Multimodal Biometric Databases

Dr. Julian Fierrez
http://atvs.ii.uam.es/fierrez
Biometrics Research Lab.
Universidad Autónoma de Madrid, Spain

Outline

1. INTRODUCTION
   - Motivation for Multimodal Databases
   - Why Multimodality? Which Modalities?
   - Other Multimodal Databases: Chimerical and Score-Level
2. ACQUISITION OF MULTIMODAL DATABASES
   - Database Design
   - Acquisition Scenario
   - Multi-Session Acquisition
   - Acquisition Software
   - Legal Issues
3. ERROR ESTIMATION WITH MULTIMODAL DATABASES
   - Test Set Size for Good Error Estimates
4. EXISTING MULTIMODAL DATABASES
5. CASE STUDIES
   - BiosecurID
6. FINAL CONSIDERATIONS
   - Challenges and Future Directions
   - Potential research lines exploiting Multimodal Databases
1. INTRODUCTION

Motivation for MM DBs

- Importance of evaluating algorithms on common benchmarks.
- Importance of having REAL multimodal data for research and advancement of the field (multi-biometrics*).
- Realistic contents, scenarios and applications.
- Include more intra- and inter-variability factors in order to understand and learn how to deal with variability.

It is often the existence of new challenging datasets that drives the research forward → We need data, more data!!

* A. A. Ross et al., Handbook of Multibiometrics, Springer, 2006
Why Multimodality?

- Multibiometrics*: use of multiple sensors, instances, realizations, algorithms and/or modalities for biometric person authentication
- Why multimodality?
  - Some subjects may experiment difficulties with a specific modality
  - Some modalities better adapted than others to specific applications
  - Exploit fusion capabilities for low-correlated sources
  - Avoid noisy conditions or sensors
  - System vulnerability, attacks for a specific modality
- 1,000s databases containing only 1 modality
- Some of them are multibiometric, but just very few are large and include several modalities

Which Modalities?

- For applications like Homeland Security, Border Control, Access Control, 3 modalities stand out:
  - Fingerprint
  - Iris
  - Face
- Why?
  - Fingerprint are very discriminative; experience in forensics
  - Iris authentication is very discriminative
  - Face recognition is transparent and convenient
Which Modalities?

• But…
  - No low-cost user-friendly iris cameras
  - Face recognition under realistic conditions is comparatively one of the less reliable modalities
  - Fingerprint sensors produce only medium-quality images and can be fooled

• And…
  - **Speech** is unbeatable for phone-based applications
  - **Hand** modalities (geometry, palmprint) are perceived as less intrusive by users
  - **Handwritten signature** is well-suited for written documents (legal, bank documents) and input-based devices (points of sale, smartphones, …)
  - **Gait** is even more transparent than face for access control

Some Pros and Cons

• **Face**: Affected by background, lighting, pose and expression; short-term and aging variability. **Transparency. Low-cost sensors.**

• **Voice**: Affected by background noise, channel; short-term and aging variability. **Phone transparency and universality. Forensic concern (wire tapping of cell phones). Remote acquisition.**

• **Iris**: focusing; eyelid occlusions; eyelash noise; user cooperation, intrusiveness. **Individuality, low intra-variability.**

• **Fingerprints**: Elastic distortion, incomplete (nail-to-nail) image; quality of sensors; user cooperation. **Forensic concern. Reliability, low intra-variability.**

• **Signature**: Easy to forge; over-the-shoulder dynamic forgeries; **very low background influence. Well adapted to written docs scenarios.**

• **Hand (geometry / palm)**: Lighting; elastic distortion (touch). **High user-acceptance; no intrusive. Complementary to fingerprint.**
### Which Modalities?

- **Some conclusions**
  - No modality is 'the winning one': it all depends on the scenario and the specific application.
  - Modalities cannot be ranked only in terms of reliability, as many other factors must be considered: user acceptance, usability, interoperability, transparency, user cooperation, easy-of-use, failure-to-enroll, etc.
  - The combination (fusion) of modalities will allow better adaptation to specific problems.
  - Some modalities are complementary for specific applications: hand/fingerprint, speech/face, iris/face, ...

