Computation EvaBio: A Tool for Performance Evaluation in Biometrics

Julien Mahier, Baptiste Hemery, Mohamad El-Abed, Mohamed El-Allam, Mohamed Bouhaddaoui, Christophe Rosenberger

Computation EvaBio: A Tool for Performance Evaluation in Biometrics

Julien Mahier, Baptiste Hemery, Mohamad El-Abed*, Mohamed T. El-Allam, Mohamed Y. Bouhaddaoui, Christophe Rosenberger

GREYC Laboratory
ENSICAEN - University of Caen Basse Normandie - CNRS
6 Boulevard Maréchal Juin, 14000 Caen Cedex - France

Abstract

Performance evaluation is a key factor in biometrics. In order to assess the quality of a new biometric system, one has in general to compute the performance on a large dataset to have significant results. It is not always an easy task as the computation complexity could be important. Moreover, the use of existing datasets is not obvious as it is represented in different ways (datasets, directory of images) and contain multiple scenarios (single enrollment process, different numbers of samples for the template generation for each user, *etc.*). We propose in this paper a new distributed-based platform facilitating the computation on different datasets in biometrics. The architecture of the proposed platform is composed of a server distributing computation tasks on its available clients. Experimental results on two biometric datasets (PolyU finger knuckle print and AR face datasets) show the benefit of proposed computing platform.

Keywords:

Biometrics, Performance Evaluation, Benchmark, Distributed Computing

*Corresponding author

Email addresses: julien.mahier@ensicaen.fr (Julien Mahier), baptiste.hemery@ensicaen.fr (Baptiste Hemery), mohamad.elabed@ensicaen.fr (Mohamad El-Abed), christophe.rosenberger@ensicaen.fr (Christophe Rosenberger)

1. Introduction

The interest for biometric technologies has increased during the last few years. The application of biometrics in industrial domains becomes a reality. Apart from the traditional modalities such as fingerprints [1] or iris recognition [2], new modalities are studied such as keystroke dynamics [3], gait [4] or finger knuckle prints [5].

As there are many research works in biometrics, there is a strong need for comparing and benchmarking biometric algorithms. Several datasets are made available in order to ease this comparison process. However, a dataset can be used within several protocols, which sets up a benchmark, to establish the performance of an algorithm. This leads to an important complexity to correctly evaluate a biometric system. Moreover, there are some well-known performance metrics presented by the International Organization for Standardization ISO/IEC 19795-1 [6] for the evaluation process such as the ROC curve (Receiver Operating Characteristic) and the EER value (Equal Error Rate). The EPC (Expected Performance Curve) is also another performance metric presented in [7] towards the performance evaluation of biometric algorithms. It requires some datasets to be split in two parts, one dealing with development and the other one with experiment. However, it is still difficult to compare biometric algorithms. It is required that the used dataset is large enough to be significant, but this leads to a very long time consuming experiment. If we have a benchmark datasets of N individuals and M samples for each user, the definition of the FRR (False Rejection Rate) requires $(M-1)\cdot N$ computations and for the FAR (False Acceptance Rate) $N(N-1)\cdot M$ ones (for a single enrollment biometric system). In order to achieve a more accurate evaluation, we need to have a large dataset which exponentially grows the computation complexity. Some platforms dedicated to biometric algorithm evaluation exist (such as the online evaluation platform FVC-onGoing), but are generally specialized for only one kind of biometric modality and only one dataset.

We present in this paper a distributed-based computation platform called EvaBio dedicated to the evaluation of biometric systems. The proposed platform is modality-independent. Thus, it needs to be able to manage any biometric dataset. Moreover, this platform needs to be easy to use and efficient. Thus, the required computation tasks can be easily distributed on several computers.

The outline of this paper is as follows: after the introduction, we present in section 2 the existing platforms dedicated to the evaluation of biometric algorithms. We also present the EvaBio platform we are working on in section 3 with a description of its computation capabilities and the scenarios generation for performance evaluation. Section 4 presents some experimental results. Finally, we conclude and give the perspectives of this work in section 5.

2. State of the art

In comparison to traditional authentication systems (such as password-based methods), biometric systems are less accurate and do not provide a 100% reliable answer. Due to this inaccuracy, many efforts are done in the literature for the performance evaluation of such systems. We provide first in this section, a brief description of the used biometric performance metrics and some well known benchmarks, followed by an overview of the well known platforms which aim to evaluate monomodal and multimodal biometric systems.

2.1. Biometric Performance Metrics

The comparison of two biometric samples produces a similarity score. If the score is higher than a predefined decision threshold, then the system accept the claim user, otherwise the claim is rejected. Figure 1 illustrates the distribution of the genuine and impostor scores. This figure shows that depending from fixed decision threshold, we obtain different kind of errors: 1) genuine users who are incorrectly rejected by the system, and 2) impostors who are considered as genuine users.

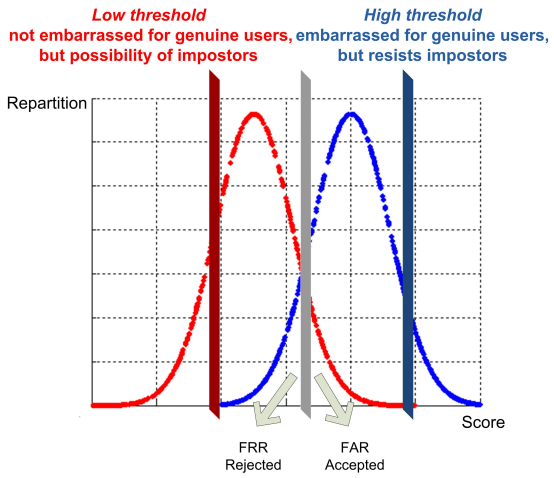


Figure 1: Distribution of genuine and impostor scores

In order to evaluate and compare the performance of biometric systems, the International Organization for Standardization ISO/IEC 19795-1 [6] defines several performance metrics such as:

- Failure-to-enroll rate (FTE): proportion of the user population for whom the biometric system fails to capture or extract usable information from biometric sample;
- Failure-to-acquire rate (FTA): proportion of verification or identification attempts for which a biometric system is unable to capture a sample or locate an image or signal of sufficient quality;
- False acceptation rate (FAR): proportion of impostors that are accepted by the biometric system;
- False rejection rate (FRR): proportion of authentic users that are incorrectly denied;
- False-match-rate (FMR): the rate for incorrect positive matches by the matching algorithm for single template comparison attempts. FMR equals FAR when the biometric system uses one attempt by a user to match its

own stored template;

- False-non-match rate (FNMR): the rate for incorrect negative matches by the matching algorithm for single template comparison attempts. FNMR equals FRR when the biometric system uses one attempt by a user to match its own stored template;

- Equal error rate (EER): it is the value where both errors rates, FAR and FRR, are equals (*i.e.*, FAR = FRR). It constitutes a good indicator, and the most used, to evaluate and compare biometric systems. In other words, lower the EER, higher the accuracy of the system.

These performance metrics may be drawn to graphically visualize the precision of the target system. A well known curve is the ROC (Receiver Operating Characteristic) curve which plots both types of errors (FAR and FRR) as depicted in figure 2. This curve provides the performance of the tested system among several decision thresholds. In addition to these metrics, usability metrics are computed within competitions and platforms such as the average enroll time, average match time, average template size. These additional metrics are important in order to have useful and usable systems.

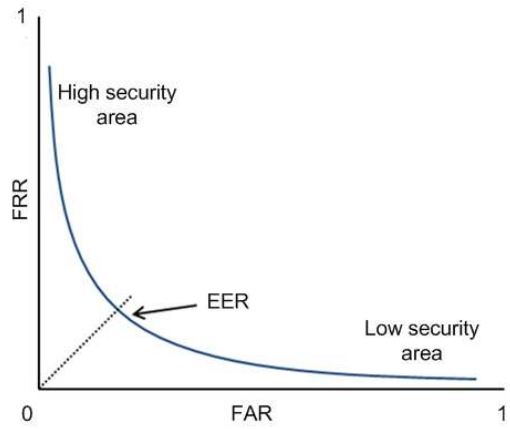


Figure 2: ROC curve

2.2. Benchmarks

In order to compute the performance metrics provided in the previous section, we need biometrics benchmarks. Several benchmarks are publicly available to researchers in order to evaluate their developed algorithms. Two categories of datasets exist in the literature: 1) monomodal datasets such as PolyU FKP[8], GREYC-Kesytroke [9], FACES94 [10], AR [11], FERET [12, 13], FRGC (Face Recognition Grand Challenge) [14], *etc.*, and 2) multimodal datasets such as XM2VTSDB [15], BANCA [16], BIOSECURE [17], *etc.* Figure 3 illustrates an example of samples from AR, ENSIB and FACES94 datasets. A detailed description of the used datasets (PolyU FKP and AR datasets) are given in section 4.1.

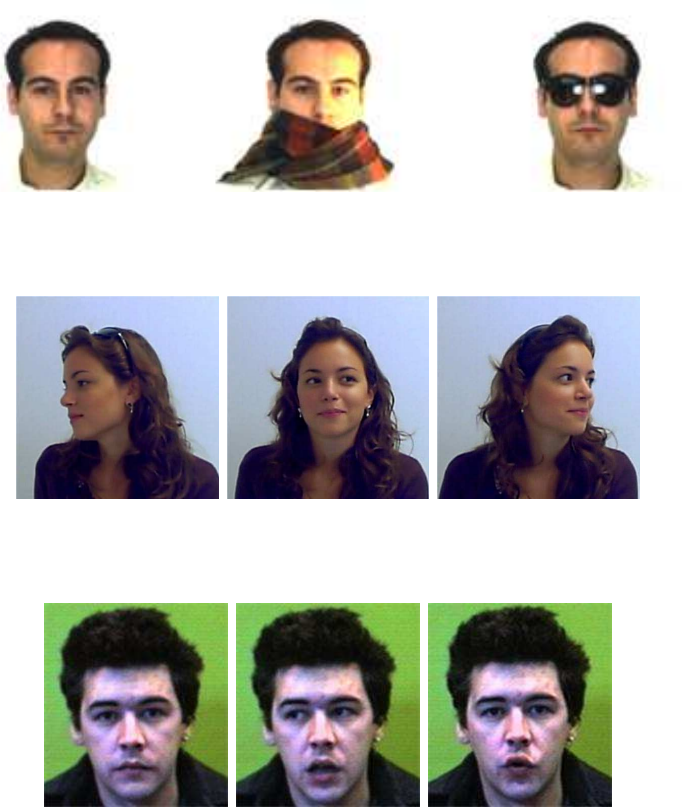


Figure 3: Biometric samples from the AR, ENSIB and FACES94 datasets

2.3. Biometric Platforms

2.3.1. Fingerprint Verification Competition-onGoing (FVC-onGoing)

FVC-onGoing is the evolution of the FVC international competitions held on 2000, 2002, 2004 and 2006. It is a web-based automated evaluation for fingerprint recognition algorithms available at `https://biolab.csr.unibo.it/FVCOnGoing/`. The tests are done using sequestered datasets. The platform uses several performance measures in terms of accuracy indications (such as the EER value), enrollment and matching failures (such as the FTE value), usability measures (such as the average enrollment time), and memory indicators (such as the average template size).

