EvaBio Platform for the evaluation biometric system : Application to the optimization of the enrollment process for fingerprint device

B Vibert, Z Yao, Sylvain Vernois, Jean-Marie Le Bars, Christophe Charrier, Christophe Rosenberger

EvaBio Platform for the evaluation biometric system : Application to the optimization of the enrollment process for fingerprint device

B.Vibert, Z.Yao, S.Vernois, JM.Le Bars, C.Charrier, C.Rosenberger
UNICAEN, GREYC F-14032 Caen, France;
ENSICAEN, GREYC, F-14032 Caen, France;
CNRS, UMR 6072, F-14032 Caen, France
{*benoit.vibert, zhigang.yao, sylvain.vernois, christophe.rosenberger*}*@ensicaen.fr*
{*jean-marie.lebars, Christophe.Charrier*}*@unicaen.fr*

Keywords: evaluation of biometric systems, biometric platform, minutiae template selection, quality metric

Abstract: Nowadays when someone wants to make a payment with smartcard, he has to enter a pin code to be identified. Only biometric is able to authenticate a user; yet biometric information is sensitive. To ensure the security and privacy of biometric data, OCC (On-Card-Comparison) has been proposed. This approach consists in storing biometric data in a secure zone on a smartcard and computing the verification decision in a Secure Element (SE). The purpose of this paper is to propose an evaluation platform for testing performance and security on biometric OCC. Based on two examples, we illustrate the different uses of the platform. The first example uses the "Quality module" which allows to choose the enrollment template with quality image. The second one addresses the reduction of the binary template when the number of minutiae is higher than expected by the OCC.

1 INTRODUCTION

Nowaday biometric is often use in our daily life, (passport, border control, smartphone to indentify and authenticate an individual). This kind of applications required the use of large online biometric databases which may cause many security and privacy problems. In order to avoid these problems, storage of biometric data and OCC verification are increasingly made on a SE (Secure Element) such as the French passport chip. The main benefit of this solution is to avoid the transmission of the biometric reference of the user. The user has also the control of its own biometric data stored in the SE. A secure element guarantees many security issues of the biometric reference (confidentiality, integrity).

The SE is frequently used for several applications such as border control or face to face bank payment. Thus to avoid misused identity for example, it becomes very important to define a general methodology for evaluating these embedded systems. Our lab have a strong link with industrial and a lot of them wants to compare and evaluate OCC or sensor to choose the best. This is why we proposed, in this paper, our evaluation platform of biometric OCC for analyzing its performance and security.

The paper is organized as follows. Section 2 is devoted to the state-of-the-art of evaluation platform. Section 3 describes the proposed platform and new modules. In Section 4, we illustrate the benefit of the proposed platform through two examples of uses cases. We conclude and give some perspectives in Section 5.

2 STATE-OF-THE-ART

In the literature, only few platform exist. We can cite the NIST platform (Grother et al., 2011), which is used in their annual research competition. It allows researchers or manufacturers to test their OCC or minutiae extractors, in term of interoperability. The NIST report disseminates information on FMR (False Match Rate) and FNMR (False Non Match Rate) for each OCC and extractor. One drawback is the difficulty to integrate new modules for example the automatic generation of results in pdf file.

We can also mention the online FVC-Ongoing platform (Biolab, 2009) dedicated to algorithms for fingerprint verification (evolution of the FCV com-

petitions). The platform offers multiple databases grouped into two parts. The first one (Fingerprint Verification) quantifies both enrollment and verification modules, while the second one (ISO Fingerprint Matching) quantifies only the verification module on ISO Templates (ISO, a) based on minutiae. Performance metrics are: the failure to acquire rate (FTA) and the failure to enroll rate (FTE), the false non match rate (FNMR) for a defined false match rate (FMR) and vice versa, the average enrollment and verification times, the maximum size required to store the biometric template on the SE, the distribution of legitimate and impostors users scores and the ROC curve with the associated equal error rate (EER). The main drawback of this platform is that it is necessary to submit the executable or source code of the MOC to the online platform which can cause confidentiality issues.

Finally the last platform actually in development, BEAT (Biometric Evaluation And Testing) project (Project, 2013) is an european project. At the end of the project, a framework to evaluate the performance of biometric technologies using several metrics and criteria (performance, vulnerabilities, privacy). The goal of this project is to have a common platform for all the industrial and researcher to evaluate their product and to have an independent and certified result with common criteria. This platform is not released actually and it is the only drawback.