### Chimerical Databases

- Chimerical databases are created by putting together separate datasets, and defining virtual subjects and experimental protocols on them.
Chimerical Databases

- Completely unrealistic procedure
- Difficulty in integrating heterogeneous datasets
- Difficulty in setting up common evaluation protocols
- Supposes independence between modalities: hand geometry/palmprint; face/lip dynamics; speech/face, ...

Score-Level MM Databases

- **NIST Biometric Scores Set (BSSR1)**
  - Raw output similarity scores from two face recognition systems (frontal face) and one fingerprint system (left and right live-scan indexes)
  - The release includes true multimodal score data, i.e., similarity scores from comparisons of faces and fingerprints of the same people

- **IDIAP\*:**
  - Built on XM2VTS (Lausanne Protocols I and II)
  - Fusion with all the possible combinations across protocols

- **MCYT\**:  
  - Built on MCYT (75 subjects, signature + fingerprint + Q_finger)
  - Adequate for research in user-dependent and quality-based fusion


2. ACQUISITION OF MULTIMODAL DATABASES

Database Design

- Sensors and equipment selection
- Acquisition protocol
- File formats; storage needs
- Backup considerations
- Legal issues; privacy concerns
- Database management of raw / formatted data
- Distribution issues; public vs. private datasets
- Evaluation protocol outline
Possible scenarios

- “Classical” scenarios:
  - Desktop acquisition
  - Supervised kiosk acquisition

- Fully unsupervised acquisition
- Internet-based acquisition
- Mobile devices (PDAs, smartphones, etc.)
- Indoors / outdoors
- Home environment (domotics)
- Car environment

- Ambient intelligence / pervasive computing

Multi-Session Acquisition

- Biometric authentication requires multi-session acquisition
- Adapted to the real operation of a recognition system (registration + use)
- Capture intra-variability of biometric samples
  - Even for low intra-variability modalities (iris, fingerprints)
  - Aging and long-term characterization of subjects (face, voice, gait)
- Train ID models with multi-session user samples
- Time span: depends on how many sessions you can acquire
  - Short-term intra-variability (1 week - 1 month)
  - Long-term (months - years)
• A software suite is usually required, in order to manage sessions, subjects, data storage, data editing, backups, ... (caption of the software tool used for BioSec, Biosecur-ID, and BioSecure)

• ... and perhaps annotation during acquisition (captions of the acquisition software for MYIDEA database)
Legal Issues (European perspective)

- **Directive 95/46/EC** of the European Parliament and the Council of 24 October 1995 sets the European requirements on the protection of individuals with regard to the processing of personal data and on the free movement of such data (+ several slightly different national legal implementations)
  - **Biometrics** are personal data:
    - Informed consent agreement before acquisition
    - “Controller” vs “Processor” of the data
    - The controller has to guarantee certain security measures, register the database to his national data protection authority, keep file of any change/access to the DB, and be readily available upon justified request of any subject to modify/delete her personal data
    - No clear legislation about transfer of DBs of personal data (i.e., biometric databases) for research purposes:
      - Delay in the distribution of several DBs already acquired!!!
      - No strict accordance to legislation in most practical cases!!!

3. **ERROR ESTIMATION WITH MULTIMODAL DATABASES**

Julian Fierrez  Multimodal Biometric Databases  19/68

Julian Fierrez  Multimodal Biometric Databases  20/68
Test Set Size for Good Error Estimates*

The minimum number of test trials, \( N \), to be performed in order to guarantee, with a risk \( \alpha \) of being wrong, that the estimated error is under an upper limit with a given confidence is:

\[
\Pr \left\{ \hat{P} > P + \epsilon(N, \alpha) \right\} < \alpha
\]

If the error distribution is Gaussian (as usually in our case), we can make a better estimation of the upper limit, as defined by the expression:

\[
\Pr \left\{ P - \hat{P} > \frac{z_{\alpha} \sigma}{\sqrt{N}} \right\} < \alpha
\]

and a good estimate of \( z_{\alpha} \) in this range of values is:

\[
z_{\alpha} \approx \sqrt{-\ln \alpha}
\]


