Evaluation of Biometric Systems : A Study of Users' Acceptance and Satisfaction

Mohamad El-Abed, Romain Giot, Baptiste Hemery, Christophe Rosenberger

Evaluation of biometric systems: a study of users' acceptance and satisfaction

Mohamad El-Abed*, Romain Giot,
Baptiste Hemery and Christophe Rosenberger

Université de Caen Basse-Normandie,
UMR 6072 GREYC, F-14032 Caen, France
and
ENSICAEN, UMR 6072 GREYC,
F-14050 Caen, France
and
CNRS, UMR 6072 GREYC,
F-14032 Caen, France
E-mail: mohamad.elabed@ensicaen.fr
E-mail: romain.giot@ensicaen.fr
E-mail: baptiste.hemery@ensicaen.fr
E-mail: christophe.rosenberger@ensicaen.fr
*Corresponding author

Abstract: This paper presents a modality-independent evaluation methodology to study users' acceptance and satisfaction of biometric systems. It uses a survey questionnaire for data collection, and some data-mining tools for their analysis. We have applied it on two biometric systems developed in our research laboratory. The results from this survey show the necessity of taking users' point of view when designing and evaluating biometric systems. A panel of 100 volunteers was more satisfied from the keystroke system than the face one. Users surprisingly considered that its perceived performance was also better, even if the used face system has a better performance with an EER of 8.76% than the keystroke one with an EER of 17.51%. The robustness of a system against attacks, computation time required during the verification phase and its easiness to use have been identified as important factors influencing their opinions regarding the tested systems.

Keywords: biometrics; evaluation; users' acceptance and satsifaction; Kruskal-Wallis test; Bayesian networks; decision trees.

Reference to this paper should be made as follows: El-Abed, M. and Giot, R. and Hemery, B. and Rosenberger, C. (xxxx) 'Evaluation of biometric systems: a study of users' acceptance and satisfaction', *Int. J. Biometrics*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Mohamad El-Abed is an Assistant Professor at ENSICAEN. He obtained his PhD from the University of Caen Basse-Normandie in 2011. He belongs to the GREYC laboratory in the e-payment and biometrics research unit. His research interests include biometrics, especially the evaluation of biometric systems.

Romain Giot obtained his Master of Science from ENSICAEN in 2008. He has been a Research Engineer in the GREYC laboratory during two years. His research interests include biometrics, especially the definition of keystroke dynamics biometric and multibiometrics systems. He is now a PhD Student at the University of Caen Basse-Normandie and works on template update strategies for biometric systems.

Baptiste Hemery is an Assistant Professor at ENSICAEN. He obtained his PhD from the University of Caen Basse-Normandie in 2009. He belongs to the GREYC laboratory in the e-payment and biometrics research unit. His research interests concern image interpretation evaluation and biometric systems.

Christophe Rosenberger is a Full Professor at ENSICAEN. He obtained his PhD from the University of Rennes I in 1999. Since 2007, he belongs to the GREYC laboratory where he leads the e-payment and biometrics research unit. His research interests concern the definition of biometric systems, their evaluation and the protection of biometric templates.

1 Introduction

Biometrics is considered as a promising solution among traditional methods based on "what we own" (such as a key) or "what we know" (such as a password). It is based on "what we are" and "how we behave". Biometric authentication systems have many applications (Jain et al., 2004): border control, e-commerce, etc. The main benefits of this technology are to provide a better security, and to facilitate the authentication process for a user. Also, it is usually difficult to copy the biometric characteristics of an individual than most of the other authentication methods such as passwords.

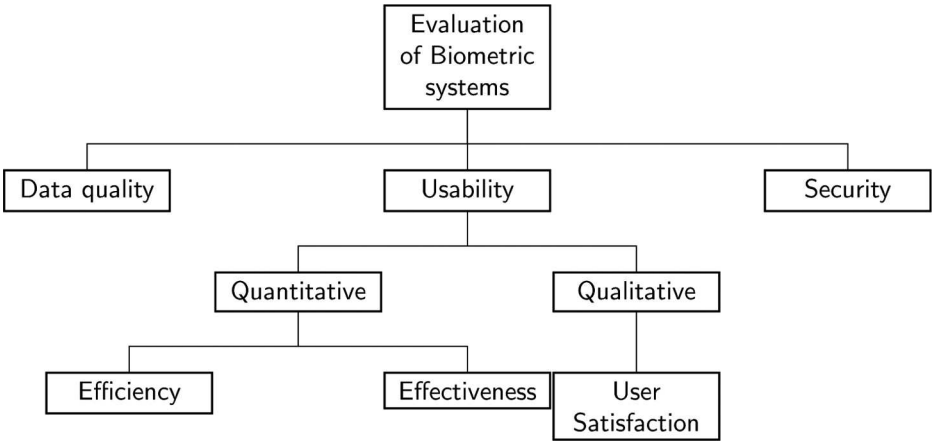
Different biometric modalities have been proposed in the literature, which can be categorised into three kinds of modalities:

- biological
- behavioural
- morphological modalities.