We have seen the platform present on the field and we have presented theirs possibility and drawbacks. However no platform answer to our criteriae (usability, modularity,...). If the BEAT platform was appeared a few years earlier, it would be a serious candidate. This is why we have decided to develop our platform we present here.

3 EVABIO PLATFORM

3.1 General scheme

The general synopsis of the proposed platform is given Figure 1. This evolution of our first platform (Vibert et al., 2013) allows us to have more functionality such as Sensor, Computing, Quality metrics, Security for MOC and Audit. The new modules yields to developer or researchers to have different kinds of methods to evaluate OCC or to choose a sensor. The proposed platform is composed of several new modules that will be defined in further subsections.

3.2 Modules

The platform is composed of different modules with specific treatments, and all modules are independent. This modularity allows us to modify a module without changing the overall operation of the platform. We may change only one and measure how this change impacts results. For example we can, quantify and determine if we improve or not the performance when we made a selection on enrollment template with image quality metrics. The platform uses active mechanisms of communication by event allowing multiple modules simultaneously access data exchanged between the client application and the OCC, thus offering "on the fly" analysis of results. All the main modules such as `Core`, `Scenario`, `Performance`, `GUI interface` are explain in (Vibert et al., 2013). In this paper, only `Sensor`, `Computing`, `Evaluation`, `Image quality assessment` modules will be described.

The `Sensor` module is a little platform which permits to acquire real and fakes fingerprint databases with real finger and specific protocols. This module is used to evaluate the performance of a sensor and to provide attacks on it. We also went on mortuary to test if sensor are able to acquire dead fingerprint (Vibert et al., 2014). This sensor platform could be used in input on `Core`, to acquire in live one or more fingerprints to compare it on OCC.

`Computing` module yields to have a distributed computation to improve the efficiency of the evaluation of OCC. For example, from three OCC on three smartcards, we are able to run three different test in parallel. We divide by three the evaluation time for a campaign.

`Image quality assessment` module is devoted to the quality metric of fingerprint images. Fingerprint quality metric is an auxiliary solution to guarantee the matching performance by dropping the bad quality samples (Grother and Tabassi, 2007) in both the enrollment and matching sessions. This purpose can be simply achieved because good quality prints could provide more precise and reliable features. Obviously, this is also beneficial to the OCC operations, especially when it is necessary to consider minutiae selection.

There will be a much higher probability that a minutiae extractor can correctly localize minutiae points within good quality images than that within bad quality prints (Chen et al., 2005). Therefore, a

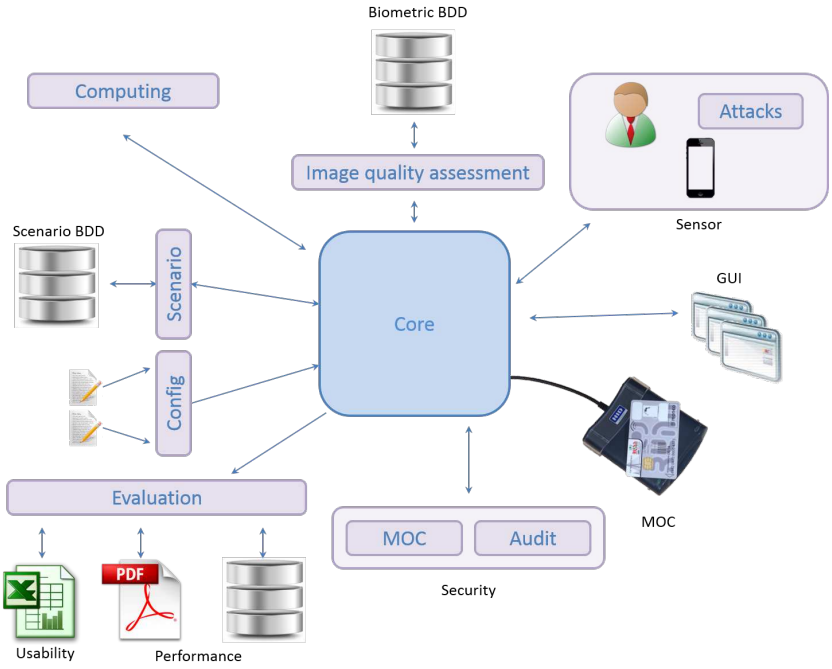


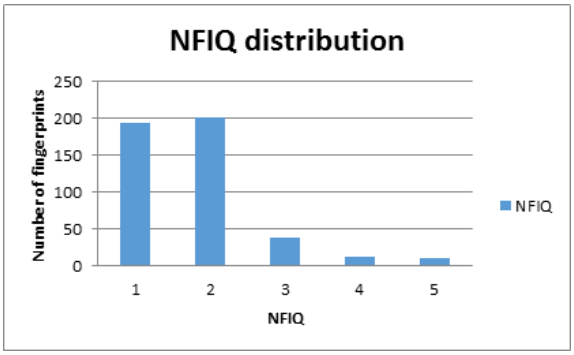
Figure 1: General scheme of EvaBio platform