The FVC-onGoing platform is a useful solution since the comparison process is done using the same protocol, and the used evaluation metrics cover the main criteria of having useful and usable systems. In addition, such kind of platforms is important since the evaluation process is done by the online platform side. This facilitates the tasks of other researchers that may do not have the required materials (*e.g.*, a cluster computing) towards the evaluation of their developed algorithms. However, FVC-onGoing is dedicated to the performance evaluation of fingerprint algorithms. Hence, it could not be exploited for other types of well used modalities such as face modality [18].

2.3.2. BioSecure Reference and Evaluation framework

BioSecure was a project of the 6th Framework Program of the European Community. Its main objective was to build and provide a common evaluation framework, which investigates and compares the biometrics-based identity authentication methods. It provides twelve benchmarking reference systems freely available at `http://share.int-evry.fr/svnview-eph/`: 2D face, 3D face, fingerprint, hand, iris, signature, speech and talking-face reference systems.

These reference systems are made of replaceable modules (preprocessing, feature extraction, model building and matching) which allow developers and re-

searchers to investigate the improvement of a specific part of the tested system. In this case, a researcher can evaluate and compare its matching algorithm just by replacing the corresponding module in the reference system. An example of the used performance measures is the distribution of genuine and impostor comparison scores. BioSecure reference systems are useful since they cover several kinds and the most used modalities. In addition, the structure of these reference systems in allowing a researcher to test a specific part of a developed algorithm is a main advantage of these systems. However, in comparison to the FVC-onGoing online platform, the computation is done on the side of the developers which may do not have the specific materials dedicated to such kinds of evaluation. Moreover, these reference systems are dependent from the used modality and use a predefined dataset and protocol. This fact limits their use if the researchers want to evaluate their developed algorithms using other protocols or even other datasets.

2.3.3. Biometric experimentation environment (BEE)

The BEE distribution [14] is the test environment of the evaluation of face recognitions systems in the Face Recognition Vendor Tests (FRVT2006). It is a computational experimental environment which easily allows researchers to compare their developed facial-based algorithms. The BEE distribution contains all the data sets for performing and scoring the experiments. It also provide a PCA-based method (known as BEE system) that has been optimized for large-scale problems as a baseline algorithm, which applies the whitened cosine similarity measure. An example of its use to compare face recognition systems in given in [19]. The BEE framework returns three kinds of ROC curves as a performance measure, corresponding to images collected within semesters, within a year, and between semesters, respectively. The BEE system is a useful framework to compare face recognition systems. However, BEE do not resolve the computation complexity of performance metrics, which is considered as a main drawback in biometrics research field. Towards resolving this problematic, we present in the next section a distributed computation platform of performance metrics.

3. Computation EvaBio platform

3.1. Proposed platform

The Computation EvaBio Platform started its development in February 2009. Its goal is to provide a low cost, scalable, adaptable, modalities independent and development tools independent platform dedicated to biometric computation. The main idea is to use the unexploited calculation power of existing computer pool in a research laboratory.

The interest for the researcher is to focus only on his/her algorithm assuming small development constraints. The main constraint is to develop a full autonomous application or script (without any interaction or IHM) dealing with an atomic computation task. The global vision of the different calculation steps is defined in the *Computation Task List*, within an XML file. Each task defined in this file describes a computation task to be done on the laboratory computers. The server is in charge of the repartition of the different computation tasks. Figure 4 shows the present architecture (v0.5). It is composed of four parts, the server and the distribution process (presented by Mahier et al. in [20]), the administration client and the computation task list generation, the atomic algorithm creation and deployment, the dataset access and the different biometrics dataset abstraction.

3.2. Biometric datasets modeling

We present in this section an original meta model suitable for any biometric dataset, we choose to perform a modeling of four publicly available datasets. It will allow us to be able to launch a performance evaluation scenario described by this meta model. We choose to use the Merise method [21] for modeling four existing biometric datasets. It is a modeling methodology for the development of information systems, which is well suited for the dataset modeling. By analysing existing benchmarks in biometrics (AR, Poly U FKP, FERET and GREYC-Kesytroke, *etc.*), we can extract some tables that are general for all

