Published as: J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano and J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database", Pattern Recognition, Vol. 40, n. 4, pp. 1389-1392, April 2007.

BioSec baseline corpus: A multimodal biometric database

Julian Fierrez-Aguilar *, Javier Ortega-Garcia,

Doroteo Torre-Toledano, Joaquin Gonzalez-Rodriguez,

Biometrics Research Lab. - ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid,

Calle Francisco Tomas y Valiente 11, 28049 Madrid, Spain

Abstract

The baseline copus of a new multimodal database, acquired in the framework of the FP6 EU BioSec Integrated Project, is presented. The corpus consist of fingerprint images acquired with three different sensors, frontal face images from a webcam, iris images from an iris sensor, and voice utterances acquired both with a close-talk headset and a distant webcam microphone. The BioSec baseline corpus includes real multimodal data from 200 individuals in 2 acquisition sessions. In this contribution, the acquisition setup and protocol are outlined, and the contents of the corpus -including data and population statistics- are described. The database will be publicly available for research purposes by mid 2006.

Key words: Multimodal, biometrics, authentication, verification, database, performance, fingerprint, iris, face, voice

* Corresponding author. Tel.: +34-91-4973363; fax: +34-91-4972235
Email address: Julian.Fierrez@uam.es (Julian Fierrez-Aguilar).

1 Introduction

Research on biometric systems has experienced an important growth in the last years. In the case of individual biometric traits such as fingerprint, face, and voice, this has been promoted by the availability of biometric databases developed for international benchmarks such as Fingerprint Verification Competitions, NIST Facial Recognition Technology Evaluations, and NIST Speaker Recognition Evaluations [1]. In the same way, the progress on multimodal biometric systems relies heavily on the availability of multimodal biometric databases.

There are already a few multimodal biometric databases publicly available. Some of them consist of only matching scores produced by several biometric systems operating on different modalities [2]. While these databases encourage research on multimodal fusion, they prevent research on individual systems and even on fusion at other levels than score fusion. In this communication we focus on the more general and common case of multimodal databases consisting of biometric signals. In this respect, prominent examples are: XM2VTS [2], including face and voice; MCYT [3], including fingerprint and handwritten signature; and BIOMET [4], which contains samples of face, voice, fingerprint, hand and handwritten signature. These previously existing databases had several limitations that the BioSec baseline corpus tries to overcome. In particular the BioSec baseline corpus tries to overcome the absence of important traits (e.g., iris), sensors (e.g., sweeping fingerprint sensors), and informed forgery simulations (e.g., voice utterances pronouncing the PIN of another user) in existing multimodal databases.

As compared to unimodal databases, the collection of multimodal databases implies some additional or extended challenges, namely: the design of the contents typically requires a complex multidisciplinary approach, the acquisition campaign is very resource- and time-consuming, there is a need for cooperation of a large group of donators spanning a period of time, and the legal issues regarding data protection are especially controversial [1]. One of the main goals of the European project BioSec [5] is to create a large multimodal database overcoming these difficulties by the integrated efforts of over 20 partners with accumulated experience in biometric database acquisition, personal data protection, performance evaluation, and usability and user acceptability issues. This effort has resulted in a comprehensive multimodal biometric database including fingerprint, iris, voice and face. The resulting database is referred to as BioSec baseline corpus. This communication describes the methodology carried out to deal with the challenges arising in multimodal database acquisition, and the resulting database with statistics of the population acquired.

2 Acquisition of BioSec baseline

Acquisition of BioSec baseline was jointly conducted by Universidad Politecnica de Madrid (UPM) and Universitat Politecnica de Catalunya (UPC) in Spain. The scenario in the acquisition was an office room, with a wide desktop for the acquisition hardware and two chairs for the donator and the supervisor of the acquisition. The supervisor ran a specific purpose tool for multimodal database acquisition developed within the project. Environmental conditions (e.g. lighting, background noise, etc.) were not controlled in order to simulate a realistic situation. The acquisition hardware included a standard personal

computer and a number of commercial biometric sensors, which are depicted in the left column of Fig. 1.

