# Fingerprint Databases and Evaluation

Fernando Alonso-Fernandez and Julian Fierrez

Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid, Campus de Cantoblanco, Madrid 28049, Spain
{fernando.alonso, julian.fierrez}@uam.es

## Synonyms

Fingerprint corpora; Fingerprint benchmark

## Definition

Fingerprint databases are structured collections of fingerprint data mainly used for either evaluation or operational recognition purposes.

The fingerprints in databases for evaluation are usually detached from the identity of the corresponding individuals, are publicly available for research purposes, and usually consist of raw fingerprint images acquired with live-scan sensors or digitized from inked fingerprint impressions on paper. These databases are the basis for research in automatic fingerprint recognition, and together with specific experimental protocols, are the basis for a number of technology evaluations and benchmarks. This is the type of fingerprint databases further developed here.

On the other hand, fingerprint databases for operational recognition are typically proprietary, usually incorporate personal information about the enrolled people together with the fingerprint data, and can incorporate either raw fingerprint image data or some form of distinctive fingerprint descriptors such as minutiae templates. These fingerprint databases represent one of the modules in operational automated fingerprint recognition systems, and will not be addressed here.

## Main Body Text

### Fingerprint databases for evaluation

Among all biometric techniques, fingerprint recognition is the most widespread in personal identification due to its permanence and uniqueness [1]. Fingerprints are being increasingly used not only in forensic investigations, but also in a large number of convenience applications, such as access control or online identification [2].

The growth that the field has experienced over the past two decades has led to the appearance of increasing numbers of biometric databases for research and evaluation purposes, either **monomodal** (one biometric trait sensed) or **multimodal** (two or more biometric traits sensed). Previous to the databases acquired within the framework of the International Fingerprint Verification Competition series, the only large, publicly available datasets were the NIST databases [3]. However, these databases were not well suited for the evaluation of algorithms operating with live-scan images [1] and will not be described here. In this section, we present some of the most popular publicly-available biometric databases, either monomodal or multimodal, that include the fingerprint trait acquired with **live-scan sensors**.

*FVC Databases.*

Four international Fingerprint Verification Competitions (FVC) have been organized in 2000, 2002, 2004 and 2006 [4, 5, 6, 7]. For each competition, four databases were acquired using three different sensors and the SFinGE synthetic generator [1]. Each database has 110 fingers (150 in FVC2006) with eight impressions per finger (twelve in FVC2006), resulting in 880

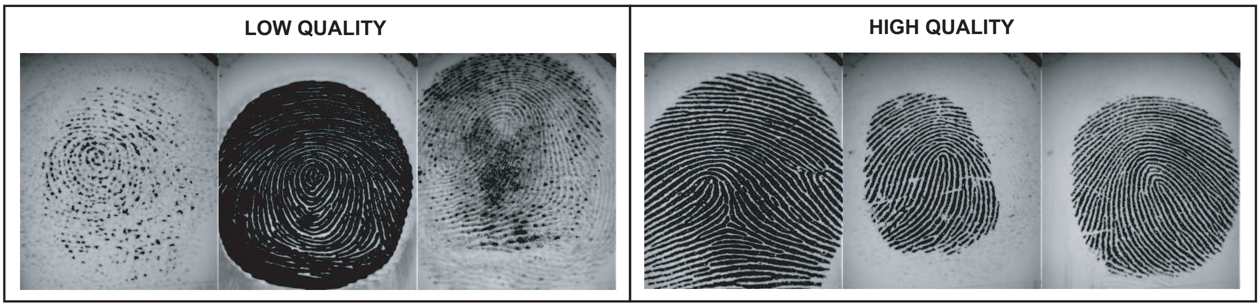

**Fig. 1.** Examples of quality degradation in fingerprint images due to factors like low/high pressure, dryness/moisture, dirt, etc.

impressions (1800 in FVC2006). In the four competitions, the SFinGe synthetic generator was tuned to simulate the main perturbations introduced in the acquisition of the three real databases.