Biological modalities are known to be the most expensive owing to the computation time, and the specific materials required to verify the user's identity. Performance of behavioural modalities provides a lower quality than the others because they depend a lot on user's feelings at the moment of the data acquisition: user may change his/her way of performing tasks owing to its stress, tiredness, concentration or illness. A survey of the literature on behavioural modalities is given by Yampolskiy and Govindaraju (2008). Each modality has its own advantages and drawbacks. A previous work (Mahier et al., 2008) summarises a comparative study of biometric modalities in terms of universality, uniqueness, permanency, collectability, acceptability and performance.

Despite the obvious advantages of biometric systems, their proliferation was not as much as attended. To be used in an industrial context, the quality of a biometric system must be precisely quantified. We need a reliable evaluation methodology to put into obviousness the benefit of a new biometric system. Nowadays, several studies have been done in the literature to evaluate biometric systems. It is generally realised within three aspects as illustrated in Figure 1:

Figure 1 Evaluation aspects of biometric systems



1 *Data quality*: Measures the quality of the biometric raw data (Tabassi and Wilson, 2005). Using quality information, the bad-quality samples can be removed during enrolment or rejected during verification. Such information could also be used in soft biometrics or multimodal approaches (Kryszczuk et al., 2009). Such type of assessment is generally used to enhance the system performance, and could be used to quantify the quality of biometric sensors.

2 *Usability*: According to the international standard ISO 13407:1999 [ISO 13407:1999], usability is defined as "*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*".

- Efficiency, which means that users must be able to accomplish the tasks easily and in a timely manner. It is generally measured as a task time.
- Effectiveness, which means that users are able to complete the desired tasks without too much effort. It is generally measured by common metrics including completion rate and number of errors such as Failure-To-Enrol rate (FTE) as (ISO/IEC 19795-1, 2006).
- User satisfaction, which measures users' acceptance and satisfaction regarding the system. It is generally measured by studying several properties such as easiness to use and trust, *etc.*

3 *Security*: Measures the robustness of a biometric system (algorithms, architecture and devices) against attacks. Such type of assessment

is important since several works in the literature (such as Ratha et al., 2001) show the vulnerabilities of biometric systems.

Traditional evaluation methods have worked well to evaluate emerging technologies, new biometric modalities and algorithm revisions. Many databases have been collected (such as ENSIB face database (Hemery et al., 2007)), many competitions (such as Fingerprint Verification Competition (Maio et al., 2004)) and platforms have been proposed (such as BioSecure (Petrovska and Mayoue, 2007)) whose objective is mainly to compare enrolment and verification/identification algorithms in the literature. Many metrics have been defined by the International Organisation for Standardization (ISO/IEC 19795-1, 2006) in terms of error computations, time computation, memory allocations, etc. These statistical measures allow in general a precise performance characterisation of a biometric system. Nevertheless, these works are dedicated to quantify the system performance (algorithms, processing time, etc.) without taking into account user's view within the evaluation process. However, the biometric process is considered as a two-way interaction, between the user and the system. Jain et al. (2004) categorise the fundamental barriers in biometrics into three main categories:

- accuracy in terms of errors
- scale or size of the database
- usability in terms of easiness to use, acceptability, etc.

One government can decide that an individual would be identified through a biometric data embedded in the passport. For logical or physical access control in a company, it is more difficult to impose a system that would not be accepted by users. As for example, DNA analysis is one of the most efficient techniques to verify the identity of an individual or to identify him/her. Nevertheless, it cannot be used for logical or physical access control not only for time computation reasons, but also because nobody would be ready to give some blood to make the verification. Therefore, taking into account user's view when designing biometric systems is considered as a crucial requirement to the widespread of use of this technology.

Nowadays, there is a lack of a generic evaluation methodology that takes into account users' acceptance within the evaluation process, which constitutes one of the main drawbacks for biometric systems proliferation. To contribute to solve this problem, we propose in this paper a modality-independent evaluation methodology to study users' acceptance and satisfaction of biometric systems. Such kind of evaluation will:

- increase system performance (Kukula et al., 2009)
- improve the accuracy of the optimistic results provided by biometric system designers in terms of errors (e.g., EER)
- reduce product complexity and increase user satisfaction (Theofanos et al., 2008b).