Biosec baseline corpus contains multimodal biometric traits of 200 different subjects. Each subject participated in two acquisition sessions separated typically by one to four weeks (see Fig. 2). For each subject and session the following information was gathered for each of the modalities considered:

Face. 4 frontal face images in neutral pose at about 30 cm distance to the camera (2 at the beginning and 2 at the end of the session). The individuals were requested to change their facial expressions between consecutive acquisitions in order to avoid identical face samples. The total number of face images in the corpus is $N_F = 1600$.

Voice. 4 utterances of a user-specific number of 8 digits (2 at the beginning and 2 at the end) and 3 utterances of other users' numbers to simulate informed forgeries in which an impostor has access to the number of a client. The 8 digits were always pronounced digit-by-digit in a single continuous and fluent utterance. The 8 digits were recorded both in Spanish and English. The total number of voice utterances is therefore $N_V = 2 \times 200 \times (4 + 3) \times 2$ sensors $\times$ 2 languages.

Iris. 4 iris images of each eye changing eyes between consecutive acquisitions, resulting in $N_I = 3200$ iris images in the corpus.

Fingerprint. 4 captures with each of 3 different sensors (see Fig. 1) of the print of 4 fingers (right and left index and middle), interleaving fingers between consecutive acquisitions. The total number of fingerprint images is $N_P = 2 \times 200 \times 4 \times 3$ sensors $\times$ 4 fingers.

——— Fig. 1 ———

Before starting the first acquisition session, the donors read and signed a consent agreement. This agreement considers biometric data as "personal data", following the European requirements on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Directive 95/46/EC of the European Parliament and the Council of 24 October 1995). First and second sessions typically took about 25 and 15 minutes, respectively.

Other personal data acquired in the first session and stored securely and independently of the biometric data included: name, age, gender, handedness, manual worker (yes/no), vision aids (glasses, contact lenses, none), and English proficiency (low, foreign language, native). A time log of the different acquisition files was also stored. The "manual worker" group includes all users having eroded fingerprints, as identified by the contributors themselves when asked about the state of their fingerprints. The use of glasses, contact lenses or none of them refers to regular use. The donators using glasses wore them for the face capture but removed them for the iris acquisition.

Acceptance and usability data was also gathered during the acquisition by using electronic questionnaires. The questionnaires included over 50 questions related to attitude, security and privacy, background, scenarios, and enrolment. More details and results of the acceptance and usability studies can be found in [5].

3 Description of BioSec baseline

Examples of typical images in BioSec baseline are given in the central part of Fig. 1 (different traits corresponding to different random subjects). Selected biometric samples with very low image quality are given in the right column of Fig. 1.

Although the corpus was carefully collected by specially designed software and a human supervisor at all times, there was still the possibility of software or human errors. In order to ensure that the BioSec baseline corpus was conformed to the acquisition protocol, all acquired biometric samples were manually verified by a human expert. The samples non-compliant with the acquisition protocol were either corrected or removed according to the following rules:

- If a user lacks an important part of his/her biometric data (approximately more than 20% of all the genuine samples), then the user is not included in BioSec baseline, i.e., it is rejected.
- If a user lacks a reduced number of genuine samples (approximately less than 20%), then the samples are copied from valid samples of the same user. Therefore some identical samples appear in BioSec baseline.
- In the particular case of utterances of a user pronouncing the 8 digit number of another user to simulate informed forgeries, the expert verifying the database produced the samples missing or invalid with his own voice.

Worth noting, low quality samples, even of the poorest quality but complaint with the acquisition protocol were neither rejected nor corrected.

Statistics about rejected and corrected users and samples are given in Fig. 2, which also includes population and session statistics.

———— Fig. 2 ————

BioSec baseline corpus was made available to BioSec partners by mid 2005. It will be made publicly available by mid 2006 either through the ATVS-Biometrics Research Lab. website (http://atvs.ii.uam.es), or through a third party like the Evaluations and Language resources Distribution Agency, ELDA (http://www.elda.org) [1] .

