Encyclopedia of Biometrics
DOI 10.1007/978-3-642-27733-7_61-2

Fingerprint Databases and Evaluation

Fernando Alonso-Fernandez*[a] and Julian Fierrez[b]
[a]Intelligent Systems Lab (IS-Lab/CAISR), Halmstad University, Halmstad, Sweden
[b]Biometric Recognition Group (ATVS), Universidad Autonoma de Madrid, Madrid, Spain

Synonyms

Fingerprint benchmark; ▸ Fingerprint corpora; ▸ Fingerprint dataset

Definition

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

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

On the other hand, fingerprint databases for operational recognition are typically proprietary, they usually incorporate personal information about the enrolled people together with the fingerprint data, and they 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 they will not be addressed here.

Fingerprint Databases for Evaluation

Among all biometric techniques, fingerprint recognition is one of the most widespread modalities in personal identification due to its permanence and uniqueness [1]. Nearly all forensics and law enforcement agencies worldwide use Automated Fingerprint Identification Systems (AFIS), and the emergence of low-cost, compact fingerprint readers has made fingerprints the preferred choice in a large number of civil and commercial applications [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

*E-mail: feralo@hh.se

DOI 10.1007/978-3-642-27733-7_61-2

Verification Competition (FVC) 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 they 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]. 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 8 impressions per finger (12 in FVC2006), resulting in 880 impressions (1,800 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, 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, 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 exaggerate displacement or rotation and to have their fingers dried or moistened.
- The FVC2004 databases were collected with the aim of creating a more difficult benchmark because, in FVC2002, top algorithms achieved accuracies close to 100 %. Therefore, more intra-class variation was introduced. During the different sessions, individuals were asked to put the finger at slightly different vertical positions, to apply low or high pressure against the sensor, to exaggerate skin distortion and rotation, and to 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 was interleaved to maximize differences in finger placement. Effects of quality degradation in fingerprint images can be observed in Fig. 1.
- For the 2006 edition, 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 [5]) by choosing the most difficult fingers according to a quality index, to make the benchmark sufficiently difficult for an evaluation.

DOI 10.1007/978-3-642-27733-7_61-2

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

MCYT Bimodal Database

A large biometric database acquisition process was launched in 2001 by four Spanish academic institutions within the MCYT project [6]. The database includes ten-print acquisition (MCYT Fingerprint subcorpus) and online 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 project. Regarding the MCYT Fingerprint subcorpus, for each individual, 12 samples of each finger were acquired using an optical and a capacitive sensor, resulting in $330 \times 12 \times 10 = 39{,}600$ captured images per 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 3 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 [7]: 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 sensor 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 and 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 (IP) 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 was the acquisition of a multimodal database. This database was acquired at four different European sites, and it

DOI 10.1007/978-3-642-27733-7_61-2

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*)

includes face, speech, fingerprint, and iris recordings. The baseline corpus [5] comprises 200 subjects with 2 acquisition sessions per subject, whereas the extended version of the BioSec database comprises 250 subjects with 4 sessions per subject (about 1 month between sessions). Each subject provided 4 samples of each of 4 fingers (left and right index and middle) per session. Fingerprints were acquired using three different sensors. The total number of fingerprint images per sensor are therefore $200 \times 2 \times 4 \times 4 = 6{,}400$ (baseline corpus) and $250 \times 4 \times 4 \times 4 = 16{,}000$ (extended version). Some example images are shown in Fig. 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 of the 7th European Framework Programme. The BMDB is comprised of three different datasets [8], namely:

- *Dataset 1 (DS1)*, acquired over the Internet under unsupervised conditions (i.e., connecting to a URL and following the instructions provided on the screen)
- *Dataset 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
- *Dataset 3 (DS3)*, acquired using two mobile **handheld 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). The BioSecure Multimodal Database has 971 subjects in DS1, 667 in DS2, and 713 in DS3. Fingerprint data in DS2 were acquired using an optical and a thermal 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. In all cases, each subject provided 2 samples of each of 6 fingers (left and right thumb, index, and middle), therefore contributing with $6 \times 2 = 12$ fingerprint samples per sensor and per session.