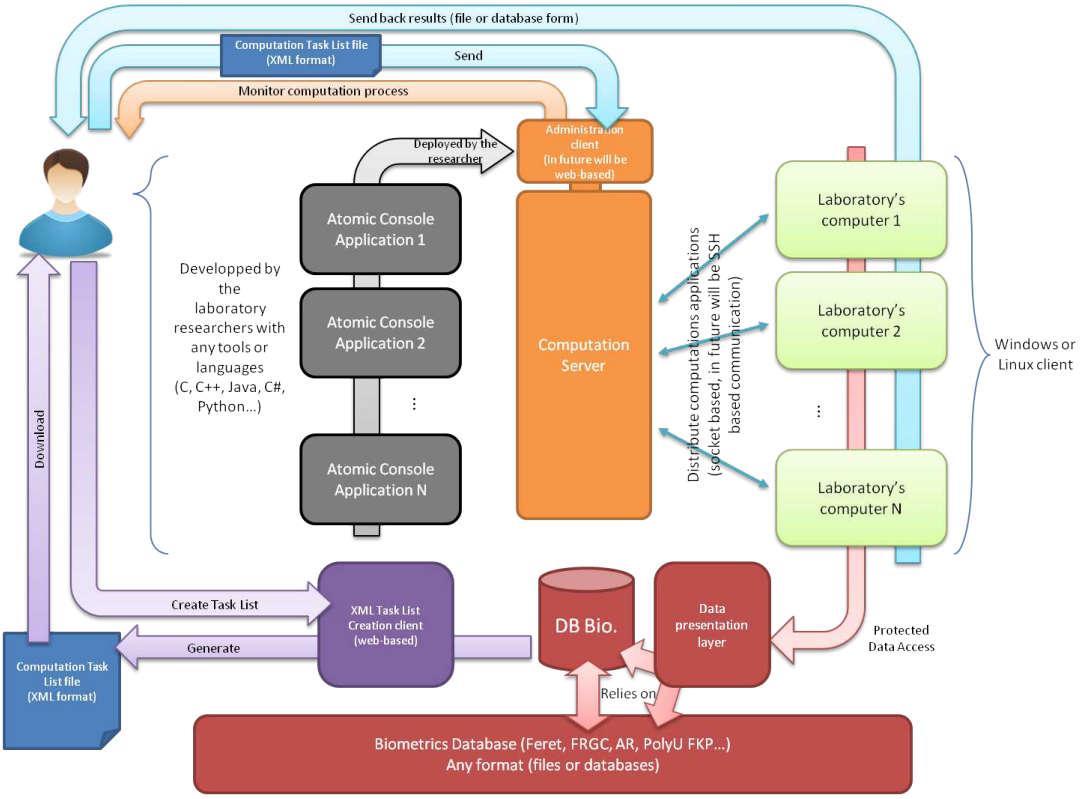


Figure 4: Computing EvaBio Platform Architecture

biometric datasets. We also had to add some tables specific in order to be able to manage several datasets. The obtained modeling can be seen at figure 5. The two top tables *Dataset* and *Modality* corresponds to specific tables we added to manage different biometric dataset. The modalityDesc field of the *Modality* corresponds to the used modality such as "face", "keystroke" and so on. It is linked to the *Dataset* table, where we store the name of the dataset and a short description. Next, we find the *Individual* table which contains data about individuals present in a dataset. It is limited to an individualNumber, which reflects the number of the individual in the dataset, and its gender. We add an externalID as a foreign key to a table with more details about this individual, such as first name, last name or date of birth. This table is not included in the

model for anonymity reasons, but the externalID can still be used to retrieve the same individual in several dataset, which can be useful for multi-modality biometrics. Finally, an active field can be used to deactivate an individual if needed.

An individual is linked to several sessions, stored in the *Session* table, via the *perform* table. The sessionID field is thus unique for each individual/session number, and can be associated to a date. These tables are associated to the table *BiometricData*, which contains information relative to the biometric sample. The data field is a link to the biometric sample and metaData can be used to store an additional information about biometric data, such as the used password for a keystroke dynamics biometric sample. The success field can be necessary to compute failure to enroll rate. The addInformation field is a link to additional data provided with some datasets, such as localization of eyes for face ones. A biometric sample is also related to the *Sensor* table, so we can manage several capture devices such as different keyboards or cameras.

Finally, we choose to add some specific information on each dataset in the *Type* table. This table contains information about the type of data contained in the dataset, such as the resolution of images, in an XML format. Thus, we add the *XSL* table, wich will contain the information about the XML format of the *Type* table. In case we want to add a new biometric modality in this model, we only have to add a new XSL entry and keep this model intact.

3.3. Use case

A researcher, named Paul, focus his work on a new face biometric algorithm. He first has to implement his algorithm for one data sample from one person (independently of the dataset and less coupled with the data format). To help him to code accurately, the platform provides design application templates on different languages. The different parameters of the algorithm will be passed as application arguments. When Paul succeeds in developing his basic console application, he will have to define the computation procedure, meaning to define