reduced minutiae template can preserve correctly detected minutiae as much as possible rather than the spurious points, and the performance could be ensured as well. The quality metric module in the platform is combined with a validation component which allows the user to measure the performance of variant metrics, which enables making a further decision to choose an appropriate metric.

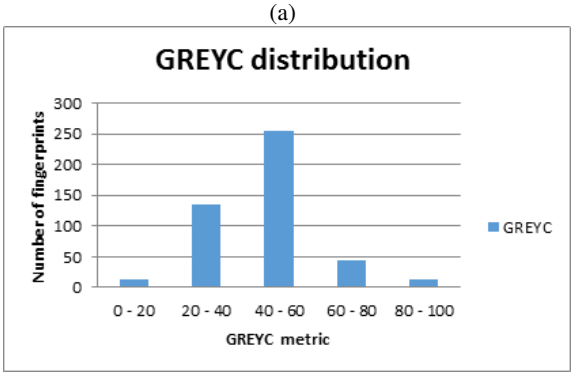
In the Figure 2, two examples of distribution on image quality metrics have been involved in the `Image quality assessment` module : 1) NFIQ (Tabassi and Wilson, 2005) and 2) GREYC Q metric. The NFIQ generates five quality levels from 1 to 5 (Figure 2(a)), where the best quality is indicated by the lowest value and the maximum level denotes the samples of very poor quality. GREYC Q metric (El-Abed et al., 2011), estimates the quality of fingerprint with five score groups (Figure 2(b)), poor (0-20), bad (20-40), medium (40-60), good (60-80) and very good (80-100). Such a continuous quality score could generate a better distribution of sample qualities than those using only few quality levels. This module is also a modular unit so that other quality metrics are also employable in the experiment.

`Evaluation` module used metrics commonly used in the literature and more specific ones:

- False Match Rate (FMR): it measures how many times the biometric data of a user provides positive verifications with biometric data of another user.
- False Non Match Rate (FNMR): it measures how many times the biometric data of a user gives a negative verification of biometric data with the same user,
- Success rate of attack: it measures the ratio of successful attacks (number of positive result over a number of transactions).
- Measuring interoperability: it quantifies the ratio of successful tests when providing an ISO template to the MOC.
- ROC curve: It describes the behavior of the biometric MOC for each value of the decision threshold (from which a test is positive). This implies that it is possible to obtain the comparison score from the OCC or to set decision threshold. For industrial OCCs, this is rarely the case but for research ones, this information is always available.
- Verification Time: we measure the time required to achieve a OCC enrollment or to obtain a verification result (after sending the ADPU (Application Data Protocol Unit defined in (ISO, b)) to the SE. It is also possible to generate several statistics on computation times such as histogram verification time, average, minimum or maximum time.
- Quality metric for images : The quality is important in images because a bad images produced



(a)



(b)

Figure 2: Image quality distribution for NFIQ and Q

bad minutiae, we used NFIQ metric witch is the standard and GREYC Q metric develop in our lab.

4 EXAMPLE OF USES CASES

In this section, we present experimental results on a commercial OCC with the selection of enrollment template when we have the quality of the original image.

4.1 Enrollment template selection

In this study, a method which permits to choose an enrollment template with the best image quality and the maximum number of minutiae accepted by the OCC has been proposed. This approach is tested with NFIQ and Q metrics, and we obtain a better result than before only with the selection of enrollment template. Figure 3, illustrates how we choose a template without quality selection 3(a) and when we use an image quality assessment process 3(b).