Test Set Size for Good Error Estimates

Supposing that the error margin is a fraction of the error rate, \( \epsilon(N, \alpha) = \beta P \), we can assure, with a confidence \((1-\alpha)\), for a given number of test set trials, \( N \):

\[
N = \left( \frac{z_{\alpha}}{\beta} \right)^2 \frac{(1-P)}{P}
\]

that the expected error \( P \) is under:

\[
\hat{P}/(1-\beta)
\]

Simplifying and taking the more strict value, for typical values of \( \alpha \) and \( \beta \) (0.05 and 0.2, respectively):

\[
N \approx \frac{100}{P}
\]
Test Set Size for Good Error Estimates

Rule of thumb:

In order to obtain an error estimation \( P \) with a statistical confidence of 95%, you need to accomplish at least \( N \) = \( \frac{100}{P} \) test trials

- \( P = 0.01 \) \( \rightarrow \) \( N = 10,000 \) test trials
- \( P = 0.001 \) \( \rightarrow \) \( N = 100,000 \) test trials

Alternative rule of thumb:

You need to have at least 100 errors in your test trials in order to have a confidence of 95% in your error estimation

BIG PROBLEM with the previous model when applied to BIOMETRICS:

- Does not consider several forms of statistical dependence (user, sensor, acquisition conditions, ...) which leads to errors not iid:

  \( \rightarrow \) More sophisticated models [Guyon et al., 98][Dass et al. 06*]
  
  \( \rightarrow \) Unfortunately, no rules of thumb!!! Not usually considered in practice

- Best practice in Biometrics:

  \( \rightarrow \) careful description of the data + error rates (+ rules of thumb regarding statistical significance considering simple models just to check that the error rates makes sense)

4. EXISTING MULTIMODAL DATABASES

Summary of Existing MM DBs*

NOTE: The table is not exhaustive (it is focused on European MM DBs). Other important MM DBs are those of NIST (e.g., MBGC) and CASIA.

<table>
<thead>
<tr>
<th>Year</th>
<th>Ref.</th>
<th>Users</th>
<th>Sessions</th>
<th>Trains</th>
<th>2Fa</th>
<th>3Fa</th>
<th>Fp</th>
<th>Hb</th>
<th>I</th>
<th>R</th>
<th>Sg</th>
<th>Sp</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioSecure</td>
<td>2008</td>
<td>971 (DS1, Internet scenario)</td>
<td>2</td>
<td>6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>BioSecure</td>
<td>2007</td>
<td>657 (DS2, Desktop scenario)</td>
<td>2</td>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>BioSecure</td>
<td>2006</td>
<td>713 (DS3, Mobile scenario)</td>
<td>2</td>
<td>6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MCYT</td>
<td>2005</td>
<td>[10]</td>
<td>320</td>
<td>3</td>
<td>2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

TABLE I

NOTE: The table is not exhaustive (it is focused on European MM DBs). Other important MM DBs are those of NIST (e.g., MBGC) and CASIA.

XM2VTS Database*

- **Context:** M2VTS project (Multi Modal Verification for Teleservices and Security applications), EU ACTS programme, dealing with access control by the use of multimodal identification based on face and voice.

- **Contents:** Microphone speech and face video sequences. At each session, two head rotation shots and six speech shots (subjects reading three sentences twice).

- **Subjects:** 295

- **Sessions:** 4 sessions over a period of 4 months.

- **Evaluation:** an evaluation protocol (Lausanne protocol) specifies training, evaluation, and test sets (algorithmic recognition performance can be assessed on the basis of comparable evaluation framework).

- **Distribution:** A variety of subsets are available from the University of Surrey; already distributed to more than 100 institutions.


Julian Fierrez  Multimodal Biometric Databases  27/68

XM2VTS Database

- **Some example snapshots of XM2VTS Database**

Julian Fierrez  Multimodal Biometric Databases  28/68
BANCA Database*

• Contents: 2 modalities (face and voice), 4 languages (English, French, Spanish, Italian); random 12-digit number along with a name, address, and date of birth.

