 8 for the degraded scenario (D) and sessions 9 to 12 for the adverse scenario (A), i.e. 4 sessions per scenario. In MyIdea, we record in one session each of the scenarios, i.e. controlled, adverse and degraded. In our session 3, we record two times the BANCA controlled, degraded and adverse scenarios to reach the corresponding 12 sessions of BANCA. The following table compares BANCA and MyIdea recordings.

Scenario	C	C	C	C	D	D	D	D	A	A	A	A
BANCA Session	1	2	3	4	5	6	7	8	9	10	11	12
MyIdea Session	1	2	3	3	1	2	3	3	1	2	3	3

During each session, the subject is prompted to record two sets of information :

- **True client information** : a random 12 digit number, his/her name, address and date of birth
- **Impostor attack information** : the true client information of another person in the group

For each session, the true client information remains the same. For different sessions the impostor attack information changed to another person in their group (hence 12 attacks in total). The sequence of impostor attacks is not randomly generated. Instead, it is designed so as to make sure that each identity is attacked 4 times in the 3 different conditions, which is similar to what is described in [1].

More formally, we have:

X_i^g	:	subject i in group g , $g \in \{mg1, \dots, mg8, fg1, \dots, fg8\}$, $i \in [1, 13]$
$y_s^m(X_i^g)$:	true client record from session s using condition m , by subject X_i of group g , $s \in [1, 4]$, $i \in [1, 13]$, $m \in \{\text{controlled, degraded, adverse}\}$
$z_t^n(X_i^g, X_j^g)$:	impostor record from subject X_j of group g claiming identity of subject X_i of the same group g during a session t and using condition n , $t \in [1, 4]$, $i \in [1, 13]$, $j \in [1, 13]$, $i \neq j$, $n \in \{\text{controlled, degraded, adverse}\}$

For example, the set of 12 impostor attempts recorded by subject X_1 in a given group g is $\{z_1^c(X_2, X_1), z_2^c(X_3, X_1), z_3^c(X_4, X_1), z_4^c(X_5, X_1), z_1^d(X_6, X_1), z_2^d(X_7, X_1), \dots, z_4^a(X_{13}, X_1)\}$ where c, d, a stands for *controlled, degraded* and *adverse*. The set of 12 impostor attempts recorded by subject X_2 is $\{z_1^c(X_3, X_2), z_2^c(X_4, X_2), z_3^c(X_5, X_2), z_4^c(X_6, X_2), z_1^d(X_7, X_2), z_2^d(X_8, X_2), \dots, z_4^a(X_{13}, X_2)\}$, etc. The impostors do not try to imitate the voice or the facial items and gestures of the true client. We are in a scenario of *informed* impostor attack, i.e. the impostor has been informed about the data of the claimed identity.

4.2.2 Content 2 - similar to Biomet database

1. **Counting from 0 to 9** The ten digits are said with no pause between the digits.
2. **Counting from 9 to 0** The ten digits are said with no pause between the digits.
3. **Two long fixed phrases** Two phonetically balanced phrases are said. These phrases are the same for each subjects and for each session.

phrase 1 - fr Alors que Monsieur Gorbatchev regagnait Moscou au terme d'un difficile voyage en Lituanie, une partie du Caucase s'est embrasée.

phrase 2 - fr Chaque jour ils reçoivent, dans la bonne humeur, la visite du commissaire des renseignements généraux, qui suit de loin l'opération.

4. **yes/no** *Yes* and *no* are said.
5. **Ten short fixed phrases** Ten short phrases are said. These phrases are the same for each subjects and for each session.

phrase 1 - fr Il se garantira du froid avec un bon capuchon.

phrase 2 - fr Annie s'ennuie loin de mes parents.

phrase 3 - fr Les deux camions se sont heurtés de face.

phrase 4 - fr Un loup s'est jeté immédiatement sur la petite chèvre.

phrase 5 - fr Dès que le tambour bat, les gens accourent.

phrase 6 - fr Mon père m'a donné l'autorisation.

phrase 7 - fr Vous poussez des cris de colère.

phrase 8 - fr Ce petit canard apprend à nager.

phrase 9 - fr La voiture s'est arrêtée au feu rouge.

phrase 10 - fr La vaisselle propre est mise sur l'évier.