- In FVC2000 [4], the acquisition conditions were different for each database (e.g. interleaving/not interleaving the acquisition of different fingers, periodical cleaning/no cleaning of the sensor). For all the databases, no care was taken to assure a minimum quality of the fingerprints; in addition, a maximum rotation and a non-null overlapping area were assured for impressions from the same finger.
- In FVC2002 [5], the acquisition conditions were the same for each database: interleaved acquisition of different fingers to maximize differences in finger placement, no care was taken in assuring a minimum quality of the fingerprints and the sensors were not periodically cleaned. During some sessions, individuals were asked to: $i)$ exaggerate displacement or rotation or, $ii)$ have their fingers dried or moistened.
- The FVC2004 databases [6] were collected with the aim of creating a more difficult benchmark because, in FVC2002, top algorithms achieved accuracies close to 100 percent [6]. Therefore, more intra-class variation was introduced. During the different sessions, individuals were asked to: $i)$ put the finger at slightly different vertical position, $ii)$ apply low or high pressure against the sensor, $iii)$ exaggerate skin distortion and rotation, and $iv)$ have their fingers dried or moistened. No care was taken to assure a minimum quality of the fingerprints and the sensors were not periodically cleaned. Also, the acquisition of different fingers were interleaved to maximize differences in finger placement. Effects of quality degradation in fingerprint images can be observed in Figure 1.
- For the 2006 edition [7], no deliberate difficulties were introduced in the acquisition as it was done in the previous editions (such as exaggerated distortion, large amounts of rotation and displacement, wet/dry impressions, etc.), but the population was more heterogeneous, including manual workers and elderly people. Also, no constraints were enforced to guarantee a minimum quality in the acquired images and the final datasets were selected from a larger database (the BioSec multimodal database [8]) by choosing the most difficult fingers according to a quality index, to make the benchmark sufficiently difficult for an evaluation.

*MCYT Bimodal Database.*

A large biometric database acquisition process was launched in 2001 by four Spanish academic institutions within the MCYT project [9]. The MCYT database includes ten-print acquisition (MCYT Fingerprint subcorpus) and on-line signature (MCYT Signature subcorpus) samples of each individual enrolled in the database. A total of 330 individuals were acquired in the four institutions participating in the MCYT project. Regarding the MCYT Fingerprint subcorpus, for each individual, 12 samples of each finger were acquired using an optical and a capacitive sensor. With the aim of including variability in fingerprint positioning on the sensor, the 12 different samples of each fingerprint were acquired under human supervision and considering three different levels of control. For this purpose, the fingerprint core had to be located inside a size-varying rectangle displayed in the acquisition software interface viewer.

*BIOMET Multimodal Database.*

Five different biometric modalities are present in the BIOMET database [10]: audio, face image, hand image, fingerprint and signature. This database was designed with the additional goal of including unusual sensors (face images captured with an infrared camera and with a 3D acquisition system). The database consists of three different acquisition sessions. The number of individuals participating to the collection of the database was 130 for the first session, 106 for the second, and 91 for the last one, resulting in 91 individuals who completed the whole acquisition process. For fingerprint acquisition, an optical

