MyIDea-multimodal biometrics database, description of acquisition protocols

Dumas, Bruno; Pugin, Catherine; Hennebert, Jean; Petrovska-Delacrétaz, Dijana; Humm, Andreas; Evéquoz, Florian; Ingold, Rolf; Von Rotz, Didier

Publication date:
2005

Document Version
Peer reviewed version

Link to publication

Citation for pulished version (HARVARD):
Dumas, B, Pugin, C, Hennebert, J, Petrovska-Delacrétaz, D, Humm, A, Evéquoz, F, Ingold, R & Von Rotz, D 2005, 'MyIDea-multimodal biometrics database, description of acquisition protocols', Paper presented at Proc. of Third COST 275 Workshop, Hatfield, United Kingdom, 27/10/05 - 28/10/05 pp. 59-62.

MYIDEA – MULTIMODAL BIOMETRICS DATABASE, DESCRIPTION OF ACQUISITION PROTOCOLS

Authors: Bruno Dumas[1], Catherine Pugin[1], Jean Hennebert[1], Dijana Petrovska-Delacrétaz[2], Andreas Humm[1], Florian Evéquoz[1], Rolf Ingold[1], and Didier Von Rotz[3]

[1] DIVA Group, Department of Informatics, University of Fribourg
Ch. Du Musée 3, 1700 Fribourg, Switzerland
[2] INT, Dept. EPH, Intermedia, 9 rue Charles Fourier 91011 Evry, France
[3] EIF, Bd de Pérolles 80 – CP 32, 1705 Fribourg, Switzerland

ABSTRACT

This document describes the acquisition protocols of MyIDea, a new large and realistic multimodal biometric database designed to conduct research experiments in *Identity Verification (IV)*. The key points of MyIDea are threefold: (1) it is strongly multimodal; (2) it implements realistic scenarios in an open-set framework; (3) it uses sensors of different quality to record most of the modalities. The combination of these three 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 BANCA, MCYT or BIOMET. MyIDea includes talking face, audio, fingerprints, signature, handwriting and hand geometry. MyIDea will be available early 2006 with an initial set of 104 subjects recorded over three sessions. Other recording sets will be potentially planned in 2006.

1. OVERVIEW

Multimodal biometrics has raised a growing interest in the industrial and scientific community. The potential increase of accuracy combined with better robustness against forgeries makes indeed multimodal biometrics a promising field. Moreover, important issues like the correlation of the modalities or the combination of multiple and potentially heterogeneous sources of biometric features make the field attractive from a research point of view. However, there is a clear need for multimodal databases, given that current existing databases offer few recorded modalities, often implement unrealistic closed-set scenarios and sometimes use even more unrealistic acquisition conditions.

MyIDea aims at providing the scientific community with a multimodal database including talking face, voice, fingerprints, signature, handwriting, palmprint and hand geometry. Real-life multimodal systems will probably not use all of these modalities at the same time. However, MyIDea will allow to group modalities according to some scenario such as, for example, biometric passports including face and fingerprint. Grouping modalities can also be performed according to the capacities of next-generation sensors that will be able to simultaneously acquire multiple modalities (hand-fingerprint, voice-handwriting, etc).

The specifications of MyIDea can be summarized as follows:

1. target of 104 subjects, recorded over three different sessions spaced in time; no control of minimum or maximum interval between sessions as in real-life applications
2. direct compatibility and potential extension of existing mono or multimodal databases: BANCA [1], BIOMET [2], XM2VTSDB [3], MCYT [4] and IAM [5]
3. different types and qualities of sensors, various and realistic acquisition scenarios, content recorded in French and English
4. organization of the recordings to allow for open-set experiments
5. impostor attempts for voice, signature and handwriting

MyIDea also offers a less-common, novel acquisition mode which is the bimodal voice-signature and voice-handwriting. It has been experimented that these modalities can be simultaneously recorded in a scenario where the user is asked to utter what he is writing. Robustness against impostor attacks is the main advantage of this bimodal approach. The act of speaking and writing at the same time is indeed perceived as a rather easy cognitive task by the user. However, it can be reasonably expected that impostors will have some difficulties to imitate the writing or the signature of someone else while speaking. Another novel point about the signature modality is the use of dedicated software which renders the dynamics of the signal to help impostors to imitate the signature.

