

Open Source Reference Systems for Biometric Verification of Identity

Aurélien Mayoue and Dijana Petrovska-Delacrétaz
TELECOM & Management SudParis,
9 rue Charles Fourier,
91011 Evry Cedex, France

Abstract. This paper focuses on the common evaluation framework which was developed by the BioSecure Network of Excellence during the European FP6 project BioSecure (Biometrics for Secure authentication). This framework, which is composed of open-source reference systems, publicly available databases, assessment protocols and benchmarking results, introduces a new experimental methodology for conducting, reporting and comparing experiments for biometric systems, participating to standardisation efforts. Its use will permit to make a link between different published works. It provides also the necessary tools to assure the reproducibility of the benchmarking biometric experiments. This framework can be considered as a re-liable and innovative process to evaluate the progress of research in the field of bio-metrics.

1 Introduction

With the widespread use of biometric systems in our society, a lot of articles about this topic have been published in the last years. Among all the algorithms presented in the literature, it seems essential to know which one performs better, hence the need for a common evaluation framework which would enable a fair comparison between biometric algorithms. Indeed, the performance of two biometric verification systems can only be fairly compared in the case where these both systems have been evaluated in the same experimental conditions, i.e. applied on the same database according to the same protocol. That's why several biometric evaluation campaigns have been launched during the last years. But in the existing ongoing evaluation programs, such as NIST-SRE [1] for speaker recognition, BANCA [2], M2VTS [3] and FRVT [4] for face, FVC [5] for fingerprint, SVC [6] for signature and ICE [7] for iris, the databases used are often not publicly available and there is no systematic usage of a state-of-the-art algorithm to calibrate both the performance of the evaluated systems and the 'difficulty' of the evaluation database and protocol. Therefore, it is impossible to reproduce the experiments and to measure the progress (in comparison with the current state of the art) done by the newly developed research prototypes which are tested during these evaluation campaigns. In this way, such evaluations can not be considered as an universal evaluation tool.

From these observations, the partners of the BioSecure Network of Excellence (NoE) [8] have brought together the expertise in a wide range of biometric modalities in order to make available to the biometric community a new open-source evaluation tool, which enables a fair comparison of the biometric algorithms being tested. In this way, the NoE has taken in charge the development or improvement of existing open-source identity verification systems (softwares) for the most widely used biometric modalities (2D and 3D faces, fingerprint, hand, iris, signature, speech and talking-face). These open-source systems are based on principles that have proven their efficiency and their purpose is to serve as comparison point to measure the real progress achieved with new research methods. Consequently, these open-source systems are known as reference systems. These systems accompanied with benchmarking (calibration) publicly available databases and benchmarking (calibration) assessment protocols define the building blocks of the evaluation framework, which also reports benchmarking results (i.e. results obtained by applying the reference system on the benchmarking database according to the benchmarking protocol).

From now on, any research laboratory can evaluate their own biometric algorithms using the benchmarking database in accordance with the benchmarking protocol and then compare their results to the performance of the reference system. This is a new way to calibrate the performance of an algorithm.

The outline of this paper is as follows. In Sect. 2, the objectives of the proposed evaluation framework are developed. Then, the open-source reference systems, which compose this framework, are briefly presented in Sect. 3 whereas in Sect. 4 we explain how to use the evaluation tool in practice. Section 5 ends up with conclusions and perspectives.

2 Objectives

2.1 Properties of the evaluation framework

In the framework of identity verification, it is very difficult to compare two different methods from two different articles in the literature, even though they deal with the very same task. It raises a real problem when one wants to know if a new original method performs better than the current state of the art for example. This can be explained by the fact that a lot of researchers have recorded their own test data-base and are the only one performing experiments on it, which are subsequently impossible to reproduce.