• Equipment: Both High and low quality microphones and cameras.

• Sessions ans Scenarios: 3 different scenarios: controlled, degraded, and adverse; 12 different sessions, in a time span of 3 months.

• Subjects: 208, with balanced gender distribution.

• Evaluation: An associated evaluation protocol is available.


---

Still face images of the same subject under 3 different scenarios: (from left to right) controlled, degraded and adverse.

• Teleworking scenario; high quality camera, uniform background, good lighting; Low noise-level, 2 mics.

• Home banking scenario; webcam, few moving elements in the background, normal lighting conditions, office noise, two mics.

• Adverse scenario: high quality camera, moving background, outdoor/indoor lighting conditions, noisy enviroment, 2 mics.
BIOMET Database*

- **Consortium:** French GET Consortium + French Swiss
- **Modalities:** 5 different modalities: speech (French), face, hand image, fingerprint and signature.
- For the face modality, also infrared illumination is used.
- **Sessions:** 3 sessions, with 3 and 5 months time-spanning from the first one.
- **Subjects:** 130 (1st session), 106 (2nd session), and 91 (3th session).


Some BIOMET Database samples:
MCYT Database*

- **Consortium:** 4 Spanish Academic institutions
- **Contents:** On-line signature, fingerprints
  - Fingerprints: 12 samples of each finger with varying position restrictions are acquired using two different sensors (optical and capacitive).
  - Signature: 25 client signatures and 25 skilled forgeries (with natural dynamics) are obtained for each individual.
- **Subjects:** 330
- **Sessions:** 1
- **Distribution:** available from the Universidad Autonoma de Madrid; already distributed to more than 60 institutions.

MYIDEA Database

- **Consortium:** MYIDEA database was acquired by the University of Fribourg in Switzerland, the Engineering School of Fribourg in Switzerland and the GET in Paris.
- **Modalities:** face, voice, fingerprints, signature, handwriting and hand geometry.
- **Subjects:** 104 subjects (aprox.)
- **Equipment:** different quality of sensors
- **Scenarios:** various realistic acquisition scenarios, and organization of the recordings to allow for open-set experimental scenarios.
- **Characteristics:** Two synchronized recordings are performed: face-voice and writing-voice. Video and handwriting-voice annotation tools are used in this case.

**BIOSEC Database***

- **Consortium:** Inside FP6 BioSec IP, 4 partners (UPM, Telefonica, Finnair, UPC)
- **Modalities:** Face, Voice, Fingerprints, Iris
- **Scenario:** Desktop acquisition, big room, no background or illumination control
- **# Subjects:** 250
- **# Sessions:** 4
- **Interoperability Issues:** 3 different fingerprints sensors used in all cases (optical, capacitive, sweeping thermal); low and high quality mics. Employed.
- **Others:** 3 languages: English, Spanish, Finish; PIN-based reply attacks

5. CASE STUDY: BIOSECURID

Julian Fierrez  Multimodal Biometric Databases  38/68
**BIOSECURID Database**

- **Consortium:** BIOSECUR-ID database was acquired by 6 Spanish Universities.
- **Modalities:** 2D face (still + video sequence), voice, fingerprints, iris, signature, handwriting hand (both geometry and palmprint), keystroking.
- **# Subjects:** 400 subjects
- **# Sessions:** 3 sessions
- **Equipment:** several sensors: fingerprints, speech, face (webcam + still photo), signature (tablet + screen), and hand modalities (doc scanner + photo)
- **Scenarios:** PC-based desktop indoor acquisition.
- **Characteristics:** synchronized audio - video recording. Compatible with BioSec database acquisition protocol.