At the time of writing this article, specific evaluation protocols were not yet defined on MyIDea. However, most of the parts of MyIDea are compatible with existing databases for which evaluation protocols are already available. Therefore, one can already use these protocols and apply them to MyIDea data. Another possibility is to use MyIDea to extend the evaluation sets defined in these protocols in order to obtain more significant results.

The overall MyIDea project is performed in the framework of collaborations between the University of Fribourg in Switzerland [6], the Engineering School of Fribourg in

Switzerland [7] and the GET in Paris [8]. In Fribourg, MyIDea is supported by the Swiss project IM2 [9]. Feedback and interest have also been raised from the EC Network of Excellence BioSecure [10].

This document is a summarized version of a detailed research report [11] available from the MyIDea website [12]. The organization is as follows: section 2 and 3 give an overview of the acquisition system and of the acquisition protocols; these sections share the same structure which is organized by modality; section 4 introduces briefly the software that has been built to perform the recordings; finally section 5 outlines the availability of the database and future work.

2. ACQUISITION SYSTEM

2.1 Video for voice and face

The settings used to record the voice and face of subjects were inspired by BANCA, BIOMET and XM2VTS. The hardware used to record the videos is of two types: on one side, "good quality" Sony Digital Camera *DCR-HC40* recording in DV quality; on the other side, regular webcams. The audio is captured by high-quality Shure microphones and directly by the integrated microphone of the webcams when applicable. This setting has been reproduced for three different acquisition scenarios similar to the one of BANCA: controlled, degraded and adverse.

In the controlled scenario, people sit in front of a uniform blue wall at a fixed distance of the camera. Lighting conditions are verified: three halogens are placed in order to avoid shadows on the face of the subjects; mean luminance value on the surface of the face is controlled around 650 lux; blinds of the room are shut. The automatic setups of the camera are turned off to avoid unwanted variations in the recordings. Both good quality and web cameras are used in this scenario.

In the degraded scenario, a work environment is simulated. The subject is sitting at a desk, reading data from a computer screen on top of which are placed the cameras. The recording conditions correspond to those of a workplace with a mix of natural light and office lamps, with background noise coming from both a nearby road and with people potentially present in the background. The default automatic settings of the cameras are turned on. Both good quality and web cameras are used for this scenario.

In the adverse scenario, the environment is highly variable, with very little control over the lighting conditions. Background noise is present due to people walking or speaking in the vicinity. The background may change from one session to another, with people passing behind (and sometimes even in front of) the subject. The uncontrolled events are sometimes distracting the subject who shows unwanted movements or changes in intonation. Only the good quality camera is used in this scenario.

2.2 Fingerprints

Two fingerprint sensors are used for the acquisition. Both sensors are controlled by the software developed for the acquisition campaign (see section 4.2) and the drivers provided by the constructor are used.

The first sensor is a high-end optical sensor *Morphosmart MSO-100* from SAGEM [13]. Fingerprints are acquired at 500 dpi on a 21 mm x 21 mm acquisition area, 8-bit grey scale. This reference sensor has been chosen to get direct compatibility with the BIOMET database.

The second sensor is the *TocaBit* scan-thermal sensor from Ekey [14] which is based on a *FCD4B14 FingerChip* from Atmel [15]. This sensor measures the temperature differential between the sensor pixels that are in contact with the finger ridges and those that are not. Provided that the fingertip has been swept at an appropriate speed and pressure, the overlap between successive frames enables an image of the entire fingerprint to be reconstructed [16]. The obtained image is typically 25 mm x 14 mm (equivalent to 500 x 280 pixels), at 8 bit grey scale resolution. This sensor presents the advantage that no cleaning is needed between acquisitions. This device has been chosen as, to our knowledge, no database uses this type of scan-thermal procedure. Furthermore, its small size and robustness make it a good candidate to be embedded on devices such as mobile phones.

2.3 Palmprints

A *perfection 1660* scanner from EPSON 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 4.2). The scanning resolution can be set to various values. As for the optical fingerprint sensor, the glass plate of the scanner needs frequent cleaning to avoid additional noise accumulating between acquisitions. Cleaning is performed with alcohol-soaked napkins at a frequency defined by the protocol.

2.4 Hand geometry

Hand geometry including the dorsum surface (top-view) and the lateral surfaces (left and right side-views) of the hand is captured using an *Olympus C-370* CCD camera. A platform consisting of a plane surface and two inclined lateral mirrors has been built. Pegs are placed on the plane surface to guide the position of the user's hand. One shot takes a picture including the top view and the side views of the hand.