6. **4 true client password phrases** These phrases are the same for each session but are different from one subject to the other (picked from a set of 104 subjects *4 phrases = 416 phrases). These phrases have a limited size between 20 and 45 chars to simulate a kind of password-based phrase scenario. The phrase corpus has been built from phrases taken in [6]⁷ and completed with short phrases taken from the BREF database [11] and from articles in newspapers Le Monde, Le Temps and La Liberté to reach a total of 832 phrases.
7. **4 impostor password phrases** These phrases correspond to true client password phrases of 4 different subjects in the same group. As we have three sessions, this leads to 104 subjects *3 sessions *4 phrases = 1248 impostor attempts. The sequence of impostor attacks is not randomly generated. Instead, it is designed so as to make sure that each subject attacks the 12 other subjects of the same group. For example, the set of 12 phrases used by subject X_1 in a given group g is $\{P_1(X_2), P_2(X_3), P_3(X_4), P_4(X_5), P_1(X_6), P_2(X_7), \dots, P_4(X_{13})\}$ where $P_i(X_j)$ stands for phrase i of subject X_j with $i \in [1, 4]$. The set of 12 phrases used by subject X_2 is $\{P_1(X_3), P_2(X_4), P_3(X_5), P_4(X_6), P_1(X_7), \dots, P_3(X_{13}), P_4(X_1)\}$, etc.
8. **5 random short phrases** These phrases are different for each session and for each subject (picked from a set of $104 * 3 * 5 = 1560$ phrases). The phrases were generated from the content of the BREF database [11]. As the phrases of BREF show a large variability of length, phrases having less than 50 and more than 90 characters were removed. Phrases having a double quoted sub-phrase, i.e. a quotation of someone else speech, were also removed. The phrases were also proofread by humans in order to remove phrases with no sense, to correct grammars or orthographic errors and to normalize the format.
9. **head rotation shots** The subject is asked to turn the head roughly 30 to the left, to the right, up, down and finally 90 to the left and right to capture his/her left and right profiles. Marks have been put on the wall, ceiling and floor to indicate the user the sequence to follow and to avoid too large variabilities in the positions of each subjects. The profile shots (45) are performed asking the subject to fully rotate his/her body on the chair while the 30 shots are obtained with a simple rotation of the head. These images can be used for profile or 3D based authentication.

4.3 Fingerprints

Two sensors are used to capture the fingerprints as described in section 2.2 of this document: an optical device and a scanning thermal device. Subjects received the following global recommendations prior each recording :

- stand up in front of the sensors
- do not move the body during the acquisition
- do not over-press on the sensors for the acquisition

The content per session is as follows:

Optical	uncontrolled	5 fingers, left-right hands	2 acquisitions
	controlled	5 fingers, left-right hands	2 acquisitions
Scan thermal	uncontrolled	5 fingers, left-right hands	2 acquisitions

For the optical sensor, the uncontrolled mode corresponds to a nominal use of the sensor, without controlling the quality and centering of the acquired image, without taking any precaution about previous fingerprint traces that could remain on the sensor glass and without preparing the

⁷Phrases with inherent difficulties to pronounce may have been modified or removed. The objective was to get a corpus which is phonetically rich but not difficult to pronounce.

fingers of the subjects. The controlled mode is performed by visually controlling the quality and centering of the acquired image. To enhance the acquisition of the controlled mode, the glass plate of the sensor is cleaned with alcohol-soaked napkins between each acquisition of each finger. Also, it has been observed that too dry fingers lead to degraded images using this sensor. Therefore, for the controlled mode, subjects are also asked to wipe their finger on their forehead prior each acquisition to somehow smooth out the dryness.

The scan thermal sensor does not need any cleaning and is supposed to be independent to the conditions of the fingers. Therefore, only the uncontrolled mode has been planned in the protocol. However, it has to be underlined that the acquisition driver may reject automatically the capture when the scanned image can not be reconstructed. The reasons for such a reject are not well documented. We have observed such rejects when fingers are swept too slow or too fast or when the temperature of the finger is too close to the temperature of the device.