Acknowledgements

This work has been supported by the European project BioSec (IST-2002-001766). The author J. F.-A. is also supported by a FPI Fellowship from Comunidad de Madrid. The authors would like to thank the valuable development work of Alberto Posse, Jaime Lopez, Diego Rodriguez, Manuel Freire, Jan Anguita, and Mireia Farrus; the contribution in the acquisition led by Javier Hernando of Universitat Politecnica de Catalunya (UPC); and the support obtained from the Iris Technology Division of LG Electronics.

References

[1] J. Wayman, A. Jain, D. Maltoni, D. Maio (Eds.), Biometric Systems: Technology, Design and Performance Evaluation, Springer, 2005.

[1] Yet to be decided as of early 2006

[2] N. Poh, S. Bengio, Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication, Pattern Recognition 39 (2006) 223–233.

[3] J. Ortega-Garcia, J. Fierrez-Aguilar, et al., MCYT baseline corpus: A bimodal biometric database, IEE Proc.–Vis. Image Signal Process. 150 (2003) 395–401.

[4] S. Garcia-Salicetti, C. Beumier, G. Chollet, et al., BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities, Lect. Notes Comput. Sc. 2688 (2003) 845–853.

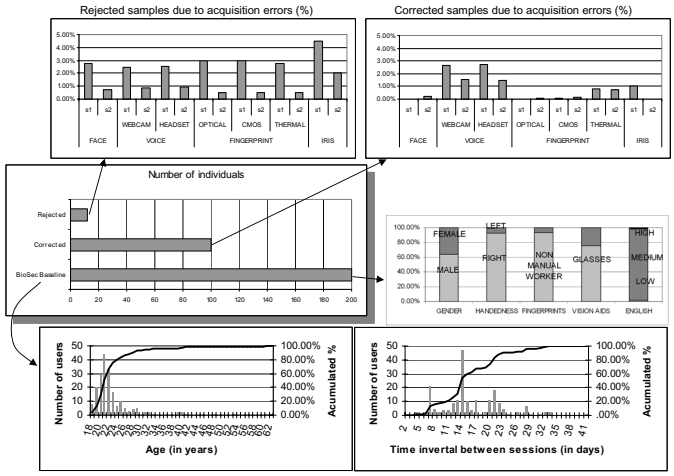
[5] BioSec, Biometrics and Security, FP6 IP IST-2002-001766 (http://www.biosec.org/).

Figure captions:

Fig. 1. BioSec baseline: sensors used (left column), example of multimodal biometric data (center column), and selected low quality samples (right column). Voice utterances are depicted as waveforms.

Fig. 2. BioSec baseline: statistics of the biometric data and population acquired.


SENSORS
BIOMETRIC SAMPLES
(Different sensors corresponding to different subjects)
seconds
minutes
days
session1
session1
session1
session2
SELECTED LOW QUALITY SAMPLES
AUTHENTEC AES4000
ATMEL FCDEMO4
BIOMETRIKA FX2000
PHILIPS ToUcam PRO II
7-3-7-7-0-0-0-9
7-3-7-7-0-0-0-9
3-9-9-6-1-0-8-0
PLANTRONICS DSP-400
7-3-7-7-0-0-0-9
7-3-7-7-0-0-0-9
8-2-9-2-2-6-0-4
LG IrisAccess EOU3000


Rejected samples due to acquisition errors (%)
Corrected samples due to acquisition errors (%)
FACE
VOICE
FINGERPRINT
IRIS
WEBCAM
HEADSET
OPTICAL
CMOS
THERMAL
Number of individuals
Rejected
Corrected
BioSec Baseline
GENDER
HANDEDNESS
FINGERPRINTS
VISION AIDS
ENGLISH
FEMALE
MALE
LEFT
RIGHT
NON-MANUAL WORKER
GLASSES
HIGH
MEDIUM
LOW
Number of users
Acumulated %
Age (in years)
Time invertal between sessions (in days)