2.5 Signature and handwriting

Signatures and handwritings are acquired with an *A4 Intuos2* graphic tablet from WACOM [17]. An *Intuos InkPen* is used to write on standard paper positioned on the tablet. This procedure presents the advantages to record on-line and off-line data in the same time and to allow a very natural writing using an almost standard pen and paper. This tablet has been chosen for direct compatibility with the BIOMET, MCYT and IAM databases

For the on-line data, the tablet records 5 parameters: x-y coordinates, pressure, azimuth and altitude at a frequency of 100 Hz. The tablet is managed by the software developed for the acquisition campaign (see section 4.2) using the drivers provided by the constructor. The software also allows performing a synchronized acquisition of writing and voice data when the subject is asked to read aloud what he/she is writing (see section 3.5 and 3.6). For this purpose, a computer microphone mounted on a headset is used.

3. ACQUISITION PROTOCOL

The strategy for the constitution of the groups is similar to the one of the BANCA database [1]. Each group has 13 subjects and is gender specific. Impostures are performed within each group, i.e. there are no cross-group impostures. The goal is to reach 8 groups of 13 subjects, i.e. a total of 104 subjects over 3 sessions. The time-interval between two sessions is not controlled and can range from days to months. The protocol is detailed for each modality in the following sub-sections.

3.1 Video for voice and face

Two types of contents are recorded. The first one is similar to the content of the BANCA database [1] and the second one is similar to the content recorded in the BIOMET database [2]:

1. BANCA: This content is recorded in controlled, degraded and adverse scenarios such as described earlier (see section 2.1). The subject is asked to read 10 random digits, a full mail address and a birth date. The subject is then asked to read the content (digits, address, birth date) of another subject of her/his group.

2. BIOMET: This content is recorded using the controlled acquisition scenario. The subject is asked to read digits from 0 to 9, digits from 9 to 0, two phonetically balanced sentences, “yes” and “no”, 10 fixed phrases for text-dependent system, 8 password-like phrases (4 genuine and 4 impostor phrases) and 5 random short phrases for text-independent system evaluation. The BIOMET session ends with head rotation movements for 3D reconstruction algorithms.

3.2 Fingerprints

Overall, 6 fingerprint images of all 10 fingers are captured per session. For each finger, 2 images are taken on the optical Sagem sensor in an “uncontrolled” manner, then 2 more images on the same sensor in a “controlled” manner, and finally 2 images on the scan-thermal Ekey sensor.

For the optical sensor, the uncontrolled mode is performed without controlling the quality and centering of the acquired image and without removing previous fingerprint traces that could remain on the sensor glass. The controlled mode is performed by visually controlling the quality and centering of the acquired image and by cleaning the glass plate between each acquisition. It has been observed that too dry fingers lead to degraded images using this sensor. Therefore, for the controlled mode, subjects are asked to wipe their finger on their forehead prior to each acquisition in order to smooth out the dryness.

The scan-thermal sensor automatically discards improper images according to its internal criterion. Acquisitions are performed until a fingerprint is declared valid by the sensor.

3.3 Palmprint

Per subject and session, palmprints of the right hand are acquired 4 times: 3 times at a resolution of 150 dpi and one last time at a resolution of 400 dpi. Subjects are asked to spread their fingers naturally and keep their rings if any.

3.4 Hand geometry

Three pictures of the dorsal and lateral sides of the right hand are taken per subject and session. These pictures are taken at a nominal resolution of 2048 x 1536 pixels. The macro mode is enabled on the CCD camera in order to enhance the quality of the image. The subject is asked to place his right hand according to the pegs. Rings are not removed from the fingers.

3.5 Signature, signature and voice

Per subject and session, 6 signature samples are acquired according to 5 different scenarios:

1. **true signature**: the subject signs using her/his own signature as naturally as possible.

2. **true signature with voice**: the subject is asked to synchronously sign and utter the content of her/his signature. If the signature contains flourish or not readable signs, the user is simply asked to utter her/his name.

3. **impostor static signature**: for each session, the subject is asked to imitate the signature of another subject of the same group. Few minutes of training are given to the subject.

4. **impostor dynamic signature**: imposture on the same signature as in 3, but this time, the subject is allowed to study the dynamics of the signature thanks to a dedicated software (see section 4.2).