For each session, 40 acquisitions are performed on the optical sensor and 20 on the scan-thermal sensor, leading to a total of 60 acquisitions per subject after three recording sessions.

4.4 Palmprint scanning

A simple scanner is used to capture the palmprints as described in section 2.3 of this document. Subjects received the following global recommendations prior each recording:

- stand while scanning is performed
- spread the fingers naturally on the glass plate
- put the right hand as flat as possible on the glass plate but without pressing too much
- do not remove rings
- try not to move during the scanning

The glass plate of the scanner is cleaned with alcohol-soaked napkins one time prior the acquisitions of a specific subject. The blinds of the room are shut and the lights are switched off to avoid too large variabilities of the illumination. For each session, the right hand of the subject is scanned 4 times, 3 times at a resolution of 150 dpi and 1 time at a resolution of 400 dpi. This leads to a total of 12 acquisitions per subject after three recording sessions.

4.5 Hand Geometry

An equipment composed of a platform and a CCD digital camera is used to capture the right hand geometry with dorsal and lateral views, such as described in section 2.4 of this document. Only the right hand geometry is recorded here. Subjects received the following global recommendations prior to each recording:

- stand and place the right hand and fingers according to pegs positions on the platform
- put the hand as flat as possible on the platform but without pressing too much
- do not remove rings

The blinds of the room are shut to avoid too large variabilities of the illumination and the platform is illuminated with the ceiling lights. The CCD camera is running in full automatic mode and we let the camera decide whether to turn on or off the flash. For each session, the right hand of the subject is captured 3 times at the nominal resolution of the CCD camera. This leads to a total of 9 acquisitions per subject after three recording sessions.

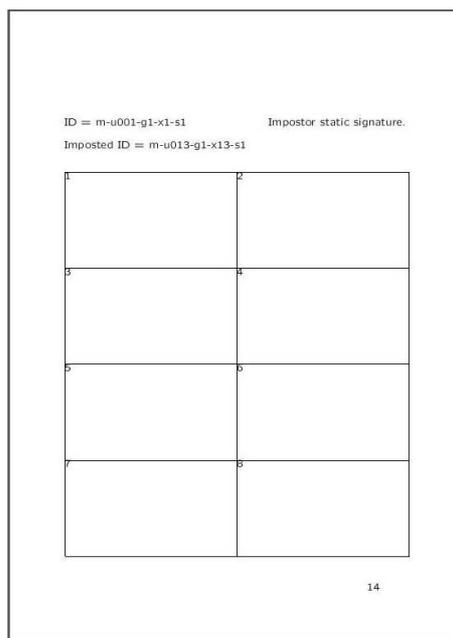


Figure 12: Example of a signature page used for the impostor static signature scenario. The subject id and imitated id are shown on the top of the page. Dimensions of the cells are 80x45 mm. These pages are automatically generated by by MyIdea generation software using \LaTeX .

4.6 Signature, signature and voice

A graphic tablet and an Ink Pen are used to capture signature and handwriting such as described in section 2.5. On-line acquisition is the main focus of MyIdea for the signature modality. However, one has to underline that the settings allow for on-line and off-line experiments. Indeed, the Ink Pen records on-line data while regular paper sheets are used to write, allowing off-line scanning of the handwriting.

The layout of the paper shows 8 numbered cells such as shown on figure 12. As defined in the protocols, subjects will sign 6 times. The 2 remaining cells are used in the case some signature would be declared as missed (see later in the text for a definition of a missed signature). The papers are uniquely identified by the subject id in the upper left corner and a title corresponding to the scenario in the upper right corner. Subjects received the following global recommendations prior each recording:

- do not sign outside of the cell
- wait for the top of the assistant, take the Ink Pen, perform the signature and put back the Ink Pen in its cradle
- turn the tablet and place fingers and hands to reach a comfortable position

4.6.1 True signature

For each session, the subject is asked to perform 6 signatures leading to a total of 18 true signatures after three sessions. The subject is asked to train for a few signatures on a separated sheet in order to accustom the hand to the Ink Pen. During the acquisition, a missed signature can be re-done if the subject declares that his/her signature is missed, if the signature goes past the borders of the cells, if there is a problem with the Ink Pen (empty of malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