Our evaluation framework brings an easy yet efficient answer to this problem. As it is composed (for each modality) of a reference system, a publicly available database and an evaluation protocol accompanied with a documentation which reports performance of the reference system, our evaluation framework can be used in two ways:

- in the case where someone publishes results on a specific database according to a specific protocol, experiments using the reference system have to be added as a way of calibrating the difficulty of this particular task.
- in the case where someone publishes results using the benchmarking data-base and protocol, the benchmarking results of the reference system can be used as a way of calibrating the performance.

In both cases, the results of the reference systems permit to measure the progress achieved with the newly developed algorithm being tested. Now, it is hoped that the published results of a new biometric algorithm will be systematically accompanied with the results obtained by the reference system in the same experimental conditions. Such a methodology for reporting experimental results will facilitate deeply advances of the state of the art.

Through our framework, an absolute reproducibility of the benchmarking results is also assured. Indeed, an explanatory documentation is available for each modality to avoid any possible misinterpretation of the proposed evaluation framework. In this way, the properties of our framework are the following:

- its universal use: for each modality, it is comprised of an open-source reference system, a publicly available database and a well defined evaluation protocol.
- it enables a fair comparison of the biometric algorithms being tested.
- it evaluates the real progress achieved with new research prototypes.
- its easy use: for each modality, a document explains in details the evaluation protocol and how to install and use the reference system.
- the full reproducibility of the benchmarking results.

2.2 Relevance of using open-source reference systems

The fact that the proposed reference systems are open-source and made of replaceable C/C++ software modules with well defined inputs/outputs is of great interest. Indeed, researchers often work on a specific part of the system and do not have the time nor the interest in building a complete system. In this way, a researcher could show the improvement of his new features extraction algorithm (for example) simply by replacing the corresponding module in the reference system and without having to bother about the pattern recognition algorithm. This new methodology for conducting research and experiments will enable to assess the contribution of each module to the final performance of a biometric system.

Using an open-source reference system as a basis for researching a specific area is also a good way to save time, human resources and money. For example, young researchers who are just starting in the field of biometrics, could immediately be faced with the reference system and they could put all their efforts to improve our efficient system rather than developing a state-of-the-art algorithm from A to Z again. In this way, the main advantages of our reference systems are the following:

- development as open-source software.

- easy implementation: in general, a reference system is composed of four replaceable modules (pre-processing, features extraction, model building and matching).
- time saving: a reference system can be considered as a starting point for the development of a more efficient system.

3 The open-source Reference Systems

At time of writing, the evaluation framework concerns eight biometric modalities. It is comprised of five publicly available databases and ten benchmarking open-source reference systems. Among these systems, eight were developed within the framework of the BioSecure NoE and only three were existing ones (the fingerprint and speech reference systems) but all systems are accompanied with a benchmarking database, assessment protocol and benchmarking results to compose our evaluation framework:

1. For the 2D face modality, the BioSecure reference system is based on the Eigenface approach [9] and the BANCA database [10] is chosen for the benchmarking experiments.
2. For the 3D face modality, the BioSecure reference system [11] uses both Iterative Closest Points [12] and Thin Plate Spline warping [14] algorithms. The associated database is the 3D RMA database [13].
3. For the fingerprint modality, the NIST reference system (NFIS2 [15]) uses a standard minutiae based approach. MCYT-100 [16] is selected as the benchmarking database.
4. For the hand modality, the BioSecure reference system [17] uses geometry of fingers for recognition purposes. The selected database is the Biosecure database [17].
5. For the iris modality, the BioSecure reference system (OSIRIS [18]) is deeply inspired by Daugman's works [19] and the database used for the benchmarking experiments is CBS [20].
6. For the signature modality, two reference systems developed by BioSecure members are available. The first one [21] is based on the Hidden Markov Model [22] whereas the second one [23] uses the Levenshtein distance [24]. The MCYT-100 [16] database is chosen for the benchmarking experiments.
7. For the speech modality, the evaluation framework is composed of two open-source softwares: ALIZE [25] and BECARs [27]. Both systems are based on the Gaussian Mixture Models approach [26]. The reference data-base is the speech part of the BANCA database [10].
8. For the talking-face modality, the BioSecure reference system [28] is based on the fusion of face and speech scores. The BANCA database [10] is used for the benchmarking experiments.