5. **impostor static signature with voice**: for each session, the subject is asked to imitate the signature of another subject (different from imposture in 3 and 4) of the same group and to synchronously utter the content of that signature.

3.6 Handwriting with voice

The subject is asked to synchronously write and utter the content of a text. The content is composed of a fixed generic phrase containing all the letters of the alphabet, followed by a text of 50 to 100 words, randomly chosen from a corpus. The fixed phrase can be used to evaluate text-dependent systems and the random text can be used to evaluate text-independent systems. First, the subject is asked to write and utter the content of the text with his own handwriting, as naturally as possible. Then (s)he is asked to imitate the handwriting of another member of her/his group. For this last part, a training time limited to few minutes is given to the subject.

4. SOFTWARE

4.1 Protocol Generation

A software has been built to automatically generate the protocols. The main idea is to limit the impact of human mistakes by generating a documented protocol for each subject and each session. These generated documents are then printed before each acquisition.

For every subject and every session, there are two types of generated documents. The first one is a protocol to help the assistant monitor the acquisition session. The assistant is supposed to check all points of the protocol before starting with the acquisition of each modality. The second part of the generated documents includes the contents that the subject needs to read or write down. Signature and handwriting are performed on these documents.

4.2 Biometrics acquisition

Two dedicated softwares have been developed to help the acquisition: *Biblios* and *SignReplay*.

The first one, Biblios, is used to monitor the acquisition and to control the different sensors. Biblios has been initially developed for the recording of BIOMET and extended here for MyIDea. Biblios guides the assistant through all the steps involved in a biometric data acquisition session. For all modalities, an on-screen live playback helps the assistant verify the quality of the acquired data. Biblios is able to monitor the following devices: scanner, graphic tablet, audio microphone and fingerprint sensors. Video sequences and hand geometry are captured independently of Biblios.

The second one, SignReplay is used to playback 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 playback speed, which can be equal or lower than the real speed. This software is used to train 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.

5. AVAILABILITY AND FUTURE WORK

MyIDea database will be distributed almost freely to research institutions. A fee will be asked to cover the reproduction costs (manpower and hardware). Institutions willing to get the database will have to sign a license agreement presently under construction. The main points of the license agreement will be as follows: (1) data must be used for research purposes only; (2) licensee has no rights to distribute the data to third parties.

Future works include the completion of the acquisition of the 104 subjects (end of 2005), the validation of all data (early 2006) and the definition of standard assessment protocols.

6. CONCLUSION

We have presented the acquisition protocols of MyIDea, a new biometrics multimodal database. MyIDea is strongly multimodal including talking-face, audio, fingerprints, signature, handwriting and hand geometry. MyIDea presents interesting features such as compatibility with existing databases, use of multiple sensors for the acquisitions and implementation of realistic scenarios in an open-set framework.

7. REFERENCES

[1] E. Bailly-Baillière, 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)*, 2003.

[2] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. Leroux les Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacrétaz. 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)*, 2003.

[3] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In *Proc. Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, 1999.

[4] J. Ortega-Garcia, et al. MCYT baseline corpus: a bimodal biometric database. *IEE Proc.-Vis. Image Signal Process*, 150(6): 395–401, December 2003.

[5] 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.

[6] Fribourg University. http://www.unifr.ch.

[7] Ecole d'Ingénieurs et d'architectes de Fribourg. http://www.eif.ch.

[8] Groupe des Ecoles des Télécommunications, Paris. http://www.get-telecom.fr.

[9] Swiss National Center of Competence in Research on Interactive Multimodal Information Management. http://www.im2.ch.

[10] BioSecure NoE. http://www.biosecure.info.

[11] B. Dumas, F. Evequoz, J. Hennebert, A. Humm, R. Ingold, D. Petrovska, C. Pugin and D. Von Rotz. MyIDea, Sensors Specifications and Acquisition Protocol. University of Fribourg Internal Publication, 05-12, June 2005.

[12] MyIDea Database. http://diuf.unifr.ch/go/myidea/

[13] Sagem S.A. http://www.sagem.com.

[14] Ekey biometric systems GmbH. http://www.ekey.net/.

[15] Atmel Corporation. http://www.atmel-grenoble.com.

[16] Peter Bishop. Atmel's fingerchip technology for biometric security. Technical report, Atmel Corporation, 2002.

[17] Wacom Technology Co. http://www.wacom.com.