The plan of the paper is organised as follows: Section 2 presents related previous research on users' acceptance and satisfaction of biometric systems. Section 3

presents the proposed evaluation methodology. We present in Section 4 the experimental results on two biometric systems developed in our research laboratory. Section 5 gives a conclusion and some perspectives of this work.

2 Background

The acceptability of biometric systems is affected by several factors. According to Smith (2003), some members of the Human-Computer Interaction (HCI) community believe that interfaces of security systems do not reflect good thinking in terms of creating a system that is easy to use, while maintaining an acceptable level of security. Nowadays, several studies have been done to quantify users' acceptability and satisfaction of biometric systems such as:

- The Opinion Research Corporation International (ORC, 2002) presents the results of a phone survey conducted in 2001 and 2002. The survey has been conducted among national probability samples of 1017 and 1046 adults, respectively, living in USA. The 2001 study showed that 77% of individuals feel that finger-imaging protects individuals against fraud. For privacy issues, 87% in 2001 and 88% in 2002 are worried for the misuse of personal information. The study indicates a good percentage of acceptance, more than 75%, for US law enforcement authorities requiring fingerprint scans to verify identity for passports, at airport check-ins and to obtain a driver licence (see ORC (2002) for more details).
- The National Institute of Standards and Technology (NIST) has performed a usability test on fingerprints (Theofanos et al., 2007). The survey was conducted on 300 adults recruited from a pool of 10,000 people. There were 151 women and 149 men ranging in ages from 18 to over 65 years. 77% of participants were in favour to provide fingerprint images as a mean of establishing identity for passport purposes. 2% of participants have expressed concerns about the cleanliness of the devices with which they would have physical contact. Another study has been done by NIST to examine the impact on fingerprint capture performance of angling the fingerprint scanners (flat, 10, 20 and 30 degrees) on the existing counter heights (99, 114.3 and 124.5 cm) as is presented in Theofanos et al. (2008a).
- Other studies presented in Deane et al. (1995), Coventry et al. (2003), Moody (2004), Jones et al. (2007), Elliott et al. (2007), Pons and Polak (2008), Giot et al. (2009b), Uzoka and Ndzinge (2009) and El Abed et al. (2010) have highlighted several points about biometrics such as:
  - Acceptance is linked to the number of uses of the biometrics in general, and information provided by the biometric device can also improve user acceptance.
  - There is a potential concern about the misuse of personal data (i.e., templates), which is seen as violating users' privacy and civil liberties. Another important concern is the probability that criminals may perpetrate heinous acts to gain access. This could include stalking or assaulting individuals to steal their biometric information.

- Individuals complain that once the biometric template is stolen, it is compromised forever.
- There are also concerns about hygiene with touching such devices and health risks for more advanced technologies such as iris or retina. According to our knowledge, no paper has emphasised physical harm to users of these systems. But despite of this, several concerns were highlighted along this interaction. Anecdotally, some users of biometrics have complained that hand geometry systems dry their hands, while military aviators participating in an experimental programme voiced concern that retinal scanning would damage their vision with extended use over time.

Most of the works done on evaluating biometric systems have focused on the performance aspect. Limited research has focused on usability issues relating to how users perceive and use biometric systems. However, most of the studies in this area are modality-dependent (such as the usability study presented by Theofanos et al. (2008a)). Hence, it could not be applied to any kind of biometric system. In addition, these studies are based on statistical answers to a questionnaire, but no mature data analysis is conducted for understanding respondents' answers to identify reasons. To contribute to solve this problem, we propose a modality-independent methodology that studies users' acceptance and satisfaction of biometric systems. As for us, taking into account users' point of view within the biometric process (hardware, software and instructional design) is not only beneficial to the end-users, but it will also help to improve the performance and effectiveness of a system.

3 Proposed method

The proposed methodology was designed to quantify users' acceptance and satisfaction when using biometric systems. To accomplish this objective, we developed a survey instrument for data collection. These kinds of surveys enable to gather information to be statistically analysed. The proposed methodology principle is as follows: It collects the data using a survey questionnaire (see Appendix). This step is followed by a pre-processing phase to extract the significant knowledge (Section 3.2). Then, the Kruskal-Wallis (KW) test (Higgins, 2003) is performed to determine if there is a significant relationship between demographic characteristics and respondents' answers (Section 3.3). Data-mining tools are used to explain these answers, to determine the reasons that influence their acceptance and satisfaction of biometric systems (Section 3.4).