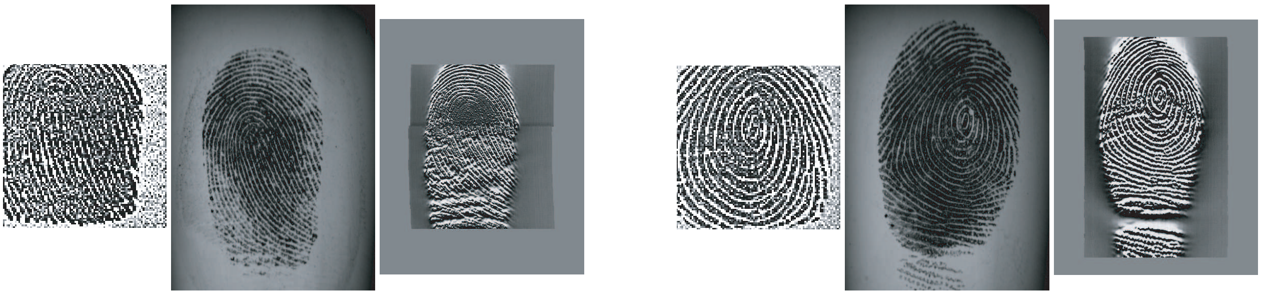

**Fig. 2.** Example fingerprint images of two fingers acquired with three different sensors (from the BioSec baseline corpus). Fingerprint images of the same finger are shown for a capacitive sensor (left of each subplot), an optical sensor (center) and a thermal sensor (right).

and a capacitive sensor were used. During the first acquisition campaign, only the optical sensor was used, whereas both the optical and capacitive sensors were employed for the second ant third campaigns. The total number of available fingerprints per sensor in the BIOMET database is 6 for the middle and index fingers of each contributor.

*BioSec Multimodal Database.*

BioSec was an Integrated Project of the 6th European Framework Programme which involved over 20 partners from 9 European countries. The goal of BioSec was to leverage the integration of biometrics in a wide spectrum of everyday's applications. One of the activities within BioSec was the acquisition of a multimodal database. This database was acquired at four different European sites and includes face, speech, fingerprint and iris recordings. The baseline corpus [8] comprises 200 subjects with 2 acquisition sessions per subject. The extended version of the BioSec database comprises 250 subjects with 4 sessions per subject (about 1 month between sessions). Each subject provided in each session 4 samples of each of 4 fingers (left and right index and middle). Fingerprints were acquired using three different sensors. Some example images are shown in Figure 2.

*BioSecure Multimodal Database.*

The acquisition of the BioSecure Multimodal Database (BMDB) was jointly conducted by 11 European institutions participating in the BioSecure Network of Excellence [11]. The BMDB is comprised of three different datasets [12], namely:

- *Data Set 1 (DS1)*, acquired over the Internet under unsupervised conditions (i.e. connecting to an URL and following the instructions provided on the screen).
- *Data Set 2 (DS2)*, acquired in a standard office room environment using a PC and a number of commercial sensors under the guidance of a human supervisor.
- *Data Set 3 (DS3)*, acquired using two mobile **hand-held devices** under two acquisition conditions (controlled/indoor and uncontrolled/outdoor).

The three datasets of the BMDB include a common part of audio and video data. Additionally, DS2 includes signature, fingerprint, hand and iris data, and DS3 includes signature and fingerprint data. The three datasets were acquired in two different sessions (approximately 2 months between them). Pending yet to be distributed publicly, the BioSecure multimodal database has approximately 1,000 subjects in DS1, and 700 in DS2 and DS3. Fingerprint data in DS2 were acquired using an optical and a capacitive sensor. Fingerprint data in DS3 were acquired with a PDA, and it is considered degraded condition with respect to DS2, since it was acquired while standing with the PDA in the hand.

## Fingerprint evaluation campaigns

The most important evaluation campaigns carried out in the fingerprint modality are the NIST Fingerprint Vendor Technology Evaluation (FpVTE2003) [13] and the four Fingerprint Verification Competitions (FVC), which took place in 2000 [4], 2002 [5], 2004 [6] and 2006 [7]. A comparative summary between FVC2004, FVC2006 and FpVTE2003 is given Table 1. An important evaluation is also the NIST Minutiae Interoperability Exchange Test (MINEX) [14].