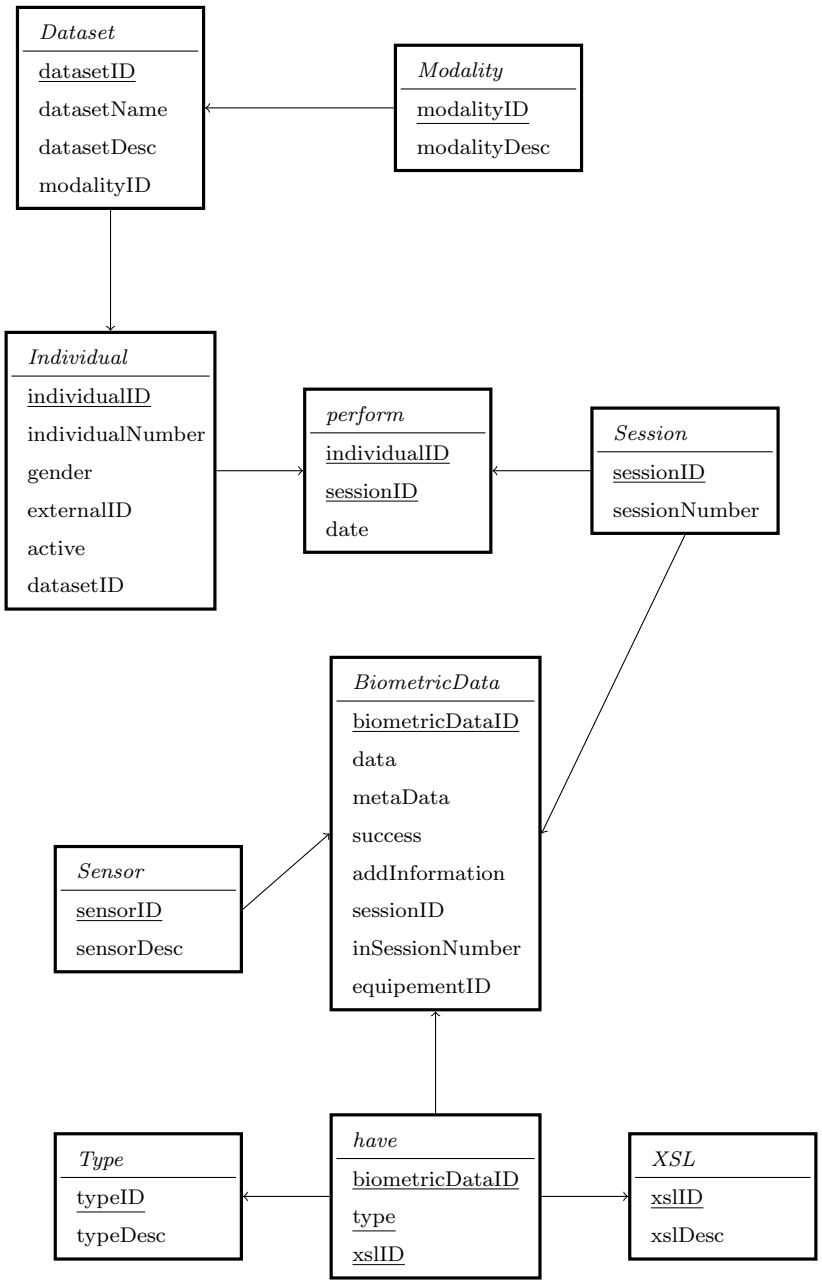


Figure 5: Meta-Model of a biometric dataset

the XML Computation Task List as depicted in table 1. This task list describes all the parameters of Paul's application will need to compute an atomic dataset. Finally, through the administration interface (see figure 6), Paul can launch its algorithm on all the predefined laboratory computers. The results are retrieved on a database or within a shared file on the network. Two weeks later, Paul need to replay some computations on another Biometric dataset. He just has to create the appropriate XML file and to launch the computation process.

```
<TaskPools>
 <TaskPool>
  <Task>
   <ExecutionCommand> ... </ ExecutionCommand>
   <ParameterList>
    <Parameter> ... </ Parameter>
   </ ParameterList>
  </ Task>
 </ TaskPool>
</ TaskPools>
```

Table 1: An XML Computation Task List description

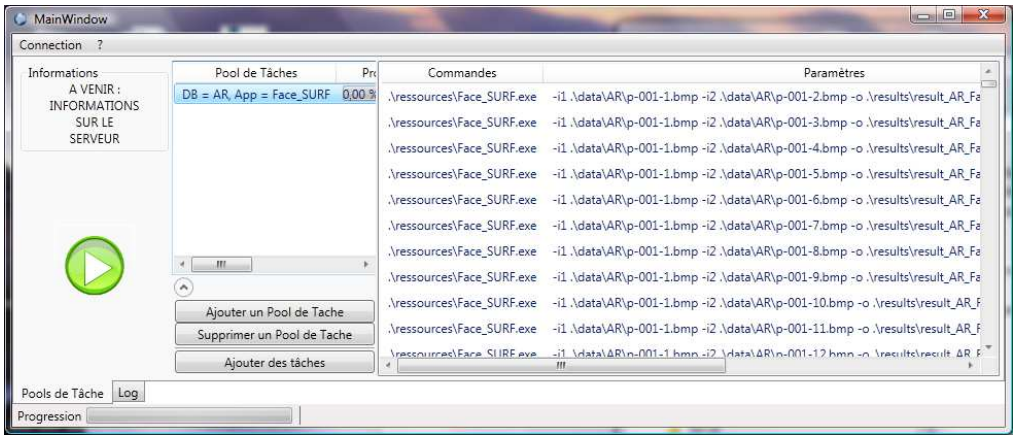


Figure 6: Administration interface

3.4. Discussion

The main interest of this solution is based on the biometric data definition format (the meta-model). Given the abstraction of the biometric data, the researcher focuses only on his algorithm and asks the platform to compute the preselected computation tasks. The abstraction enables a pre-selection from one part of a single dataset (for example only women and illumination altered samples of the AR dataset) to multiple parts of multiple datasets (for example only men of the AR and Feret datasets). The key point is on the tool which allows the researcher to question the meta-biometric-dataset and generates the XML task list. This tool will be presented in another paper.

4. Results

We present in this section some results of the computing EvaBio platform. First, we present how an existing biometric dataset can be fitted into the proposed model. Then, we present some results from the existing distributed computation tool.

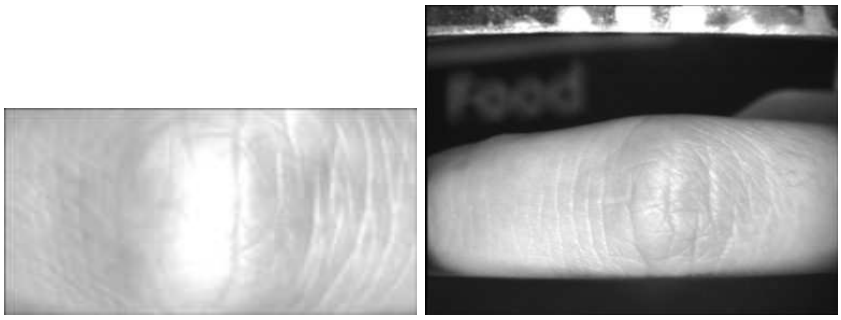
4.1. Fitting existing datasets

4.1.1. PolyU FKP dataset

The PolyU FKP images were collected from 165 volunteers (125 men and 40 women). Among them, 143 subjects were 20 to 30 years old and the others were 30 to 50. The images have been taken in two different sessions. In each session, the subject has provided 6 images for the right index finger, for the right middle finger, the left index finger and the left middle finger. There are therefore 48 pictures, from the 4 fingers from each person. In total, the dataset contains 7,920 images from 660 fingers. The interval average time between the first and second session is 25 days. The maximum interval and minimum were 96 days and 14 days respectively. The dataset consists of several files. Each file is written in a folder named as "nnn_fingertype": nnn represents the identity of the person and fingertype can be left index, right index, left middle or right middle.

Each of these folders contains 12 pictures, from 01 to 06 belong to session 1 and 07 to 12 belong to the session 2. There is actually two datasets. The dataset "FKP Database.zip" contains the original images and the "FKP ROI.zip" one contains region of interest pictures [22]. However, the construction of the two datasets is the same, and only the filename name changed, "ROI" being added at the end of the filename for the ROI dataset. An example of these datasets can be found in figure 7.