Concerning the protocol, the used biometric data have been collected in earlier experiment with 39 individuals. We have made three captures sessions with

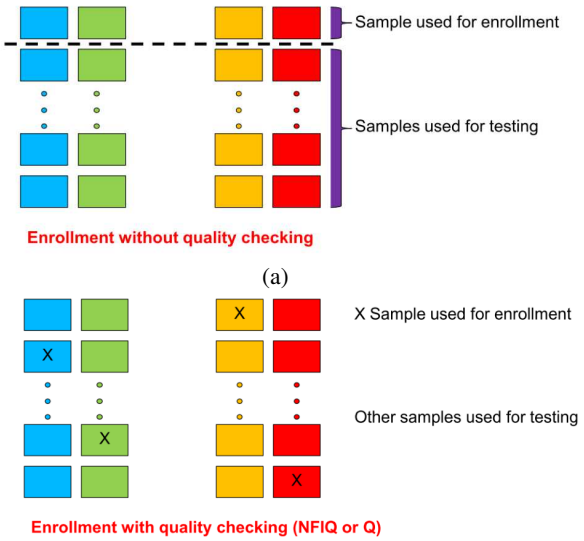


Figure 3: Illustration of protocol selection for enrollment template with and without quality metric where one column corresponds to an individual.

two fingers : left and right index finger, with five captures per session per individual. In total we captured 1170 fingerprint images and ISO Compact Card templates. On each images we compute the NFIQ and Q quality metrics and we save the results in a file.

To select the enrollment templates we have selected, for each person, the template with the best quality metric and the most number of minutiae under the maximum size allows by the OCC.

The performance without template selection has been done in (Vibert et al., 2013). In Table 1, we present the results of the comparison on commercial OCC with or without quality selection.

	FMR	FNMR
Without selection	0.41%	17.36%
NFIQ selection	0.05%	14.36%
Q selection	0.003%	4.75%

Table 1: Performance values for each metric selection method

As a conclusion of the performance, there is a fairly good robustness to imposture for all methods and the FNMR is relatively good with Q selection but not with others.

We observed quite satisfactory performance without quality selection, in terms of false match rate compared to those found in (Grother et al., 2011) and a very good performance with quality selection. The

false non match rate appears too high, however by improving the selection of the enrollment template, we are able to reduce the FNMR around 10% with NFIQ selection and an other 10% with Q selection in comparison with NFIQ. This experiment shows that the selection of enrollment template is very important to achieve good performance.

4.2 Minutiae selection

We also have develop a module to reduce the ISO Compact Card template when we have too many minutiae. The truncation method defined in the state-of-the-art (Grother and Salamon, 2007) has been initially embedded in the module. To determine if truncation is the best method in computation time and performance, two methods have been tested to be compared with it.

4.2.1 No selection

The first method keeps all minutiae in the template. The performance associated to the initial template is used as reference for the experimental results. We could expect that the other methods have a lower performance.

4.2.2 Selection by truncation

This method is based on a simple truncation *i.e.*, we only keep minutiae from the initial template the first N_{max} minutiae. The efficient of this simple approach depends on the method used to generate the fingerprint template. For many commercial biometric systems, a fingerprint template is generated with a specific method. It can be generated considering minutiae with the ascending locations Y as for example. In the case where multiple captures have been made, high quality minutiae (always present in the different captures as for example) can be placed at the beginning of the template. Selecting the N_{max} first minutiae could be in this case and very efficient an simple.

4.2.3 Barycentre selection

This method based on a pruning mechanism is simple and fast (few milliseconds). It has been proposed by the NIST for minutiae selection in (Grother and Salamon, 2007). It has been shown that minutiae located near the core of a fingerprint are the most useful ones for the matching process (Weiwe and Wang, 2002). Given a fingerprint template, the core location is usually unknown. However the centroid of minutiae can be a good estimate (when no other information is available). This minutiae selection approach tends to only keep minutiae near the centroid for this reason. We have four steps for the computation process:

1. Compute the centroid of the minutiae from the fingerprint template (containing N_j minutiae);

$$Centroid = (X_c, Y_c) = \frac{1}{N_j}(\sum_{i=1}^{N_j} X_i, \sum_{i=1}^{N_j} Y_i) \quad (1)$$

2. Compute the distance of each minutiae to the centroid;

$$r_i = \sqrt{(X_i - X_c)^2 + (Y_i - Y_c)^2},\ i = 1 : N_j \quad (2)$$

3. Sort in ascending order minutiae according to the distance $r_i,\ i = 1 : N_j$;
4. Select the first N_{max} minutiae.

4.2.4 Performance evaluation

Concerning the protocol, we have used the FVC2002DB2 (Maio et al., 2002) databases to illustrate our purpose. This database is composed of 8 fingerprints per person and 100 individuals, with a total of 800 fingerprint.