---

**Acquisition Environment**

- Relaxed acquisition scenario (variability between samples acquired in different sites → real-world conditions).
  - Neutral lighting (no preponderant focuses)
  - Indoor noise conditions (no excessive background noise)
  - Pose: sitting in a non-revolving chair
Acquisition Devices

- **PLANTINOCS 20P-009 (Audio Reader)**
  - Sensor: Pre-Processed
  - Microphone: Antimicrobial
  - Encoder: Compression de audio: 10 kHz - 10 MHz

- **BIOQUEST 2000 (Optional)**
  - Resolution (dpi): 120 x 75
  - Number of images: 600 x 500
  - Area captured (pixels): 4.12 x 3.05

- **VIDA (Sensor based on the capture of pictures)**
  - Resolution (dpi): 70 x 35
  - Number of images: 500
  - Area captured (pixels): 4.32 x 3.02

- **PHILIPS TeleCom PRO II (Facial)**
  - Resolution (dpi): 10 x 9 x 34
  - Number of images: 640 x 480
  - Area captured (pixels): 1

**LG Imager 3000**
- Resolution (dpi): 1000 x 800
- Number of images: 640 x 480
- Area captured (pixels): 1

**VACOON NTUOS 9 (Touchplate digitalizer)**
- Resolution (dpi): 1000 x 600
- Number of images: 5000
- Area captured (pixels): 4.32 x 3.02

**Labs Standard Keyboard 6 (Touchpad)**
- Resolution (dpi): 1000 x 600
- Number of images: 1

**Scanner LGNOS P888 0999 Flats**
- Resolution (dpi): 277 x 507
- Number of images: 4300 x 3500
- Area captured (pixels): 2.16 x 2.97

---

Acquisition Protocol

- **Summary of the biometric data available for each user in BiosecurID**

<table>
<thead>
<tr>
<th>Modality</th>
<th>Samples</th>
<th>α Samples</th>
<th>Storage space (Mib)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech</td>
<td>10 short sentences</td>
<td>30</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>4 x 4 PIN genuine</td>
<td>16</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>3 x 4 PIN imitations</td>
<td>12</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>4 x 4 thermal</td>
<td>38</td>
<td>32.7</td>
</tr>
<tr>
<td>Fingertips</td>
<td>4 x 4 x 4 optical</td>
<td>64</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>4 x 4 x 4 thermal</td>
<td>64</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>128</td>
<td>22.5</td>
</tr>
<tr>
<td>Iris</td>
<td>2 x 4 x 4</td>
<td>32</td>
<td>9.4</td>
</tr>
<tr>
<td>Hand</td>
<td>2 x 4 x 4</td>
<td>32</td>
<td>11.6</td>
</tr>
<tr>
<td>Face</td>
<td>4 x 4 still faces</td>
<td>16</td>
<td>14.1</td>
</tr>
<tr>
<td></td>
<td>4 x 4 talking faces</td>
<td>4</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>8.28</td>
</tr>
<tr>
<td>Writing</td>
<td>1 x 4 lower-case text</td>
<td>4</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>1 x 4 upper-case words</td>
<td>4</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>1 x 4 number sequence</td>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>3.7</td>
</tr>
<tr>
<td>Signatures</td>
<td>4 x 4 genuine signatures</td>
<td>16</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>3 x 4 skilled forgeries</td>
<td>12</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>1.0</td>
</tr>
<tr>
<td>Keystroking</td>
<td>4 x 4 genuine name</td>
<td>16</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>3 x 4 skilled forgeries</td>
<td>12</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Acquisition Examples (I)

- Hand
- Face

Acquisition Examples (II)

- Signature - Genuine
- Signature - Forgery

- Keystroking - Genuine
- Keystroking - Imitation

Julian Fierrez

Multimodal Biometric Database
Acquisition Examples (III)

Handwriting – LowerCase

Handwriting – UpperCase

Handwriting – Numbers

Compatibility with other MM DBs

- BIOSEC → 1 year → BIOSECUR-ID → 1 year → BIOSECURE
  - Research on Aging and Time Adaptation!!