3.1 Data collection

The first step of the proposed method consists of creating a satisfaction questionnaire. Existing works presented in Section 2 highlighted different important factors impacting their acceptance of biometric systems such as:

- *Socio-demographic factors*: Such as age and gender

- *Learnability and memorability*: They mainly concern how rapidly a user can use the system after instruction or training.
- *Confidence or trust*: Indicates how the performance of the system is perceived by users. It depends mainly on feedbacks from users and their experience.
- *Easiness to use*: Depends on the quality of the biometric sensor and the ergonomic interface. It may also depend on the time required for verification or identification. For example, if the biometric system takes several minutes between the acquisition of the required data and user identification, users may feel that the biometric system is not easy to use.
- *Privacy issues*: There is a potential risk concerning the misuse of the personal collected data, which is seen as violating user's privacy and civil liberties. Many debates have been conducted over the central storage of biometric templates vs. holding the personal template on a smart card where the verification is locally processed.
- *Physical invasiveness*: The acquisition of biometric data requires a user interaction with the biometric sensor. Depending on the used method, the acquisition of the biometric raw data is performed with or without contact with the biometric sensor.
- *Cultural issues*: The acceptability denotes the way how users perceive the biometric system and interact with it. Acceptability is highly dependent on the culture of users. As for example, cultures with an aversion to touch public surfaces would prefer to use biometric systems with contactless sensors (e.g., iris or palm veins).

The developed questionnaire (see Appendix) aims to extract these factors to study users' acceptance and satisfaction of biometric systems. It was designed to collect demographic, experiential and attitudinal characteristics that might have an impact on or a relationship to respondents' views on the use of biometrics. The questionnaire was created owing to the results of extensive desk research presented in the literature. We took into account these studies and we added new questions to complete it. It also noted unsolicited questions that we found it valuable when collecting two biometric databases (Hemery et al., 2007; Giot et al., 2009a) (both databases are publicly available to the biometric community), during the recent usability studies of biometric systems presented in previous works (Giot et al., 2009b; El Abed et al., 2010), and the opinions of two experts working in the social psychology research topic. This help was important especially for the wording and ordering of the questions. Indeed, the question must be as neutral as possible to avoid bias answers. A 4-point Likert-type scale is used to evaluate respondents answers on the satisfaction questions. This scale requires for the respondent to make a positive or negative choice, which avoids to many neutral answers. The survey questionnaire contains 18 questions divided into two sets:

- *General perception of biometric systems* (Appendix, part B), which contains 7 questions aiming to understand users' experience on biometric technology.

- *Perception of the tested system* (Appendix, part C), which contains 11 questions aiming to measure users' acceptance and satisfaction of the tested system.

In addition to these questions, we request some information on the individual such as gender (Appendix, part A). These demographic characteristics are requested to determine if there are significant relationships between them and respondents' answers on biometric technology and the tested system. We also request the question 16 (Appendix, part C) to identify where the use of the tested system would be appropriate for the user.

3.2 Data pre-processing

Before analysing the pilot data, we use a technique to enhance the accuracy and the reliability of the extracted knowledge. It consists of deleting answers having a predefined number of questions without answers.

3.3 Respondent demographics analysis

To determine whether there is a significant relationship between demographic characteristics and respondents' answers on biometric technology and the tested system, we use the KW test. It is a non-parametric (distribution free) test, which is used to decide whether K independent samples are from the same population. In other words, it is used to test two hypothesis given by equation (1) the null hypothesis H_0 assumes that samples originate from the same population (i.e., equal population means) against the alternative hypothesis H_1 which assumes that there is a statistically significant difference between at least two of the subgroups.

$$\begin{cases} H0 : \mu_1 = \mu_2 = \cdots = \mu k \\ H1 : \mu_i \neq \mu_j \ \exists \ (i,j) \text{ with } i \neq j \end{cases} \tag{1}$$

The KW test statistic H is given by equation (2), and the p-value is calculated using a χ^2 distribution with $k-1$ degrees of freedom. The decision criterion to choose the appropriate hypothesis is given in equation (3).

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{g} n_i \bar{r}_{i.}^2 - 3(N+1) \tag{2}$$

where n_i is the number of observations in group i, r_{ij} is the rank of observation j from group i and N is the total number of observations across all groups.

$$\bar{r}_{i.}^2 = \frac{\sum_{j=1}^{n_i} r_{ij}}{ni} \quad \text{and} \quad \bar{r} = \frac{1}{2}(N+1)$$

$$\begin{cases} p\text{-value} \ \geq 0.05 & \text{accept } H_0 \\ \text{otherwise} & \text{reject } H_0 \end{cases} \tag{3}$$

3.4 Data-mining analysis

To analyse respondents' answers, we use two types of classifiers: *Bayesian networks* (Friedman et al., 1997) and *decision trees* (Breiman et al., 1984). They are formal graphical tools for representation of decision scenarios requiring reasoning under uncertainty. We present in sections 3.4.1 and 3.4.2 the Bayesian networks and the decision trees, respectively. Section 3.4.3 presents its performance evaluation to choose the best decision model.