Some of the open-source reference systems available in the benchmarking framework have been already used in major existing evaluation campaigns. For example, the ALIZE software [25] is consistently used during the NIST-SRE [1] and it exhibits state-of-the-art performance. The iris reference system [18] has been tested with the ICE'2005 database and it has a quite good ranking compared to the results re-reported in [7]. Endless, the signature reference system based on HMM [21] has obtained good performance at the SVC'2004 competition [6]. More details about performance of all reference systems can be found in [29].

4 Use of the evaluation tool

The source code of the ten reference systems described in Sect. 3 are available at [30]. Scripts (required to run a full experiment with the reference system), lists of tests (impostor and genuine trials) to be done and all the necessary information needed to fully reproduce the benchmarking results are also available at this address. Endless, for each modality, a document explains how to install and use the reference system. In practice, the user of our evaluation framework can be confronted with three different scenarios:

1. He wants to evaluate his new biometric algorithm using our reference data-base and protocol to calibrate the performance of his algorithm. In this case, he has to process as follows:
 - run the evaluated system on the reference database using the list of trials provided within the evaluation framework.
 - compare the obtained results to the benchmarking results provided within the evaluation framework to measure the progress achieved with the new algorithm.
2. He wants to test the reference system on his own database according to his own protocol in order to calibrate this particular task. In this case, he has to process as follows:
 - run the reference system using his own database and protocol.
 - compare the obtained results to those provided within the evaluation framework to evaluate the 'difficulty' of these particular database and protocol.
 - use the obtained results as comparison point for the other experiments carried out in the same conditions.
3. He wants simply to reproduce the benchmarking results to be convinced of the good installation and use of the reference system:
 - run the reference system on the reference database using the scripts and lists of trials provided within the evaluation framework.
 - verify that the obtained results are the same as the benchmarking results.

Since the reference systems are open-source, other scenarios are of course conceivable. Some parameters or even some modules of the reference systems can be changed in order to improve their basic performances.

5. Conclusions and Perspectives

In the field of biometry, the BioSecure NoE has developed an universal evaluation framework based on open-source reference systems, publicly available data-bases, evaluation protocols and benchmarking results. This framework enables a fair comparison of the performance of biometric systems and its widespread use would facilitate the advances of the biometric state of the art.

This framework has been used for the first time during the BioSecure Multimodal Evaluation Campaign (BMEC) 2007. The organizers of this competition have systematically used the reference systems to calibrate the 'difficulty' of the newly constituted BioSecure database and protocol. Then, the results of the reference systems have been used for measuring the improvements acquired with each submitted re-search system. All results of this evaluation are available on the BMEC's web site [31].

As the research advances, the reference systems should be updated in order to remain competitive. The development of reference systems for the newly appeared modalities (such as palm vein) is also part of our future plans.

Acknowledgments

Biosecure is a project of the 6th Framework Programme of the European Union. We thank in particular all the universities which participated in the development of the BioSecure reference systems.

References

1. NIST Speaker Recognition Evaluations. <http://www.nist.gov/speech/tests/spk>.
2. BANCA: Biometric Access control for Networked and e-Commerce Applications. <http://www.ee.surrey.ac.uk/CVSSP/banca/>.
3. M2VTS: Multi Modal Verification for Teleservices and Security applications. <http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>.
4. FRVT: Face Recognition Vendor Test. <http://www.frvt.org/>.
5. FVC: Fingerprint Verification Competition. <http://bias.csr.unibo.it/fvc2006/>.
6. SVC: Signature Verification Competition. <http://www.cse.ust.hk/svc2004/>.
7. ICE: Iris Challenge Evaluation. <http://iris.nist.gov/ice/>.
8. BioSecure Network of Excellence. <http://biosecure.info/>.