Table 5. Summary of the main compatibilities of BiosecurID with other existing multimodal databases (total number of subjects in brackets)

<table>
<thead>
<tr>
<th></th>
<th>Fa</th>
<th>Fp</th>
<th>Ha</th>
<th>Hw</th>
<th>Ir</th>
<th>Ks</th>
<th>Sg</th>
<th>Sp</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOSEC Desktop (ca. 700)</td>
<td>20</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOSEC mobile (ca. 700)</td>
<td>20</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOSEC (250)</td>
<td>37</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MyIDEA (ca. 104)</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCYT (330)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2Fa face 2D, 3Fa face 3D, Fp fingerprint, Ha hand, Hw handwriting,Ir iris, Ks keystroking, Sg signature, and Sp speech.
5. CASE STUDY: BIOSECURE

DATASETS:
- DS1 (Internet): Voice, face
- DS2 (Desktop): Voice, face, signature, fingerprint, iris, hand
- DS3 (Mobile): Voice, face, signature, fingerprint

STATISTICS:
- 11 acquisition sites across Europe
- 2 acquisition sessions for each DS (2 months between them)
- Subjects: 971 DS1, 667 DS2, 713 DS3 (400 common)

AVAILABILITY:
- Through the Biosecure Association (http://www.biosecure.info)

### Internet Dataset (DS1)

**DS1: Voice, face**

- PC-based, on-line, unsupervised (Internet)
- Equipment: low-cost webcam and Bluetooth microphone

---

#### Internet Dataset (DS1)

**Acquisition protocol (per session, total duration per session around 20 minutes, COMMON to the 3 DSs):**

<table>
<thead>
<tr>
<th>Mode ID</th>
<th>Sample ID</th>
<th>Data Type</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1-2</td>
<td>Image</td>
<td>2 still frontal face images</td>
</tr>
<tr>
<td>C</td>
<td>1-2</td>
<td>AV</td>
<td>2 repetitions of a 4-digit PIN code (the same between DSs) from a set of 100 different PINs in English</td>
</tr>
<tr>
<td>C</td>
<td>3-4</td>
<td>AV</td>
<td>2 repetitions of a 4-digit PIN code (different to C1-2, the same between DSs) from a set of 10 different PINs in native language</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>AV</td>
<td>Digits from 0 to 9 in English</td>
</tr>
<tr>
<td>S</td>
<td>1-2</td>
<td>AV</td>
<td>2 different phonetically rich sentences in English (different between DSs)</td>
</tr>
<tr>
<td>S</td>
<td>3-4</td>
<td>AV</td>
<td>2 different phonetically rich sentences in native language (different to S1-2, different between DSs)</td>
</tr>
</tbody>
</table>
Desktop Dataset (DS2)

DS2: Voice, face, signature, fingerprint, iris, hand

<table>
<thead>
<tr>
<th>Mode</th>
<th>Sample</th>
<th>Data Type</th>
<th>Sensor</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI</td>
<td>1-5</td>
<td>Signatures</td>
<td>Tablet</td>
<td>5 genuine of donor n</td>
</tr>
<tr>
<td>SI</td>
<td>6-10</td>
<td>Signatures</td>
<td>Tablet</td>
<td>5 dynamic imitations of donor n – 1 (n–3 session 2)</td>
</tr>
<tr>
<td>SI</td>
<td>11-15</td>
<td>Signatures</td>
<td>Tablet</td>
<td>5 genuine of donor n</td>
</tr>
<tr>
<td>SI</td>
<td>16-20</td>
<td>Signatures</td>
<td>Tablet</td>
<td>5 dynamic imitations of donor n – 2 (n–4 session 2)</td>
</tr>
<tr>
<td>SI</td>
<td>21-25</td>
<td>Signatures</td>
<td>Tablet</td>
<td>5 genuine of donor n</td>
</tr>
<tr>
<td>COMMON – AUDIO / VIDEO (simultaneously with the webcam and the Bluetooth earbud)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>1-4</td>
<td>Iris images</td>
<td>Iris cam</td>
<td>(Right eye → Left eye) x 2 times</td>
</tr>
<tr>
<td>FO</td>
<td>1-12</td>
<td>Fingerprint</td>
<td>Optical</td>
<td>(R_thumb → R_index → R_middle → L_thumb → L_index → L_middle) x 2</td>
</tr>
<tr>
<td>FT</td>
<td>1-12</td>
<td>Fingerprint</td>
<td>Thermal</td>
<td>(R_thumb → R_index → R_middle → L_thumb → L_index → L_middle) x 2</td>
</tr>
<tr>
<td>HA</td>
<td>1-8</td>
<td>Hand</td>
<td>Camera</td>
<td>(Right hand x 2 times → Left hand x 2 times) without flash → (THE SAME) with flash</td>
</tr>
<tr>
<td>FA</td>
<td>1-4</td>
<td>Face</td>
<td>Camera</td>
<td>2 photos without flash → 2 photos with flash (ISO-like conditions)</td>
</tr>
</tbody>
</table>
Mobile Dataset (DS3)