*Fingerprint Verification Competitions (FVC).*

| | FVC 2004 | FVC 2006 | FpVTE 2003 |
|---|---|---|---|
| **Participants** | 43 | 53 | 18 |
| **Algorithms** | Open Category: 41<br>Light Category: 26 | Open Category: 44<br>Light Category: 26 | Large Scale Test (LST): 13<br>Medium Scale Test (MST): 18<br>Small Scale Test (SST): 3 |
| **Population** | Students | Heterogeneous (including manual workers and elderly people) | Operational data from a variety of U.S. Government sources |
| **Fingerprint format** | Flat impressions from low-cost scanners | Flat impressions from low-cost scanners | Mixed formats (flat, slap and rolled) from various sources (paper cards, scanners) |
| **Perturbations** | Deliberately exaggerated perturbations | Selection of the most difficult images according to a quality index | Intrinsic low quality fingers and/or non-cooperative users |
| **Data collection** | Acquired for this event | From the BioSec database | From existing U.S. Government sources |
| **Database size** | 4 databases, each containing 880 fingerprints from 110 fingers | 4 databases, each containing 1800 fingerprints from 150 fingers | 48105 fingerprints from 25309 subjects |
| **Anonymous participation** | Allowed | Allowed | Not allowed |
| **Best average EER**<br>(over all the databases used) | 2.07 %<br>(Open Category) | 2.16 %<br>(Open Category) | 0.2 %<br>(MST, the closest to the FVC Open Category) |

**Table 1.** Comparative summary between FVC2004, FVC2006 and FpVTE2003.

The Fingerprint Verification Competitions were organized with the aim of determining the state of the art in fingerprint verification. These competitions have received great attention both from academic and commercial organizations, and several research groups have used the FVC datasets for their own experiments later on. The number of participants and algorithms evaluated has increased in each new edition of the FVC. Also, to increase the number of participants, anonymous participation was allowed in 2002, 2004 and 2006. Additionally, the FVC2004 and FVC2006 were subdivided into: $i)$ *open category* and $ii)$ *light category*. The light category aimed at evaluating algorithms under low computational resources, limited memory usage and small template size.

For each FVC competition, four databases were acquired using three different sensors and the SFinGE synthetic generator [1]. The size of each database was set at 110 fingers with 8 impressions per finger (150 fingers with 12 impressions per finger in FVC2006). A subset of each database (all the impressions from 10 fingers) was made available to the participants prior to the competition for algorithm tuning. The impressions from the remaining fingers were used for testing. Once tuned, participants submitted their algorithms as executable files to the evaluators. The executable files were then tested at the evaluator's site and the test data were not released until the evaluation concluded. In order to benchmark the algorithms, the evaluation was divided into: $i)$ **genuine attempts**: each fingerprint image is compared to the remaining images of the same finger, and $ii)$ **impostor attempts**: the first impression of each finger is compared to the first image of the remaining fingers. In both cases, symmetric matches were avoided.

In Table 2, results of the best performing algorithm in each FVC competition are shown. Data in the 2000 and 2002 editions were acquired without special restrictions and, as observed in Table 2, error rates decrease significantly from 2000 to 2002, demonstrating in some sense the maturity of fingerprint verification systems. However, in the 2004 and 2006 editions, it is observed that error rates increase with respect to the 2002 edition due to the deliberate difficulties and/or low quality sources introduced in the data, thus revealing that degradation of quality has a severe impact on the recognition rates.

| **database** | **2000** | **2002** | **2004** | **2006** |
|---|---|---|---|---|
| DB1 | 0.67% | 0.10% | 1.97% | 5.56% |
| DB2 | 0.61% | 0.14% | 1.58% | 0.02% |
| DB3 | 3.64% | 0.37% | 1.18% | 1.53% |
| DB4 | 1.99% | 0.10% | 0.61% | 0.27% |
| average | 1.73% | 0.19% | 2.07% | 2.16% |

**Table 2.** Results in terms of Equal Error Rate (EER) of the best performing algorithm in each of the four databases of the FVC competitions.