We present in figure 8 how the PolyU FKP dataset fits the meta-model. We can see that the Poly U FKP fits well with this meta model. Some fields cannot be completed such as the date where an individual performs a session. Moreover, some fields have default value for publicly available datasets such as the activate field of the *Individual* table, which is set to true. Apart from theses fields, we can see that other fields are filled with data related to the biometric sample. Some specific data related to a dataset can also be stored in the database. For example, the height and with of the images from the Poly U FKP dataset can be stored in the XML field of the *Type* table.



(a) 001_left index/01ROI (b) 002_left index/12

Figure 7: Two biometric samples from the PolyU FKP datasets

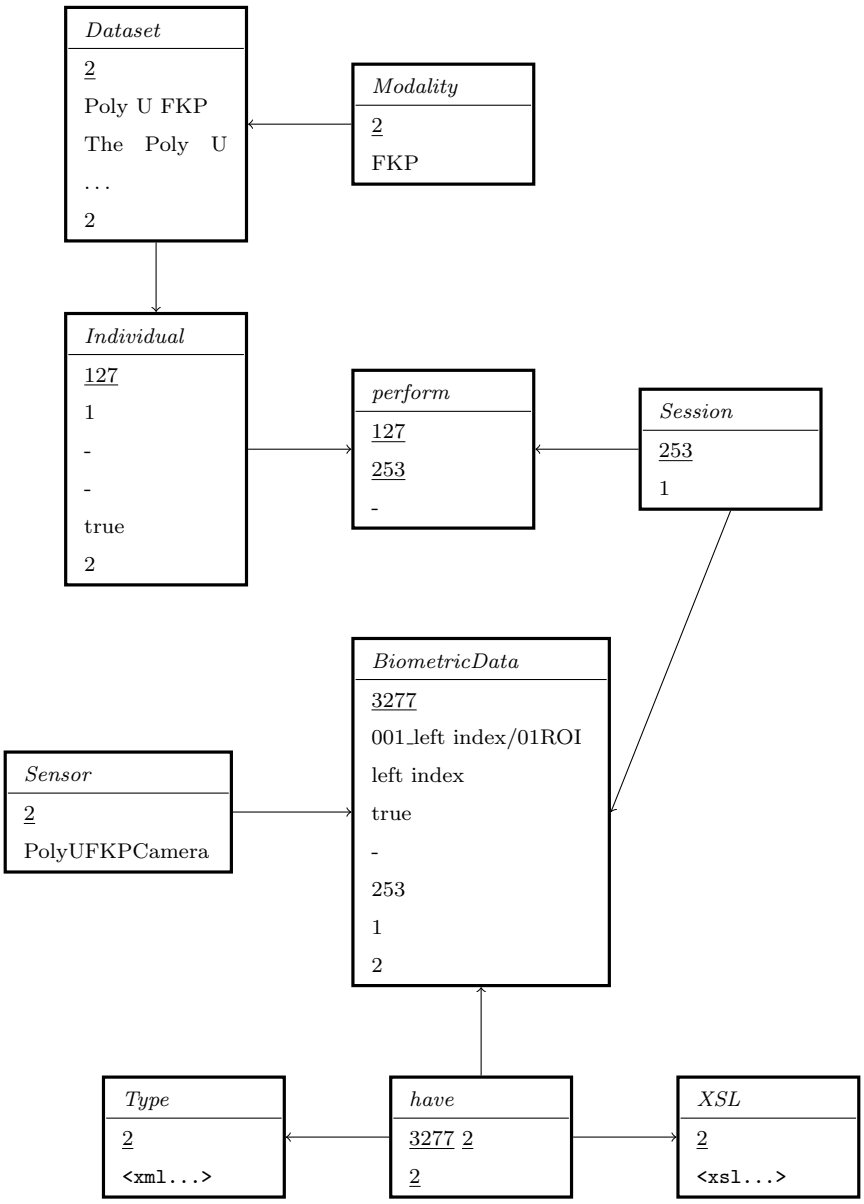


Figure 8: Application of the meta model to the FKP dataset

4.1.2. AR face dataset

The AR dataset contains over 4,000 color images, acquired with the same system, corresponding to 126 different faces (70 men and 56 women). The

images were taken from the front with various facial expressions, lighting conditions varied, and sometimes with sunglasses or a scarf. Also, no restrictions on clothing (clothes, glasses, *etc.*), makeup, and hair cut was imposed on participants. Each person has participated in two sessions, separated by 14 days. The same images were taken in both sessions. All image are in the same folder, and only the file name enable distinguishing biometric data. The name of the images is built this way "x-nnn-mm.jpg": x can take the value m or f (male or female), nnn identifies the person (can range from 001 to 070 for male and 001 to 056 for female) and mm identifies the session and the expression (01 to 26). Examples of such images are presented in figure 9.

(a) m-016-01 (b) m-016-17

Figure 9: Two biometric samples from the AR face dataset

Similarly to the Poly U FKP dataset, we present in figure 10 how the meta-model can be used to represent biometric data from the AR face dataset. We can see that the AR face dataset fits well the meta-model. We still have some fields that cannot be completed, but most of them are. We can also see that it is simple to create requests on this database. For example, if we want only male individuals from the AR dataset, we can create a SQL request containing two conditions such as `Individual.gender=male` and `Dataset.datasetName=AR`. Moreover, as the same model is applied, we can use the same SQL requests for both datasets. This clearly shows the interest of this meta model.