3.4.1 Bayesian networks

A Bayesian network (B_S, B_P) is a probabilistic graphical model that represents a set of random variables $U = \{x_1, x_2, \ldots x_n\}$ and their conditional independencies. The Bayesian structure B_S is a Directed Acyclic Graph (DAG) where nodes represent propositional variables in a domain, and the arcs between nodes represent the dependency relationships among the variables. The Bayesian probability distributions B_P is a set of probability tables $B_P = \{p(u \mid pa(u)) | u \in U\}$ where $pa(u)$ is the set of parents of u in B_S.

The method used to learn the Bayesian network structure B_S is based on *conditional independence tests* as described in Bouckaert et al. (2009). This method mainly stem, from the goal of uncovering causal structure. The assumption is that there is a network structure that exactly represents the independencies in the distribution that generated the data. The method is divided into two stages:

- *Find a skeleton:* Starting with a complete undirected graph, the method tries to find conditional independencies $\{x \to y\} \cup \forall z \in Z z \to y$ in the data. If an independency is identified, the edge between x and y is removed from the skeleton. We use the following conventions to identify counts in the database D of a network structure B_S:

  - Let r_i $(1 \leq i \leq n)$ be the cardinality of the variables x_i.
  - We denote by q_i the cardinality of the parent set of x_i in the network structure B_S. Hence, q_i can be calculated as the product of cardinalities of nodes in $pa(x_i)$, $q_i = \prod_{x_j \in pa(x_i)} r_j$.
  - We denote by $N_{ij} (1 \leq i \leq n, 1 \leq j \leq q_i)$ the number of records in D for which $pa(x_i)$ takes its jth value.
  - We denote by $N_{ijk} (1 \leq i \leq n, 1 \leq j \leq q_i, 1 \leq k \leq r_i)$ the number of records in D for which $pa(x_i)$ takes its jth value and for which x_i takes its kth value. Hence, $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$.
  - We use N to denote the number of records in D.

  To test whether variables x and y are conditionally independent given a set of variables Z, a network structure with arrows $\forall z \in Z z \to y$ is compared with one with arrows $\{x \to y\} \cup \forall z \in Z z \to y$. A test is performed by using a predefined score metric. In this study, we use four score metrics as defined here

- Entropy metric $H(B_S, D)$ defined as

$$H(B_S, D) = -N \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} \frac{N_{ijk}}{N} \log \frac{N_{ijk}}{N_{ij}} \tag{4}$$

- Akaike Information Criterion metric $Q_{AIC}(B_S, D)$ defined as

$$Q_{AIC}(B_S, D) = H(B_S, D) + K \tag{5}$$

  where $K = \sum_{i=1}^{n} (r_i - 1) \cdot q_i$

- Minimum Description Length metric $Q_{MDL}(B_S, D)$ defined as

$$Q_{MDL}(B_S, D) = H(B_S, D) + \frac{K}{2} \log N \tag{6}$$

- Bayesian metric $Q_{Bayes}(B_S, D)$ defined as

$$Q_{Bayes}(B_S, D) = \prod_{i=0}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(r_i - 1 + N_{ij})!} \prod_{k=1}^{r_i} N_{ijk}! \tag{7}$$

- *Direct acyclic graph (DAG):* the second stage consists in directing all the edges in the skeleton to get a DAG. The first step in directing arrows is to check for every configuration $x - -z - -y$ where x and y not connected in the skeleton whether z is in the set Z of variables that justified removing the link between x and y. If $z \notin Z$, we can assign direction $x \rightarrow z \leftarrow y$. Then, a set of rules (presented in Bouckaert et al., 2009) is applied to direct the remaining edges.

3.4.2 Decision trees

Introduced by Breiman et al. (1984), decision trees are one of the few knowledge representation schemes, which are easily interpreted and may be inferred by very simple learning algorithms (Baldwin and Xie, 2005). A decision tree is a tree in which:

- each internal node tests an attribute
- each branch corresponds to an attribute value
- each leaf node assigns a classification. Decision trees are powerful predictors and provide an explicit concept description for a data set.

Nowadays, several methods have been proposed for learning decision trees. For this study, we have used the most well known and used methods in the literature: C4.5 developed by Quinlan (1993) and CART (Breiman et al., 1984) algorithms. The learning algorithms of decision trees contain three main functions:

- deciding if a node is a leaf
- selecting an attribute for a test node

- associating a label to a leaf.