*NIST Fingerprint Vendor Technology Evaluation (FpVTE2003).*

The NIST Fingerprint Vendor Technology Evaluation (FpVTE2003) [13] aimed at: $i)$ comparing systems on a variety of fingerprint data and identifying the most accurate systems; $ii)$ measuring the accuracy of fingerprint matching, identification, and verification on actual operational fingerprint data; and $iii)$ determining the effect of a variety of variables on matcher accuracy. Eighteen different companies competed in the FpVTE, and 34 systems were evaluated.

Three separate subtests were performed in the FpVTE2003: $i)$ the Large-Scale Test (LST), $ii)$ the Medium-Scale Test (MST), and $iii)$ the Small-Scale Test (SST). SST and MST tested matching accuracy using individual fingerprints, whereas LST used sets of fingerprint images. The size and structure of each test were designed to optimize competing analysis objectives, available data, available resources, computational characteristics of the algorithms and the desire to include all qualified participants. In particular, the sizes of MST and LST were only determined after a great deal of analysis of a variety of issues. Designing a well-balanced test to accommodate heterogeneous system architectures was a significant challenge.

Data in the FpVTE2003 came from a variety of U.S. Government sources, including low quality fingers of low quality sources. 48,105 sets of flat slap or rolled fingerprint sets from 25,309 individuals were used, with a total of 393,370 fingerprint images. The systems that resulted in the best accuracy performed consistently well over a variety of image types and data sources. Also, the accuracy of these systems was considerably better than the rest of the systems. Further important conclusions drawn from the FpVTE2003 included: $i)$ the number of fingers used and the fingerprint quality had the largest effect on system accuracy; $ii)$ accuracy on controlled data was significantly higher than accuracy on operational data; $iii)$ some systems were highly sensitive to the sources or types of fingerprints; and $iv)$ accuracy dropped as subject age at time of capture increased.

*NIST Minutiae Interoperability Exchange Test (MINEX).*

The purpose of the NIST Minutiae Interoperability Exchange Test (MINEX) [14] was to determine the feasibility of using minutiae data (rather than image data) as the interchange medium for fingerprint information between different fingerprint matching systems, and to quantify the verification accuracy changes when minutiae from dissimilar systems are used for matching fingerprints. **Interoperability** of templates is affected by the method used to encode minutiae and the matcher used to compare the templates. There are different schemes for defining the method of locating, extracting, formatting and matching the minutiae information from a fingerprint image [1]. In the MINEX evaluation, proprietary template formats were compared to the ANSI INCITS 378-2004 template standard [15].

The images used for this test came from a variety of sensors, and included both live-scanned and non live-scanned rolled and plain impression types. No latent fingerprint images were used. Participants submitting a system had to provide an algorithm capable of extracting and matching a minutiae template using both their proprietary minutiae format and the ANSI INCITS 378-2004 minutiae data format standard [15]. The most relevant results of the MINEX evaluation are:

- In general, proprietary templates lead to better recognition performance than the ANSI INCITS 378-2004 template.
- Some template generators produce standard templates that are matched more accurately than others. Some matchers compare templates more accurately than others. The leading vendors in generation are not always the leaders in matching and vice-versa.
- Authentication accuracy of some matchers can be improved by replacing the vendors template generator with that from another vendor.
- Performance is sensitive to the quality of the dataset. This applies to both proprietary and interoperable templates. Higher quality datasets provide reasonable interoperability, whereas lower quality datasets do not.

## Related Entries

Biometric Acquisition, Fingerprint Device, Interoperability Performance, Performance Evaluation

## References

1. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: *Handbook of Fingerprint Recognition*, Springer, New York (2003)
2. Jain, A., Ross, A., Pankanti, S.: Biometrics: A tool for information security, *IEEE Trans. Information Forensics and Security*, **1** (2002) 125–143
3. NIST Special Databases and Software from the Image Group, http://www.itl.nist.gov/iad/894.03/databases/defs/dbases.html