DOI 10.1007/978-3-642-27733-7_61-2

CASIA Fingerprint Image Database

The CASIA fingerprint database (currently version 5) has been acquired by the Institute of Automation, Chinese Academy of Sciences (CASIA) [9]. It contains 20,000 fingerprint images of 500 subjects captured using an optical fingerprint sensor in one session. Each volunteer contributed with 5 samples of each of 8 fingers (left and right thumb, index, middle, and ring), totaling 40 samples per person. The volunteers were asked to rotate their fingers with various levels of pressure to generate significant intra-class variations.

Fingerprint Evaluation Campaigns

The most important evaluation campaigns in the fingerprint modality are the series of Fingerprint Verification Competitions (FVC) [4] and the different evaluations carried out by the US National Institute of Standards and Technology (NIST) [10]. The Fingerprint Verification Competitions took place in 2000, 2002, 2004, and 2006. Since 2009, a new online evaluation campaign, FVC-onGoing, offers web-based automated evaluation of fingerprint recognition algorithms, where participants can upload algorithms at any time and automatically obtain performance results. The US National Institute of Standards and Technology has also conducted several biometric evaluation campaigns in the last decade, not only in the fingerprint modality, but also in others. As for the fingerprint modality, they include the series of the Fingerprint Vendor Technology Evaluation (FpVTE2003, FpVTE2012), the Proprietary Fingerprint Template Evaluations (PFT2003, PFTII-2011), and the Minutiae Interoperability Exchange Test (MINEX2004, MINEXII-2007). Other evaluation series, not covered here, are the Evaluation of Latent Fingerprint Technologies (ELFT2007, ELFT-EFS2009).

Fingerprint Verification Competitions (FVC)

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, 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. In Table 1, results of the best-performing algorithm in each FVC are shown. Data in the 2000 and 2002 editions were acquired without special restrictions and, as observed in Table 1, 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

DOI 10.1007/978-3-642-27733-7_61-2

Table 1 Results in terms of Equal Error Rate (EER) of the best performing algorithm in each of the four databases of the FVC

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

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.

Since 2009, the FVC have been substituted by the web-based automated evaluation FVC-onGoing. Here, participants can upload algorithms at any time and obtain performance results automatically. The system is always open to new participants, and the participant can decide to publish the results of its algorithms on the public result section. FVC-onGoing provides various benchmarks to evaluate fingerprint algorithms, and each benchmark is based on a sequestered dataset that will not evolve over time. The addition of new evaluation benchmarks over time is also contemplated. Currently, FVC-onGoing allows the evaluation of the following six subproblems: one-to-one fingerprint verification, one-to-one palmprint verification, one-to-one fingerprint matching of ISO minutia-based template format (ISO/IEC 19794-2) [11], fingerprint indexing over a large database, fingerprint orientation extraction, and one-to-one fingerprint verification using protected templates.

NIST Fingerprint Vendor Technology Evaluation (FpVTE)

The NIST Fingerprint Vendor Technology Evaluation is a one-to-many fingerprint evaluation whose first edition was conducted in 2003 (FpVTE2003) [12]. The second FpVTE edition (FpVTE2012) is currently underway, with results to be published by mid-2013.

FpVTE consists of multiple tests performed with combinations of fingers (e.g., single fingers, two index fingers, four to ten fingers) and different types and qualities of operational fingerprints (e.g., rolled and flat inked fingerprints, multi-finger flat live-scan images, and single flat live-scan images). Data in FpVTE come from a variety of US government sources, including low-quality fingers of low-quality sources. Rolled fingerprints are captures obtained by rolling the full finger from side to side, whereas flat fingerprints are captured by pressing the finger against the sensor. Multi-finger flat sensors capture the four left/right fingers all at the same time (the thumb can also be included in newer sensors), and single-finger flat sensors only allow captures of individual fingers. Multi-finger flat captures are not segmented in FpVTE; therefore, participants are required to implement this step in their submissions.

Data in FpVTE2003 comprised 48,105 sets of flat slap or rolled fingerprint sets from 25,309 individuals, 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

DOI 10.1007/978-3-642-27733-7_61-2

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.

The plan for FpVTE2012 is to use enrolled sample sizes extending into the multiple millions, in accordance with the current requirements of one-to-many large-scale applications [2]. Another goal is to enable evaluation on operational datasets captured with newer live-scan ten-print sensors, as well as the cross-comparison (sensor **interoperability**) with live-scan single- and multi-finger sensors, and the historically significant scanned inked fingerprints. FpVTE2012 also contemplates the possibility of using data from mobile devices. Results are expected to be published by mid-2013.