Before presenting the used algorithms, we introduce the following notions. For a database of observations D, a target variable $C = \{1, \ldots, c\}$ and a decision tree t, we define for each position p in the tree t:

- $N(p)$ is the cardinality of the set of observations associated to the position p in the database D.
- $N(k/p)$ is the cardinality of the set of observations associated to the position p belonging to the class k.
- $P(k/p)$ is defined as:

$$P(k/p) = \frac{N(k/p)}{N(p)} \tag{8}$$

- The diversity functions used as a splitting criterion are the information content (IC) and the Gini index defined as follows:

$$IC(p) = -\sum_{k=1}^{c} P(k/p) \times \log(P(k/p)) \tag{9}$$

$$Gini(p) = 1 - \sum_{k=1}^{c} P(k/p)^2 \tag{10}$$

3.4.2.1 CART algorithm

CART method, introduced by Breiman et al., consists of learning binary decision trees. The CART learning algorithm consists of:

1 *Deciding if a node is a leaf*: A node p is a leaf if $Gini(p) \leq i_0$ or $N(p) \leq n_0$, where i_0 and n_0 are initial parameters of the learning algorithm.

2 *Selecting an attribute for a test node*: For a position p, the algorithm selects the attribute that maximises the gain defined in equation (11) using the $Gini$ index as a splitting criterion.

$$Gain(p, test) = Gini(p) - (P_{left} \times Gini(p_1) + P_{right} \times Gini(p_2)) \tag{11}$$

where P_{left} (respectively P_{right}) represents the proportion of the elements associated to position p and going to the node in position p_1 (respectively, p_2).

3 *Associating a label to a leaf*: The majority rule is used to label each leaf in the decision tree.

3.4.2.2 C4.5 algorithm

C4.5 is an extension of the *ID3* (Quinlan, 1986) algorithm developed by Ross Quinlan in 1986. C4.5 builds decision trees from a set of training data using the concept of information content defined in equation (9). At each node of the tree, C4.5 chooses one attribute of the data that most effectively splits its set of samples

into subsets enriched in one class or the other. Its criterion is based on the gain ratio, defined in equation (12), that results from choosing an attribute for splitting the data. The attribute with the highest gain ratio is chosen to make the decision. The C4.5 algorithm then recurses on the smaller sublists.

$$GainRatio(p, test) = \frac{Gain(p, test)}{SplitInfo(p, test)} \tag{12}$$

where $Gain(p, test)$ and $SplitInfo(p, test)$ are defined as:

$$Gain(p, test) = IC(p) - \sum_{i=1}^{n} P_i \times IC(p_i) \tag{13}$$

$$SplitInfo(p, test) = -\sum_{i=1}^{n} P\prime(i/p) \times \log(P\prime(i/p)) \tag{14}$$

where $test$ is the test attribute having n values, $P\prime(i/p)$ is the proportion of elements in D at position p and satisfying the ith test attribute value.

3.4.3 Performance metrics

Classifiers are useful tools, which are commonly used in decision analysis, to help identifying a strategy most likely to reach a goal (Edwards et al., 2008). They provide a highly effective and simple structure that can be explored to make predictions and decisions. Despite the obvious advantages of these tools, they do not provide a 100% accuracy result. Because of this inaccuracy, several performance criteria have been proposed in the literature (Rakotomalala, 1997) to identify the quality of a classifier. The main criteria used are:

- *Accuracy*: Denotes the percentage of the correctly classified instances.
- *Area under the ROC curve (AUC)*: It is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. In this study, AUC is estimated using Mann-Whitney statistic test as presented by Ling et al. (2003). The AUC of a classifier G is defined as:

  $$\widehat{\mathrm{AUC}} = \frac{S_0 - n_0(n_0 + 1)/2}{n_0 n_1} \tag{15}$$

  where n_0 and n_1 are the numbers of positive and negative examples respectively, and $S_0 = \sum r_i$, where r_i is the rank of the ith positive example in the ranked list. Ling et al. (2003) suggest that its use should replace accuracy when measuring and comparing classifiers: the best classifier is the one with the largest *AUC* value.
- *Comprehensibility*: Qualifies the exploitability of the produced model. For example, in a Bayesian network, the important number of a node's parents affects the identification of its strong relations with them.

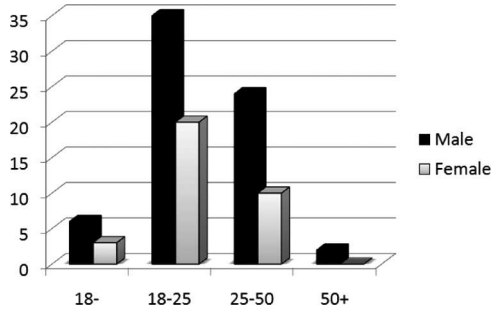
4 Experimental results

In this section, we detail the experimental protocol and the results we obtained. We present first the test protocol and test materials used in this study, followed by the pre-processing phase prior to the data analysis. Then, we present the data analysis phase to study users' acceptance and satisfaction of the two tested biometric systems (GREYC-Keystroke (Giot et al., 2009a) and GREYC-Face (Hemery et al., 2010)) developed in the GREYC research laboratory.