4.6.2 True signature with voice

The subject is asked to synchronously sign and utter the content of his/her signature. Invariant synchronization is of course not possible, however, the subject is asked to sign in such a way that the written symbols correspond roughly in time with the uttered phonemes. If the signature is far away from any logical representation of the subject's name, the subject is asked to simply utter his/her name during the signature. For each session, the subject is asked to perform 6 signatures and voice inputs leading to a total of 18 true acquisitions after three sessions. The subject is asked to train for a few signatures and voice inputs on a separated sheet in order to get used to the procedure. During the acquisition, a missed signature can be re-done if the subject declares that his/her signature is missed, if the signature goes past the borders of the cells, if there is a problem with the Ink Pen (empty of malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

As this particular acquisition is critical from a security point of view (the identity of the writer may be more easily discovered by listening to the voice part of the acquisition), the subject may decline to utter his/her signature. In such a case, the acquisition is still performed, but without the voice part.

4.6.3 Impostor static signature

For each session, the subject X_i^g is asked to imitate the signature of another subjects X_j^g on the basis of his/her acquired **static** data. A sheet such as illustrated on figure 12 including 6 valid signatures of subject X_j^g is presented to the subject. The subject has a limited time of two minutes to train on a separate sheet to imitated the signatures. More formally, for each session, subject X_i^g is asked to perform 6 signatures of another subjects X_j^g in the same group with $j = i - s$, s indicating the session index ($s \in [1, 3]$). As groups have 13 subjects, the following rotation rule is applied: if $j \leq 0$ then $j = 13 + j$. This procedure leads to a total of 18 impostor signatures on 3 different subjects after the three sessions. During the acquisition, a missed signature can be re-done only if the signature goes past the borders of the cells, if there is a problem with the Ink Pen (empty of malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

4.6.4 Impostor dynamic signature

The subject X_i^g is asked to imitate the signature of another subject X_j^g on the basis of his/her acquired **static** and **dynamic** data. First, a sheet such as illustrated on figure 12 including 6 valid signatures of subject X_j^g is presented to the subject. Second, the on-line acquisition of the same signatures are replayed on a computer screen. The subject can decide to play repeatedly the signature at various speed to get the details of the dynamics. As in the previous impostor attempts, the subject has a limited time of two minutes to train on a separate sheet to imitated the signature of subject X_j^g . More formally, for each session, subject X_i^g is asked to perform 6 signatures of another subject X_j^g in the same group with $j = i - s$, s indicating the session index ($s \in [1, 3]$). As groups have 13 subjects, the following rotation rule is applied: if $j \leq 0$ then $j = 13 + j$. This procedure leads to a total of 18 impostor signatures on 3 different subjects after the three sessions. For a subject X_i^g , the set of 3 imitated subjects X_j^g is included to the set of imitated subjects of impostor static signature scenario. During the acquisition, a missed signature can be re-done only if the signature goes past the borders of the cells, if there is a problem with the Ink Pen (empty of malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

4.6.5 Impostor static signature with voice

For each session, the subject X_i^g is asked to imitate the signature of another subjects X_j^g and to synchronously utter the content of the signature. Subject X_i^g has access to the **static** signature data of subject X_j^g and to the verbal **content** of what subject X_j^g has said during a former acquisition. In other words, access to the recording is not given. A sheet such as illustrated on figure 12 including 6 valid signatures of subject X_j^g and the verbal content of the signature is

	session 1	session 2	session 3
1. True signature	$y_1^s(X_i^g)$	$y_2^s(X_i^g)$	$y_3^s(X_i^g)$
2. True signature and voice	$y_1^{s+v}(X_i^g)$	$y_2^{s+v}(X_i^g)$	$y_3^{s+v}(X_i^g)$
3. Impostor static signature	$z_1^{ss}(X_{i-1}^g, X_i^g)$	$z_2^{ss}(X_{i-2}^g, X_i^g)$	$z_3^{ss}(X_{i-3}^g, X_i^g)$
4. Impostor dynamic signature	$z_1^{ds}(X_{i-1}^g, X_i^g)$	$z_2^{ds}(X_{i-2}^g, X_i^g)$	$z_3^{ds}(X_{i-3}^g, X_i^g)$
5. Impostor static signature and voice	$z_1^{ss+v}(X_{i-4}^g, X_i^g)$	$z_2^{ss+v}(X_{i-5}^g, X_i^g)$	$z_3^{ss+v}(X_{i-6}^g, X_i^g)$