9. M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71-86, 1991.
10. E. Bailly-Baillière, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, and J.-P. Thiran. The BANCA Database and Evaluation Protocol. In 4th International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA'03), volume 2688 of Lecture Notes in Computer Science, pages 625-638, Guildford, UK, January 2003. Springer.
11. M.O Irfanoglu, B. Gokberk, and L. Akarun. 3d Shape-based Face Recognition Using Automatically Registered Facial Surfaces. In Proc. 17th International Conference on Pattern Recognition, Cambridge, UK, 2004.
12. P. Besl and N. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239-256, 1992.
13. 3D RMA database. http://www.sic.rma.ac.be/~beumier/DB/3d_rma.html.
14. F.L. Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567-585, 1989.
15. C. Watson, M. Garris, E. Tabassi, C. Wilson, R. McCabe, and S. Janet. Guide to NIST Finger-print Image Software 2 (NFIS2) - <http://fingerprint.nist.gov/NFIS>.
16. J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, M.F.J. Gonzalez, V. Espinosa, A. Satue, I. Hernaez, J. J. Igarza, C. Vivaracho, D. Escudero, and Q. I. Moro. MCYT baseline corpus: A bi-modal biometric database. *IEE Proceedings Vision, Image and Signal Processing, Special Issue on Biometrics on the Internet*, 150(6):395-401, December 2003.
17. G. Fouquier, L. Likforman, J. Darbon, and B. Sankur. The Biosecure Geometry-based System for Hand Modality. In the Proceedings of 32nd International Conference on Acoustics, Speech and Signal Processing (ICASSP), Honolulu, Hawaii, USA, april 2007.
18. E. Krichen. Reconnaissance des personnes par l'iris en mode dégradé. Thèse de doctorat, Institut National des Télécommunications, 2007.
19. J.G. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Patt. Ana. Mach. Intell.*, 15(11):1148-1161, 1993.
20. CASIA iris image database. <http://www.cbsr.ia.ac.cn/IrisDatabase.htm>.
21. B. Ly Van, S. Garcia-Salicetti, and B. Dorizzi. On using the Viterbi Path along with HMM Likelihood Information for On-line Signature Verification. In Proceedings of the IEEE Transactions on Systems, Man and Cybernetics, Part B, to appear.
22. L. Rabiner and B.H. Juang. *Fundamentals of Speech Recognition*. Prentice Hall Signal Processing Series. 1993.
23. S. Schimke, C. Vielhauer, and J. Dittmann. Using Adapted Levenshtein Distance for On-Line Signature Authentication. In Proceedings of the ICPR 2004, IEEE 17th International Conference on Pattern Recognition, ISBN 0-7695-2128-2, 2004.
24. V.I. Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics*, 10:707-710, 1966.
25. ALIZE software. <http://www.lia.univ-avignon.fr/heberges/ALIZE/>.
26. D. A. Reynolds. A Gaussian Mixture Modeling Approach to Text-Independent Speaker Identification. Ph.D. Thesis, Georgia Institute of Technology, 1992.
27. BECARS software. <http://www.tsi.enst.fr/becars/index.php>.
28. Hervé Bredin, Guido Aversano, Chafic Mokbel, and Gérard Chollet. The Biosecure Talking-Face Reference System. In 2nd Workshop on Multimodal User Authentication (MMUA'06), Toulouse, France, May 2006.
29. D. Petrovska-Delacrétaz, G. Chollet, and B. Dorizzi, editors. *Guide to Biometric Reference Systems and Performance Evaluation*. Springer-Verlag, London, to be published.

30. BioSecure Reference Systems. <http://share.int-evry.fr/svnview-eph/>.

31. BioSecure Mutimodal Evaluation Campaign 2007 (BMEC'2007). <http://biometrics.it-sudparis.eu/BMEC2007/>.