4. Maio, D., Maltoni, D., Capelli, R., Wayman, J., Jain, A.: FVC2000: Fingerprint Verification Competition, *IEEE Trans. Pattern Anal. Mach. Intell.*, **24** (2002) 402–412
5. Maio, D., Maltoni, D., Capelli, R., Wayman, J., Jain, A.: FVC2002: Second Fingerprint Verification Competition, *Proc. Intl. Conf. Pattern Recognition*, **3** (2002) 811–814
6. Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., Jain, A.K.: Performance evaluation of fingerprint verification systems, *IEEE Trans. Pattern Anal. Mach. Intell.*, **28** (2006) 3–18
7. FVC2006, 2006. Fingerprint Verification Competition, http://bias.csr.unibo.it/fvc2006/default.asp.
8. Fierrez, J., Ortega-Garcia, J., Torre-Toledano, D., Gonzalez-Rodriguez, J.: BioSec baseline corpus: a multimodal biometric database, *Pattern Recognition*, **40** (2007) 1389–1392
9. Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J., Vivaracho, C., Escudero, D., Moro, Q.: MCYT baseline corpus: a bimodal biometric database, *IEE Proceedings on Vision, Image and Signal Processing*, **150** (2003) 395–401
10. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., les Jardins, J., Lunter, J., Ni, Y., Petrovska-Delacretaz, D.: BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities, *Lect. Notes in Computer Science*, **2688** (2003) 845853
11. BioSecure, 2004. Biometrics for Secure authentication, FP6 NoE, IST - 2002-507634 (http://www.biosecure.info)
12. Alonso-Fernandez, F., Fierrez, J., Ramos, D., Ortega-Garcia, J.: Dealing with sensor interoperability in multi-biometrics: the UPM experience at the BioSecure Multimodal Evaluation 2007, *Proc. SPIE*, **6944** (2008)
13. Wilson, C., *et al*: Fingerprint Vendor Techonology Evaluation 2003: Summary of results and analysis report. NISTIR 7123, http://fpvte.nist.gov (2004)
14. Grother, P., *et al.*: MINEX - Performance and interoperability of the INCITS 378 fingerprint template, NISTIR 7296, http://fingerprint.nist.gov/minex (2005)
15. ANSI-INCITS 378, Fingerprint Minutiae Format for Data Interchange, American National Standard, 2004

## Definitional Entries

### Monomodal/multimodal database

A monomodal database is a database which only has one biometric trait sensed. A multimodal database is a database which has more that one biometric trait from the same individual.

### Live-scan sensor

A live-scan sensor is a sensor that allows to capture and digitize biometric data in real time. As opposed to live-scan acquisition, in off-line acquisition, data is not digitized in real time (e.g. when a fingerprint image is first obtained by smearing ink on the fingertip and creating an inked impression on paper, and then the inked impression is digitized by scanning the paper).

### Hand-held devices

A hand-held device is a pocket-sized computing device, typically comprising a small visual display screen for user output and a miniature keyboard or touch screen for user input. New hand-held devices include a number of sensors that can be used to acquire biometric data, e.g.: touch-screens (signature and handwritting), fingerprint sensors, microphones (speech), cameras (face, video), etc.

### Genuine/impostor attempt

In a genuine attempt, a biometric sample is compared against other biometric samples from the same subject. If similarity between the samples is not high enough, the subject will be wrongly rejected by the system. In an impostor attempt, a biometric sample is compared against biometric samples from other subjects. If similarity between the samples is high enough, the subject will be wrongly accepted by the system. It should be noted that biometric samples from the same user are not necessarily similar (e.g. temporary injuries in the finger) and on the other hand, biometric samples from different users can be quite similar (e.g. signature forgeries).

### Interoperability

Interoperability in biometrics can be defined as the capability of a recognition system to operate with data from different sources (e.g. data acquired using different sensors or features extracted using systems from different vendors). Most biometric systems are designed under the assumption that the data to be compared are obtained from a unique source and are restricted in their ability to match or compare biometric data originated from different sources. As a result, changing the source may affect the performance of the system.