Table 3: Sequence of signatures acquisition and true/impostor indexes per session for a given subject X_i^g .

presented to the subject. The subject has a limited time of two minutes to train on a separate sheet to imitated the signatures. More formally, for each session, subject X_i^g is asked to perform 6 signatures of another subjects X_j^g in the same group with $j = i - s$, s indicating the session index ($s \in [1, 3]$). As groups have 13 subjects, the following rotation rule is applied: if $j \leq 0$ then $j = 13 + j$. This procedure leads to a total of 18 impostor signatures on 3 different subjects after the three sessions. During the acquisition, a missed signature can be re-done only if the signature goes past the borders of the cells, if there is a problem with the Ink Pen (empty of malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

Table 3 shows the sequence of signatures acquisition and summarizes the true-impostor indexes per session for a given subject X_i^g .

4.7 Handwriting with voice

A graphic tablet and an Ink Pen are used to capture handwriting such as described in section 2.5. A simple microphone mounted on a headset is used to capture the voice. On-line acquisition is the main focus of MyIdea for handwriting modality. However, one has to underline that the settings allow for on-line and off-line experiments. Indeed, the Ink Pen records on-line data while regular paper sheets are used to write, allowing off-line scanning of the handwriting.

The layout of the paper forms used for guiding the acquisitions is shown on figure 13. The forms were automatically generated in L^AT_EX using MyIdea protocol generation software (see section 3.1). This procedure guarantees that all forms are processed and generated in the same way. The layout of the form is inspired from the IAM database [12]. The form consists of four parts. The first part comprises the title such as *True handwriting* or *Impostor Handwriting*, the unique id of the subject/session and a reference assigned to the text. This reference shows from which category the text belongs to and allows to identify the document source from which the text has been retrieved. For example, D_2_402_77-1 indicates that the text of the form is extracted from text 402 in the text category *D-Education* and sub-category *2-science*. The next digits of the reference indicates starting line and word indices in the source text. In the second part of the form, the text the individual persons are asked to write is printed. The second part of the form is separated from the first and third part by a horizontal line. The third part of the form is a blank zone where the writers have to put in their handwriting. The blank zone has a constant vertical size of 17 cm, which has been experienced to be large enough for most of the handwriting.

As the main focus of MyIdea is the use of handwriting data to perform identity verification, we wanted to make image preprocessing as easy as possible. Therefore, we decided that the writers have to use rulers. These guiding lines are 15 mm spaced and are printed on a separate sheet of paper which is put under the form.

The sentences of each text fragment are extracted from a French corpus including various text categories such as listed in table 4. The text fragments contain between 50 and 100 words each and were sampled uniformly from the main text categories.

Subjects received the following global recommendations prior to each recording:

- use the every day writing in order to get the most natural and unconstrained way of writing
- use the punctuation as in the reference text
- write if possible in one shot without going back on potential mistakes

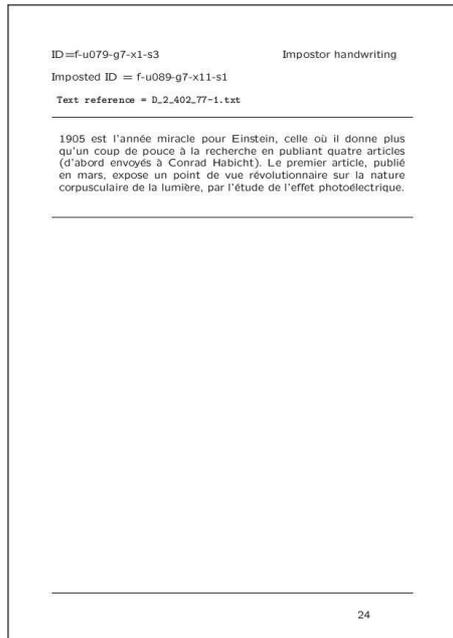


Figure 13: Example of a handwriting acquisition page used for the impostor handwriting scenario. The subject id and imitated id are shown on the top of the page, as well as the original text reference. These pages are automatically generated by MyIdea generation software using L^AT_EX.