**DS3: Voice, face, signature, fingerprint**

- Equipment: mobile devices (PDA and Ultra-Mobile PC)
- Indoor and outdoor conditions

<table>
<thead>
<tr>
<th>Mode ID</th>
<th>Place (standing)</th>
<th>Sample ID</th>
<th>Data Type</th>
<th>Sensor</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI</td>
<td>Indoor</td>
<td>1-5</td>
<td>Sign</td>
<td>iPAQ</td>
<td>5 signatures of donor n</td>
</tr>
<tr>
<td>SI</td>
<td>Indoor</td>
<td>6-10</td>
<td>Sign</td>
<td>iPAQ</td>
<td>5 dynamic imitations of donor n – 1 (n=3 session 2)</td>
</tr>
<tr>
<td>SI</td>
<td>Indoor</td>
<td>11-15</td>
<td>Sign</td>
<td>iPAQ</td>
<td>5 signatures of donor n</td>
</tr>
<tr>
<td>SI</td>
<td>Indoor</td>
<td>16-20</td>
<td>Sign</td>
<td>iPAQ</td>
<td>5 dynamic imitations of donor n – 2 (n=4 session 2)</td>
</tr>
<tr>
<td>SI</td>
<td>Indoor</td>
<td>21-25</td>
<td>Sign</td>
<td>iPAQ</td>
<td>5 signatures of donor n</td>
</tr>
<tr>
<td>FT</td>
<td>Indoor</td>
<td>1-12</td>
<td>Finger</td>
<td>iPAQ</td>
<td>(R-thumb → R-index → R-middle → L-thumb → L-index → L-middle) x 2</td>
</tr>
</tbody>
</table>

COMMON – AUDIO / VIDEO (Q1 + WebCam) – INDOOR

COMMON – AUDIO / VIDEO (Q1 + WebCam) – OUTDOOR

Julian Fierrez  Multimodal Biometric Databases  54/68
BioSecure MM DB: Examples

Fingerprints - Optical

Iris

Fingerprints - Thermal

BioSecure MM DB: Low Q Samples

Julian Fierrez  Multimodal Biometric Databases  55/68

Julian Fierrez  Multimodal Biometric Databases  56/68
"Mobile" scenario (DS3)*: talking faces, signature, fingerprint
Objective: to test the robustness of mono and multimodal systems
The participants were provided with raw data (monomodal) and development scores (multimodal)

“Access control” scenario (DS2)**: face, fingerprint, iris
- **Quality-based evaluation**: aimed at achieving the best verification performance using q-based score fusion algorithms
- **Cost-based evaluation**: aimed at minimizing a criterion combining verification error rates with the cost of deployment (the use of each biometric trait is associated with a given cost)

The participants were provided with development scores and biometric data quality information for each trait

17 laboratories, 50 different fusion systems submitted

BMEC 2007 Q-based Evaluation

Traits and devices:

<table>
<thead>
<tr>
<th>Mode</th>
<th>Data type</th>
<th>Sensor</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>fn1</td>
<td>Face still</td>
<td>Digital camera (high resolution)</td>
<td>FrONTAL face images</td>
</tr>
<tr>
<td>ft1</td>
<td>Fingerprint</td>
<td>Webcam (low resolution)</td>
<td></td>
</tr>
<tr>
<td>f01, f02, f03</td>
<td>Fingerprint</td>
<td>Optical</td>
<td>1 right thumb, 2 right index</td>
</tr>
<tr>
<td>ft1, ft2, ft3</td>
<td>Fingerprint</td>
<td>Thermal</td>
<td>3 right middle finger</td>
</tr>
</tbody>
</table>