All minutiae templates used in the experiments have been extracted using the NBIS tool, MINDTCT (Watson et al., 2007) from the NIST. In order to realize the matching of fingerprint templates, we used a very well known minutiae matching algorithm proposed in 1997 by Jain et al (Jain et al., 1997). This method consists of an alignment stage (translation and rotation estimation between the two templates to compare) and a matching stage after transformation.

To evaluation the performance of minutiae selection algorithms, we use the AUC (Area Under the Curve) metric since it is often considered as global performance criterion. We use this value to quantify the efficiency of a minutiae selection method.

We compute the AUC value for each selection method with N_{max} varying from 30 to 50 by step of 2.

We show in Figure 4 the result of minutiae selection using the two tested methods. Selected minutiae are represented by a red star and others with a blue circle. We can see with the barycentre approach, select minutiae are near the estimated CORE. With truncation method, we loose the right part of the template.

Table 2 details the AUC value for each minutiae selection method for the FVC2002 DB2 database. On this database, most of selection methods permit to

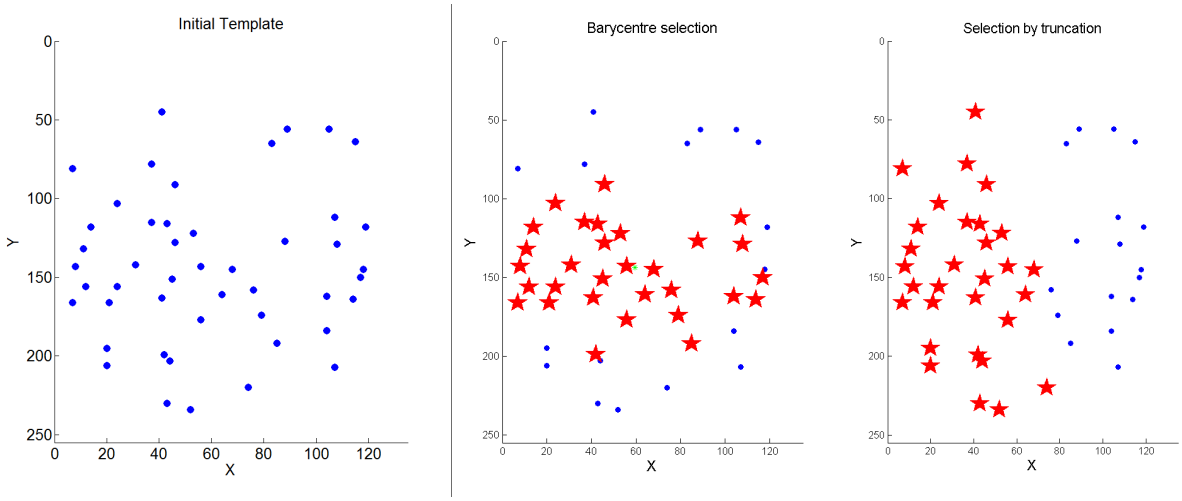


Figure 4: Example of minutiae selection on a fingerprint sample: stars represent selected minutiae by the different methods. For the barycenter selection approach, the green point represents the estimated CORE point (barycenter of minutiae)

N_{max}	30	34	38	42	46	50
No selection	11.2%	11.2%	11.2%	11.2%	11.2%	11.2%
Truncation	10.2%	9.97%	9.29%	**8.93%**	9.41%	9.48%
Barycentre	**8.73%**	**9.01%**	**9.00%**	9.26%	**9.17%**	**9.47%**

Table 2: AUC values for each minutiae selection method for different values of N_{max} on FVC2002DB2