4.1 Test protocol

The pilot study was distributed on a paper sheet to a sample of 100 volunteers, including students (70% of the population) coming from different countries and employees. Tests have been conducted in public places over a 3-month period. It consists in testing both systems (*enrolment* then multiple *verification* attempts playing the role of an impostor and a legitimate user). Then, they were requested to answer a questionnaire (see Appendix): part A, part B and two times part C (one for each tested system). Volunteers completed the survey voluntarily and received none remuneration. During the tests, volunteers were informed about the purpose of the study, and their responses would be confidential and anonymous. The age and gender distribution of the volunteer crew are shown in Figure 2.

Figure 2 Age and gender distribution of the volunteer crew



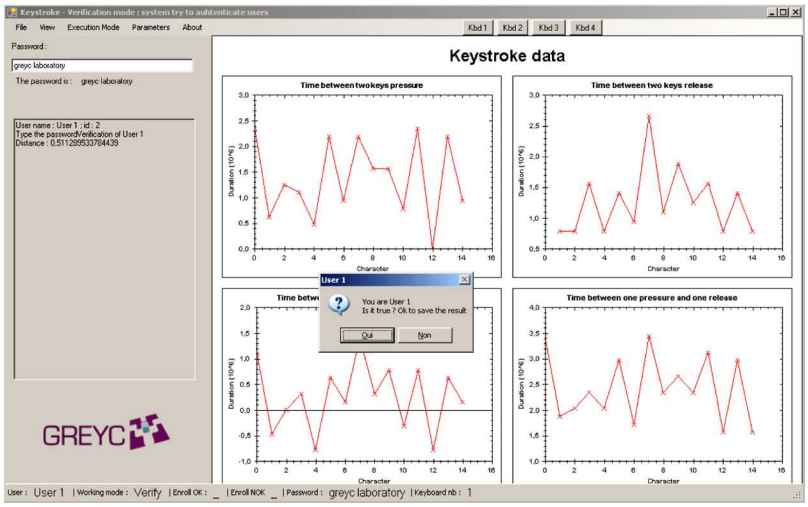
4.2 Test materials

In this study, we have used two biometric verification systems developed in our research laboratory: GREYC-Keystroke and GREYC-Face. The first is a behavioural-based analysis, while the second is a morphological-based analysis. We have chosen to illustrate the proposed methodology on these two biometric systems for many reasons. First, we would like to quantify the acceptability of our developed systems to improve them (in terms of performance and ergonomic interface based on users feedbacks). We are mainly interested of these modalities since they belong to the possible candidates that may be implemented in an Automated Teller Machine (ATM), and can be used for e-commerce applications. Second, it would be important to test which kind of modality (morphological or behavioural) is better perceived by the volunteers. Obviously, we have to

include other types of systems (e.g., iris verification) to test in an accurate way this hypothesis. Third, we would like to see which type of modality is better perceived to be used to manage physical (e.g., border control) or logical (e.g., e-commerce) access. Finally, evaluation results provided by these two very different biometric systems should be also as different as possible. Therefore, we believe that it is a good choice to illustrate this evaluation methodology on these systems. The performances of tested systems are calculated with captures provided by our volunteer crew. We plot their Detection Error Trade-off (DET) curves in Figure 5:

- *GREYC-Keystroke verification system*: It is a biometric system based on behavioural analysis (see Figure 3). The main goals of this software are to allow the creation of a keystroke dynamics database and to compare different algorithms in the literature, within the same conditions (e.g., acquisition conditions), for evaluation issues. The system provides an EER value equal to 17.51% on a database composed of 70 individuals with 3 vectors used for *enrolment* and 2 for the *tests*. The system implements a score-based method presented by Hocquet et al. (2007). To achieve the enrolment in the system, users type 5 times a predefined password 'greyc laboratory'. For the verification process, users tried freely (i.e., in term of number of attempts) both genuine and impostor attempts. For each user, we have at least two genuine attempts and one impostor attempt.

Figure 3 GREYC-Keystroke verification system (see online version for colours)



- *GREYC-Face verification system*: It is a biometric system based on morphological analysis (see Figure 4). The system implements a SIFT-based (Lowe, 2004) algorithm. The used matching similarity principle is described in a previous work (Hemery et al., 2010). The system provides an EER value equal to 8.76% on a database composed of 70 individuals with 1 image used for *enrolment* and 2 for the *tests*. To achieve the enrolment in the system,

a single picture is captured per user. For the verification process, users tried the face system using the same process than the previous one.