Possible mixtures for each access:

<table>
<thead>
<tr>
<th>Mixture</th>
<th>Modalities</th>
<th>Face</th>
<th>Fingerprint</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fn1/f01/f02/f03</td>
<td>Good quality</td>
<td>Good quality</td>
</tr>
<tr>
<td>2</td>
<td>fn1/ft1/xf1/xf2/xf3</td>
<td>Good quality</td>
<td>Bad quality</td>
</tr>
<tr>
<td>3</td>
<td>xfa1/fo1/fo2/fo3</td>
<td>Bad quality</td>
<td>Good quality</td>
</tr>
<tr>
<td>4</td>
<td>xfa1/xf1/xf2/xf3</td>
<td>Bad quality</td>
<td>Bad quality</td>
</tr>
</tbody>
</table>

- 1 face score, 3 fingerprint scores per access
- xft/xfα: template image is acquired using the good quality sensor and query image is acquired using the bad quality sensor
- All fingerprints are acquired with the same device for each access

Face quality measures (14 in total):

- Face detection reliability, Brightness, Contrast, Focus, Bits per pixel, Spatial resolution, Illumination, Uniform Background, Background Brightness, Reflection, Glasses, Rotation in plane, Rotation in Depth, and Frontalness

Fingerprint quality measure (only one):

- Based on local gradient (minutiae extractability)

Reference systems for matching:

- Face: Omniperception’s Affinity SDK, LDA-based matcher
- Fingerprint: NIST fingerprint system

Protocol:

- DEVELOPMENT: aprox. 50 subjects
- EVALUATION: aprox. 150 subjects
BMEC 2007 Q-based Eval: UAM Fusion

Fusion architecture:

- Face device estimation
- Calibration of face score
- Calibration of fingerprint score
- Fusion of calibrated scores

Log-likelihood ratios:

- $>0$ accept
- $<0$ reject

We choose the score which **stronger** supports the acceptance or rejection decision:

$$\pm \max(|s_{\text{face}}|, |s_{\text{fingerprint}}|)$$

Q-based + handling missing data!!

Julian Fierrez  Multimodal Biometric Databases  61/68

---

BMEC 2007 Q-based Eval: UAM Fusion

**Linear Logistic Regression fusion:**

$$f = a_0 + a_1s_1 + \cdots + a_Ns_N$$

$s = (s_1, \ldots, s_N)$ → scores of individual systems

$\{a_0, a_1, \ldots, a_N\}$ → weights trained by linear logistic regression*, solving (conjugate gradient algorithm):

$$\arg \min_{a_0, \ldots, a_N} = \frac{1}{N_u} \sum_{i=1}^{N_u} \log(1 + e^{-f_i}) + \frac{1}{N_i} \sum_{j=1}^{N_i} \log(1 + e^{-f_j})$$

$N_u, N_i$: number of user and impostor training scores

$f_u, f_i$: fused user and impostor training scores

**Score normalization property:** fused scores $\rightarrow$ log-likelihood ratios (LLR):

$$f \approx \log \left( \frac{p(s | \text{genuine})}{p(s | \text{impostor})} \right)$$

when $N = 1 \rightarrow$ score normalization of a given system


Julian Fierrez  Multimodal Biometric Databases  62/68
BMEC 2007 Q-based Eval: Results

Fusion performance (EER)

Fusion performance (DET curve)
Challenges & Future Directions

- Privacy concerns and legal issues: consent agreement with subjects, notification and registration of the data, national regulations, IPR issues, revocation
- Aging in biometrics, long-term acquisition
- Unsupervised acquisition
- Evaluation on multimodal databases (NIST MBGC)
- Database availability
- Fully realistic conditions, pervasive computing
- Long-term storage, security concerns
- ID managing, anonymity
- Biometrics for user-convenience applications
- Compatibility between datasets
Potential Research on MM DBs

- Individual modalities
- Fusion and multi-biometric systems
- Aging and template update
- Biometric quality and its application in multi-biometrics
- Demographic groups (age groups, gender, ...)
- Sensor interoperability
- Attacks and countermeasures
- Template security