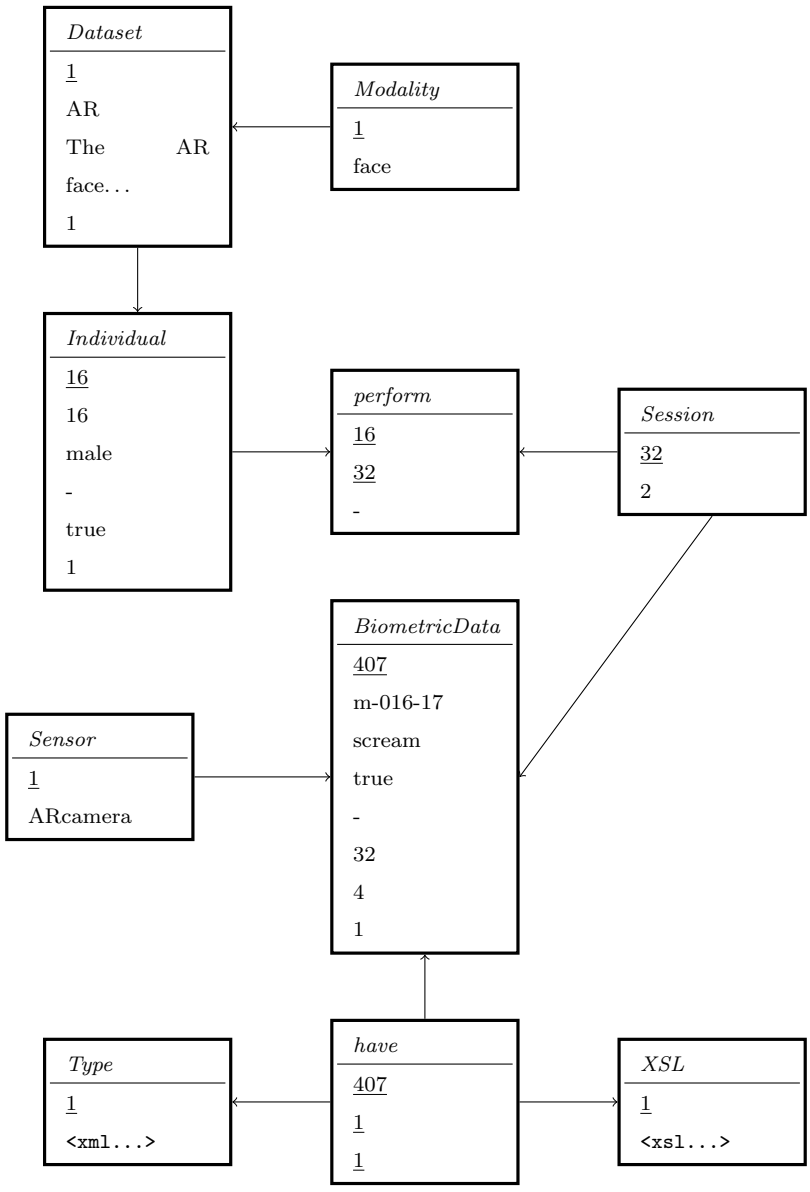


Figure 10: Application of the meta model to the AR dataset

4.2. Distributed computing

In order to show the efficiency of the distributed computing part of the EvaBio platform, we present the time gain provided by this platform using a

face verification algorithm based on SURF descriptor [23]. The used verification algorithm is a single-based enrollment. The comparison of two face images is considered as the comparison of two sets of keypoints detected from each image. Each SURF keypoint is characterized by a vector in 64 dimensions. The matching similarity principle is presented in previous works [24]. The matching between two images im_1 and im_2 corresponds to compute a similarity between two sets of features $X(im_1)$ and $X(im_2)$. We thus use the following matching method which is a modified version of a decision criterion first proposed by Lowe [25]. Given two keypoints $x \in X(im_1)$ and $y \in X(im_2)$, we say that x is associated to y iff:

$$d(x,y) = min_{\{z \in X(im_2)\}} d(x,z) \text{ and } d(x,y) \leq C\ d(x,y\prime) \tag{1}$$

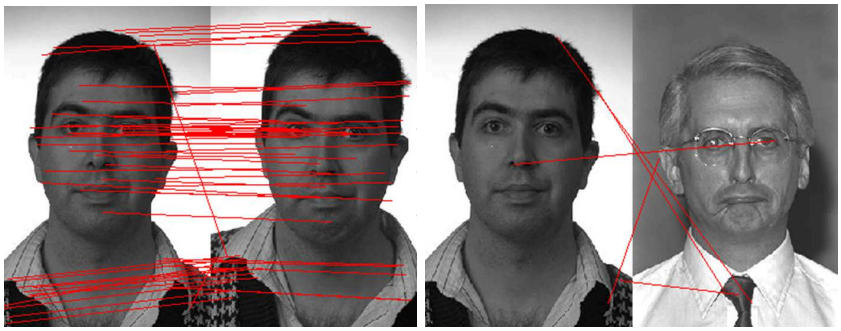
where C is an arbitrary threshold, $d(\cdot,\cdot)$ denotes the Euclidean distance between the SIFT descriptors and $y\prime$ denotes any point of $X(im_2)$ whose distance to x is minimal but greater than $d(x,y)$:

$$d(x,y\prime) = min_{\{z \in X(im_2),\ d(x,z)>d(x,y)\}} d(x,z) \tag{2}$$

In other words, x is associated to y if y is the closest point from x in $X(im_2)$ according to the Euclidean distance between SIFT descriptors and if the second smallest value of this distance $d(x,y\prime)$ is significantly greater than $d(x,y)$. The significance of the necessary gap between $d(x,y)$ and $d(x,y\prime)$ is encoded by the constant C. Then, we consider that keypoint x is matched to y iff x is associated to y and y is associated to x. An example of a genuine and an impostor matching result is given in 11(a) and 11(b), respectively. The number of associations is used here as a similarity measure.