NIST Proprietary Fingerprint Template Evaluations (PFT)

The Proprietary Fingerprint Template evaluations is a program aimed at measuring the performance of fingerprint matching software by using vendor proprietary fingerprint templates. Unlike the FpVTE program, PFT is intended to assess the core algorithmic capability of the technology in one-to-one verification. Also, PFT evaluations are ongoing and new SDKs can be tested at any time. There is also one additional difference between the PFT tests and FpVTE. In FpVTE, each vendor is provided with a set of test images and asked to return the matching scores, so testing is run on the vendor's own hardware and using its own matching software. In PFT, on the other hand, vendors supply their SDK libraries to NIST, so matching is run on NIST hardware.

The first PFT evaluation started in 2003 and concluded in 2010. The last published report is from 2005 [13], although updated plots and performance figures are reported in the NIST evaluation web site until 2010 [10]. Data comprised samples of 5,888 subjects from different US government sources, and they included rolled and plain fingerprints from inked paper and live-scan devices. The datasets were grouped by finger position, so only the right index is compared to the right index and so on for other fingers. Considering all the datasets and fingers, the total number of matches performed by each SDK was 758,638,238. Results showed that the most accurate SDKs were consistent across all datasets while others had a wide range of results depending on which dataset was being used. This effect was also seen in the FpVTE tests [12]. Not surprisingly, the datasets perceived to have better quality performed better than the other datasets. Also as a general observation, as matcher performance increased, speed decreased. If the data quality is good enough, however, a faster matcher could do almost as well as the slower matchers. Performance of vendors that participated both in FpVTE2003 and PFT2003 was also compared, with most of them having similar performance and ranking on both tests.

The newer PFTII evaluation started in 2011 and it is currently accepting SDKs for evaluation. A first report has been released with the results of ongoing experiments [14]. The original PFT2003 only reported the matching algorithm's accuracy. The PFTII-2011 evaluation will report, in addition, template extraction times, template size information, and matcher timings. The sample dataset sizes have been increased to 120,000 subjects. Two of the fingerprints datasets from PFT-2003 will be used in the ongoing PFTII-2011, but with added samples. A new dataset with ten-print rolled images will also be included. Also, the fingerprint matches will be performed between different types of print impressions: plain vs. plain images, plain vs. rolled images, and rolled vs. rolled images.

Encyclopedia of Biometrics
DOI 10.1007/978-3-642-27733-7_61-2

NIST Minutiae Interoperability Exchange Test (MINEX)

MINEX is series of NIST-coordinated development efforts aimed at improving the performance and interoperability of implementations of the INCITS 378 and ISO/IEC 19794-2 fingerprint minutia standards [11, 15]. Minutiae data (rather than image data) is used in MINEX as the interchange medium for fingerprint information between different fingerprint matching systems. There are different schemes for defining the method of locating, extracting, formatting, and matching the minutiae information from a fingerprint image [1], and the **interoperability** of templates is affected by the method used to encode minutiae and the matcher used to compare the templates.

The first edition of MINEX in 2004 was intended to assess the viability of the INCITS 378 template as the interchange medium for fingerprint data [16]. Specific objectives of MINEX2004 included the following: to determine if standardized minutiae templates can be matched against templates extracted by another vendor and to estimate the verification accuracy when standardized templates are compared to proprietary templates. This way, proprietary and standard template formats were compared, and verification accuracy changes when minutiae from dissimilar systems are used for matching fingerprints were also quantified. 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 INCITS 378 minutiae data format standard. The most relevant results of the MINEX2004 evaluation were:

- In general, proprietary templates lead to better recognition performance than the INCITS 378 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 vendor's 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.

The second MINEX edition in 2007 maintained the specific objectives of MINEX2004. The main difference is that matching in MINEXII-2007 was done on ISO/IEC 7816 smart cards and the data template was the ISO/IEC 19794-2 compact card fingerprint minutia standard [17]. The MINEXII-2007 evaluation spanned from 2007 to 2010, with several rounds of testing during this period. Based on the results, match-on-card remains a difficult tasks, and the porting of algorithms running on general-purpose computers to smart cards is not trivial. However, during the 3 years spanned by MINEXII, several providers of match-on-card implementations showed considerable improvements in both accuracy and speed, demonstrating significant progress toward the possibility of applications involving fast and accurate matching of compact biometric templates.