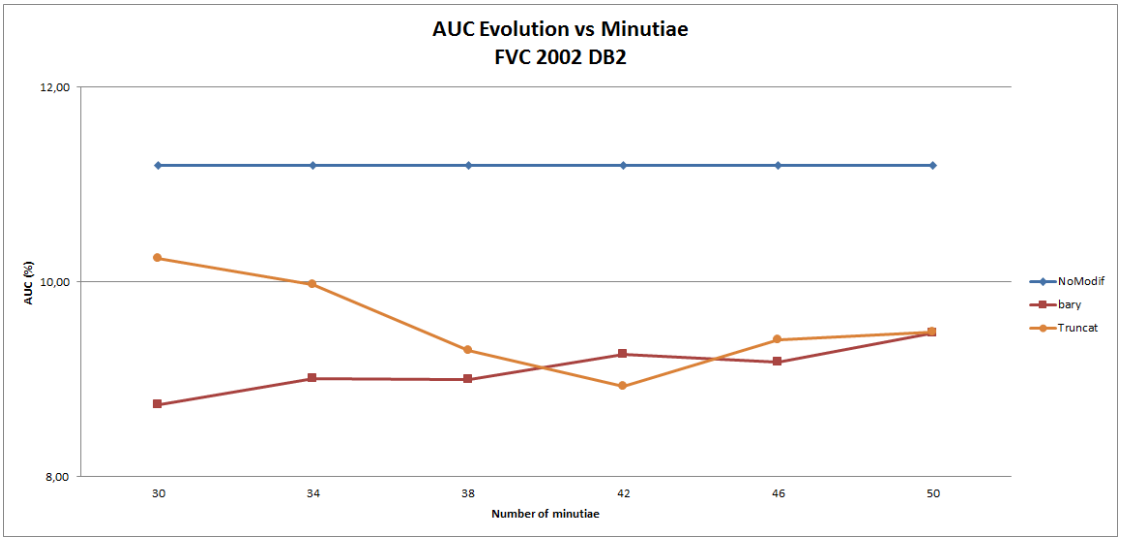


Figure 5: Evolution of the AUC value face to minutiae selection on FVC2002DB2

obtain a better performance (Cf. Figure 3) compared to the initial template (no selection).

To conclude with this illustration, we observe than barycenter is better than truncature method which is even the standard method. This use case illustrate the interest of the evabio platform, and a dedicated paper on this problematic will be publish soon.

5 CONCLUSIONS

In this paper, we have presented our biometric evaluation platform, and we have illustrate with two examples the capability of the platform. The first one quantifies how the use of image quality metrics on enrollment template selection is influences performance. The second is a small comparative study of fingerprint minutiae selection algorithms. To conclude, we demonstrate the facility to obtain result with our platform and the facility to use it.

In perspective, we plan to develop new modules to evaluate the OCC and sensor on smartphone and to design new sorts of attacks on OCC and sensors. We will also improve the scenario module to propose new tests.

REFERENCES

ISO/IEC 19795-2. information technology - biometric data interchange format - part 2 : Finger minutiae data, 2004.

ISO/IEC 7816-1 to 15: Identification cards - Integrated circuit(s) cards with contacts(Parts 1 to 15). ISO/IEC, http://www.iso.org.

Biolab (2009). FVConGoing. https://biolab.csr.unibo.it/FVCOnGoing.

Chen, Y., Dass, S. C., and Jain, A. K. (2005). Fingerprint quality indices for predicting authentication performance. In *Audio-and Video-Based Biometric Person Authentication*, pages 160–170. Springer.

El-Abed, M., Hemery, B., Charrier, C., Rosenberger, C., et al. (2011). Evaluation de la qualité de données biométriques. *Revue des Nouvelles Technologies de l'information (RNTI)*, pages 1–22.

Grother, P. and Salamon, W. (2007). Interoperability of the ISO/IEC 19794-2 compact card and 10 ISO/IEC 7816-11 match-on-card specifications 11.

Grother, P., Salamon, W., Watson, C., Indovina, M., and Flanagan, P. (2011). Minex ii "performance of fingerprint match-on-card algorithms" phase iv : report NIST interagency report 7477 (revision ii).

Grother, P. and Tabassi, E. (2007). Performance of biometric quality measures. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):531–543.

Jain, A. K., Hong, L., Pankanti, S., and Bolle, R. (1997). An identity-authentication system using fingerprints. *Proceedings of the IEEE*, 9:1365–1388.

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., and Jain, A. K. (2002). FVC2002: Second fingerprint verification competition. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 3, pages 811–814. IEEE.

Project, B. (2013). Beat project. https://www.beat-eu.org/.

Tabassi, E. and Wilson, C. L. (2005). A novel approach to fingerprint image quality. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II–37. IEEE.

Vibert, B., Lebouteiller, J., Keita, F., Rosenberger, C., et al. (2014). Biometric sensor and match-on-card evaluation platform. In *International Biometric Performance Testing Conference (IBPC)*.

Vibert, B., Rosenberger, C., and Ninassi, A. (2013). Security and performance evaluation platform of biometric match on card. In *Computer and Information Technology (WCCIT), 2013 World Congress on*, pages 1–6. IEEE.

Watson, C. I., Garris, M. D., Tabassi, E., Wilson, C. L., Mccabe, R. M., Janet, S., and Ko, K. (2007). Users guide to nist biometric image software (nbis). Technical report, NIST.

Weiwe, Z. and Wang, Y. (2002). Core-based structure matching algorithm of fingerprint verification. *International Conference on Pattern Recognition*.