A	press	
	1	international
	2	national
	3	regional
	4	miscellaneous
B	fiction	
	1	short story
	2	crime
	3	S-F
C	religion	
D	education	
	1	literature
	2	science
E	letters	
F	hobby	

Table 4: Examples of text categories and sub-categories in the corpus used for handwriting acquisition.

	session 1	session 2	session 3
1. True handwriting and voice	$y_1^{h+v}(X_i^g)$	$y_2^{h+v}(X_i^g)$	$y_3^{h+v}(X_i^g)$
2. Impostor static handwriting with voice	$z_1^{h+v}(X_{i-1}^g, X_i^g)$	$z_2^{h+v}(X_{i-2}^g, X_i^g)$	$z_3^{h+v}(X_{i-3}^g, X_i^g)$

Table 5: Sequence of handwriting acquisition and true/impostor indexes per session for a given subject X_i^g .

- stop to write if there is not enough space left to write the whole text (in order to avoid to get pressed and deformed words)
- turn the tablet and place fingers and hands to reach a comfortable position
- wait for the top of the assistant, take the Ink Pen, perform the signature and put back the Ink Pen in its cradle

4.7.1 True handwriting with voice

The subject is asked to write one fixed phrase which contains all the letter of the alphabet⁸ ("pangram") similar from session to session and one text fragment different from session to session. This leads to a total of 3 true handwriting acquisitions per subject. The subject is asked to synchronously write and utter the content of the text he/she is writing. The subject is asked to train for a few lines on a separated sheet in order to accustom the hand to the Ink Pen and to accustom to talking and writing in the same time. An acquisition can be re-done if the subject goes past the borders of the page, if there is a problem with the Ink Pen (empty or malfunctioning), if he/she performs too many mistakes or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

4.7.2 Impostor static handwriting with voice

For each session, the subject X_i^g is asked to imitate the handwriting of another subjects X_j^g and to synchronously utter the content of the text. Subject X_i^g has access to the **static** handwriting data of subject X_j^g but does not have access to the voice recording of subject X_j^g . A sheet such as illustrated on figure 13 including the handwriting of subject X_j^g is presented to the subject. The subject has a limited time of two minutes to train on a separate sheet to imitated the handwriting while uttering the content of the text. The same subjects as described in the previous section are imitated. More formally, for each session, subject X_i^g is asked to perform the imitation of the handwriting of another subjects X_j^g in the same group with $j = i - s$, s indicating the session index ($s \in [1, 3]$). As groups have 13 subjects, the following rotation rule is applied: if $j \leq 0$ then $j = 13 + j$.

As a part of the text fragment to write is different from session to session and as users are not sequentially recorded for session 2 and 3, impostor attempts are performed using the text fragment of the first acquisition session (and not of the second and third acquisition session). This procedure has the advantage to avoid that a user in session has no data to impost for the handwriting.

This procedure leads to a total of 3 impostor attempts on 3 different subjects after the three sessions. During the acquisition, a missed imitation can be re-done only if the handwriting goes past the borders of the cells, if there is a problem with the Ink Pen (empty or malfunctioning) or if the page which is supposed to be fixed on the graphic tablet moves for some reasons.

Table 5 shows the sequence of handwriting acquisition and summarizes the true-impostor indexes per session for a given subject X_i^g .

5 Acknowledgments

We are very grateful to the following people for the precious comments they provided during the definition of MyIdea protocol and the settings of the sensors: Dr. Samy Bengio, Prof. Horst

⁸In French: *Portez ce vieux whisky au juge blond qui fume.*

6 Revision

Last known revision information:

\$Id: MyIdea-sensors-protocol-final.tex 328 2005-06-14 09:33:47Z henneber \$

List of Figures

1	Picture of the visual audio acquisition equipments. Controlled scenario.	2
2	Picture of the visual audio acquisition equipments. Degraded scenario.	4
3	Pictures of the Ekey (left) and SAGEM (right) fingerprint sensors.	4
4	Picture of the EPSON scanner used for the acquisition of the hand images.	5
5	Schema of the hand sensor.	6
6	Picture of the hand sensor.	7
7	Picture of the A4 Intuos graphic tablet from WACOM used for handwriting and signature acquisition.	7
8	Map of MyIdea Biometric Lab	8
9	Setup screen of Biblios	10
10	Exemple of acquisition in Biblios : fingerprint	10
11	SignReplay in action	11
12	Example of a signature page used for the impostor static signature scenario. The subject id and imitated id are shown on the top of the page. Dimensions of the cells are 80x45 mm. These pages are automatically generated by by MyIdea generation software using L ^A T _E X.	16
13	Example of a handwriting acquisition page used for the impostor handwriting scenario. The subject id and imitated id are shown on the top of the page, as well as the original text reference. These pages are automatically generated by MyIdea generation software using L ^A T _E X.	19

List of Tables

1	Summary of sensors, scenarios and potential complementarities with existing databases for each modality acquired in MyIdea.	2
2	Conditions of video acquisition for the contents similar to BANCA and BIOMET	12
3	Sequence of signatures acquisition and true/impostor indexes per session for a given subject X_i^g	18
4	Examples of text categories and sub-categories in the corpus used for handwriting acquisition.	19
5	Sequence of handwriting acquisition and true/impostor indexes per session for a given subject X_i^g	20

References

[1] E. Bailly-Baillièrè, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, and J.-P. Thiran. The BANCA database and evaluation protocol. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA*. Springer-Verlag, 2003.

[2] S. Bengio, F. Bimbot, J. Mariéthoz, V. Popovici, F. Porée, E. Bailly-Baillièrè, G. Matas, and B. Ruiz. Experimental protocol on the BANCA database. IDIAP-RR 5, IDIAP, 2002.

- [3] Ekey biometric systems GmbH. <http://www.ekey.net/>.
- [4] Peter Bishop. Atmel's fingerchip technology for biometric security. Technical report, Atmel Corporation, 2002.
- [5] Wacom Technology Co. <http://www.wacom.com>.
- [6] P. Combescure. 20 listes de dix phrases phontiquement quilibres. *Revue d'Acoustique*, 56:34–38, 1981.
- [7] Atmel Corporation. <http://www.atmel-grenoble.com>.
- [8] Hochschule für Technik un Architektur Freiburg Ecole d'Ingénieurs et d'architectes de Fribourg. <http://www.eif.ch>.
- [9] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. Leroux les Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacrtaz. Biomet: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In UK University of Surrey, Guildford, editor, *4th International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA)*. Springer-Verlag, 2003.
- [10] Paris Groupe des Ecoles des Télécommunications. <http://www.get-telecom.fr>.
- [11] L.F. Lamel, J.-L. Gauvain, and M. Eskénazi. Bref, a large vocabulary spoken corpus for french. In *Proceedings of the European Conference on Speech Technology, EuroSpeech*, pages 505–508, Genoa, September 1991.
- [12] U.-V. Marti and H. Bunke. A full english sentence database for off-line handwriting recognition. In *Proc. of the 5th Int. Conf. on Document Analysis and Recognition (ICDAR'99)*, pages 705–708, 1999.
- [13] K. Messer, J. Matas, J. Kittler, J. Luetttin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In *Proc. Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA '99)*, 1999.
- [14] Swiss National Center of Competence in Research (NCCR) on Interactive Multimodal Information Management. <http://www.im2.ch>.
- [15] BioSecure Network of Excellence. <http://www.biosecure.info>.
- [16] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.-J. Igarza, C. Vivaracho, D. Escudero, and Q.-I. Moro. Mcyt baseline corpus: a bimodal biometric database. *IEE Proc.-Vis. Image Signal Process.*, 150(6):395–401, December 2003.
- [17] Nikola Pavešić, Slobodan Ribarić, and Dami Ribarić. Personal authentication using hand-geometry and palmprint features - the state of the art. In Claus Vielhauer, editor, *Biometrics: challenges arising from theory to practice*, pages 17–26, 2004.
- [18] Sagem S.A. <http://www.sagem.com>.
- [19] Universität Freiburg Université de Fribourg. <http://www.unifr.ch>.