DOI 10.1007/978-3-642-27733-7_61-2

Related Entries

- Biometric Data Acquisition
- Fingerprint Device
- Interoperability Performance
- Performance Evaluation

References

1. D. Maltoni, D. Maio, A. Jain, S. Prabhakar, *Handbook of Fingerprint Recognition*, 2nd edn. (Springer, New York, 2009)
2. A.K. Jain, A. Kumar, Biometrics of next generation: an overview, in *Second Generation Biometrics* (Springer, Heidelberg, 2010)
3. NIST special databases and software from the image group, http://www.nist.gov/itl/iad/ig/special_dbases.cfm
4. FVC-onGoing, On-line evaluation of fingerprint recognition algorithms (2009), https://biolab.csr.unibo.it/fvcongoing
5. J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, J. Gonzalez-Rodriguez, BioSec baseline corpus: a multimodal biometric database. Pattern Recognit. **40**, 1389–1392 (2007)
6. J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. Igarza, C. Vivaracho, D. Escudero, Q. Moro, MCYT baseline corpus: a bimodal biometric database. IEE Proc. Vis. Image Signal Process. **150**, 395–401 (2003)
7. S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. les Jardins, J. Lunter, Y. Ni, D. Petrovska-Delacretaz, BIOMET: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities, in *Proceedings of the International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA*, Guildford, UK 2003, pp. 845–853
8. J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J. Alba-Castro, E. Gonzalez-Agulla, E. Otero-Muras, S. Garcia-Salicetti, L. Allano, B. Ly-Van, B. Dorizzi, J. Kittler, T. Bourlai, N. Poh, F. Deravi, M. Ng, M. Fairhurst, J. Hennebert, A. Humm, M. Tistarelli, L. Brodo, J. Richiardi, A. Drygajlo, H. Ganster, F. Sukno, S. Pavani, A. Frangi, L. Akarun, A. Savran, The multi-scenario multi-environment biosecure multimodal database (BMDB). IEEE Trans. Pattern Anal. Mach. Intell. **32**, 1097–1111 (2010)
9. BIT (Biometrics Ideal Test), http://biometrics.idealtest.org
10. NIST Biometric Evaluations Homepage, http://www.nist.gov/itl/iad/ig/biometric_evaluations.cfm
11. ISO/IEC 19794-2 biometric data interchange formats – part 2: finger minutiae data. JTC1/SC37 biometrics (2011), http://isotc.iso.org/isotcportal
12. C. Wilson, R. Hicklin, H. Korves, B. Ulery, M. Zoepfl, M. Bone, P. Grother, R. Micheals, S. Otto, C. Watson, Fingerprint vendor techonology evaluation 2003: summary of results and analysis report. NISTIR 7123 (2004), http://www.nist.gov/itl/iad/ig/fpvte03.cfm
13. C. Watson, C. Wilson, K. Marshall, M. Indovina, R. Snelick, Studies of one-to-one fingerprint matching with vendor SDK matchers. NISTIR 7221 (2005), http://www.nist.gov/itl/iad/ig/pft.cfm

DOI 10.1007/978-3-642-27733-7_61-2

14. S.L. Cheng, G. Fiumara, C. Watson, PFTII report. Plain and rolled fingerprint matching with proprietary templates. NISTIR 7821 (2011), http://www.nist.gov/itl/iad/ig/pft.cfm
15. ANSI-INCITS 378, *Fingerprint Minutiae Format for Data Interchange* (American National Standard, New York, 2004)
16. P. Grother, M. McCabe, C. Watson, M. Indovina, W. Salamon, P. Flanagan, E. Tabassi, E. Newton, C. Wilson, MINEX – performance and interoperability of the INCITS 378 fingerprint template. NISTIR 7296 (2005), http://fingerprint.nist.gov/minex
17. P. Grother, W. Salamon, C. Watson, M. Indovina, P. Flanagan, MINEX II – performance of fingerprint match-on-card algorithms. Phase IV report. NISTIR 7477 (2011), http://fingerprint.nist.gov/minex