This experiment represents the computation required for 30 individuals from the AR face dataset, and thus, 900 task were present in an XML file. For a 26 samples per individual, the definition of the FRR requires 750 comparisons and for the FAR 22620 ones. To do so, we used different numbers of clients to realize this computation tasks (some are present on the same computer). Results are

presented in table 2. For the two lines with local, the server and client were on the same computer, and thus does not require any network communication. For lines with a $n + n$, it means that two clients were used on n computer in order to benefit from the dual processor of computers. We can see that using two clients (line 1+1) on the same computer increases the performance. This is due to the fact that the computer used is a dual core processor and each task is single threaded. We can see that using as few as 3 computers is enough to clearly reduce the computation time.



(a) Genuine User (b) Impostor User

Figure 11: Face verification attempts using SURF descriptor: red lines correspond to a match between two associated keypoints

Number of clients	Time
1 (Local)	57m39s
1+1 (Local)	30m01s
1	70m27s
2	51m07s
3	32m19s
4	25m48s
4+4	14m07s

Table 2: Computation time results for different number of clients.

5. Conclusion and perspectives

We proposed in this paper a new platform dedicated to researchers and engineers in the field of biometrics to develop new algorithms. The computing EvaBio platform permits to launch on different clients some computation tasks to quantify the performance of an algorithm in biometrics. A meta model has also been proposed to characterize a dataset and to facilitate the creation of benchmark and complex scenarios. This meta-model can be used to store the data related to users and biometric sample acquisitions for any biometric modality. It aims to ease the relation between biometric samples and data related to these samples. Thus, a single database, using this meta model, allow researchers to have a single access point for each biometric benchmark stored in it.

We plan to use this model in a platform dedicated to performance evaluation. This tool, containing a distributed computation tool, enables an easy performance evaluation system. The perspective of this work is to create a tool that will enable to easily create the XML file containing the difference computation tasks, used by the distributed computation tool. This tasks list creation tool will rely on this meta-model to have a unique access point to all kind of biometric datasets. We plan to provide this tool for the research community in biometrics.

Acknowledgements

The authors would like to thank the French Research Ministry for their financial support of this work.

References

[1] A. K. Jain, L. Hong, S. Pankanti, R. Bolle, An identity-authentication system using fingerprints, Proceedings of the IEEE 85 (9) (1997) 1365–1388.

[2] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, S. McBride, A system for automated iris recognition, in: IEEE Workshop on Applications of Computer Vision (WACV'94), 1994, pp. 121–128.

[3] R. Giot, M. El-Abed, C. Rosenberger, Keystroke dynamics with low constraints svm based passphrase enrollment, in: IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS'09), 2009, pp. 1–6.

[4] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, K. W. Bowyer, The humanID gait challenge problem: data sets, performance, and analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2005) 162–177.

[5] A. Kumar, C. Ravikanth, Personal authentication using finger knuckle surface, IEEE Transactions on Information Forensics and Security 4 (2009) 98 – 110.

[6] ISO/IEC 19795-1, Information technology – biometric performance testing and reporting – part 1: Principles and framework (2006).

[7] S. Benjio, J. Mariethoz, M. Keller, The expected performance curve, in: In Proceedings of the 22nd International Conference on Machine Learning, 2005, pp. 9 – 16.

[8] Polyu fkp database, `http://www.comp.polyu.edu.hk/~biometrics/FKP.htm`.

[9] R. Giot, M. El-Abed, C. Rosenberger, Greyc keystroke : a benchmark for keystroke dynamics biometric systems, in: IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS), 2009.

[10] U. of Essex, Faces94 database, face recognition data (1994).
URL `http://cswww.essex.ac.uk/mv/allfaces/faces94.html`

[11] A. Martinez, R. Benavente, The ar face database, Tech. rep., CVC Technical Report 24 (1998).

[12] P. Phillips, H. Wechsler, J. Huang, P. Rauss, The FERET database and evaluation procedure for face recognition algorithms, Journal of Image and Vision Computing 16 (5) (1998) 295–306.

[13] J. P. Phillips, H. Moon, S. A. Rizvi, P. J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (2000) 1090–1104.

[14] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, 2005, pp. 947–954.

[15] K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre, XM2VTSDB: The Extended M2VTS Database, in: Second International Conference on Audio- and video-based Biometric Person Authentication (AVBPA'99), 1999, pp. 72–77.

[16] V. Popovici, J. Thiran, E. Bailly-Bailliere, S. Bengio, F. B. M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. M. B. Ruiz, F. Poiree, The BANCA database and evaluation protocol, in: 4th International Conference on Audio- and Video-Based Biometric PersonAuthentication, Guildford, UK, Vol. 2688, 2003, pp. 625–638.

[17] Biosecure Multimodal Biometric Database, http://www.biosecure.info/ (2008).

[18] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) (1997) 711–720.

[19] J. Yang, C. Liu, Horizontal and vertical 2dpca-based discriminant analysis for face verification on a large-scale database, IEEE Trans. o 2 (4) (2007) 781–792.

[20] J. Mahier, M. El-Abed, B. Hemery, C. Rosenberger, Toward a distributed benchmarking tool for biometrics, in: In IEEE International Conference on High Performance Computing & Simulation (HPCS'11), 2011, pp. 1–6, supplied as additional material `Article_hpbench.pdf`.

[21] D. Avison, Merise: A european methodology for developing information systems, European Journal of Information Systems 1 (3) (1991) 183–192.

[22] L. Zhang, L. Zhang, D. Zhang, Finger-knuckle-print: a new biometric identifier, in: Proceedings of the IEEE International Conference on Image Processing, 2009, pp. 1981–1984.

[23] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, Surf: Speeded up robust features, Computer Vision and Image Understanding (CVIU) 110 (2008) 346–359.

[24] B. Hemery, J.-J. Schwartzman, C. Rosenberger, Study on color spaces for single image enrolment face authentication, in: IAPR International Conference on Pattern Recognition (ICPR), 2010, pp. 1249–1252.

[25] D. G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vision 60 (2004) 91 – 110.