Figure 4 GREYC-Face verification system. An example of a matching attempt resulting from a genuine user (see online version for colours)

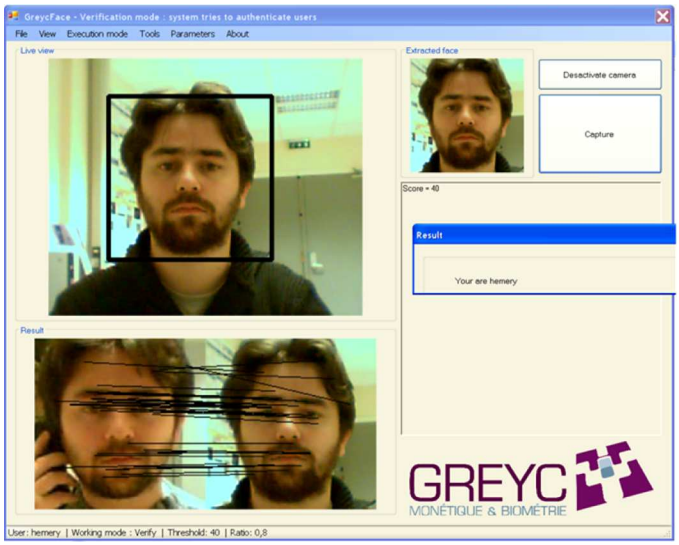
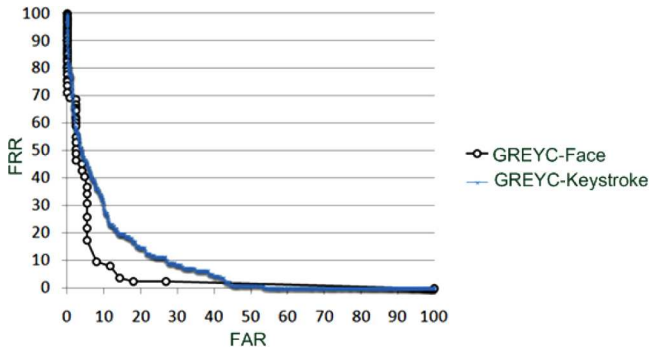


Figure 5 DET curves for the two tested biometric systems (see online version for colours)



4.3 Data pre-processing

The first step of the proposed methodology consists in the deletion of respondents' answers that did not answer a certain number of questions of the questionnaire. For three unanswered questions, two vectors of answers (one from each system) have been eliminated from this study. Therefore, the results presented in the next subsections are done using 99 vectors of answers on both systems.

4.4 Respondent demographics analysis

We study in this section the relationship, on each system, between respondents' demographic characteristics and their answers on biometric technology and satisfaction questions (Appendix, parts B and C). Table 1 shows the results for a confidence degree equal to 95%. Bold values indicate significant relationships based on the criterion defined in equation (3). From this table, we can put into obviousness these significant relationships:

- Age was significantly related to their answers about keystroke's robustness against attacks: aged respondents (≥ 28) considered that the system is less robust against attacks than youngest ones.
- For keystroke system, education level was significantly related to the disturbed, threats to privacy, verification quickness and correct answer factors:
    - high school graduate respondents were less disturbed than the others.
    - none graduated respondents have expressed much more concerns about their privacy than the others.
    - high school graduate respondents considered that the computation time during the verification phase is faster than the college graduate respondents.
    - keystroke performance was perceived better by the high school graduate respondents than the college graduate respondents.

Table 1 Respondents' answers and demographic factors, Kruskal-Wallis analysis: lines with two p-values correspond to systems' specific questions (face/keystroke)

Questions	*p-value* *Gender*	*Age*	*Education*	*Profession*
Biometric technology knowledge	0.089	0.652	0.134	0.911
Awareness about fraud identity	0.247	0.831	0.14	0.578
Secret-based against fraud	0.482	0.804	0.213	0.778
Biometric-based against fraud	0.71	0.324	0.74	0.546
Disturbed	0.419/0.385	0.509/0.473	0.096/**0.007**	0.696/0.428
Threats to privacy	0.206/0.556	0.65/0.649	0.196/**0.044**	0.756/0.36
Easiness to use	0.145/0.438	0.063/0.522	**0.012**/0.392	0.46/0.406
Verification quickness	0.135/0.144	0.226/0.294	0.195/**0.039**	0.095/0.202
Correct answer	0.982/0.075	0.779/0.717	0.78/**0.034**	0.369/0.058
System can be easily attacked	0.276/0.226	0.492/**0.02**	0.558/0.204	0.405/0.22
Use in the future	0.079/0.48	0.604/0.883	0.721/0.821	0.33/0.255
Trust	0.414/0.689	0.469/0.218	0.709/0.947	0.372/0.068
General appreciation	0.078/0.129	0.984/0.873	0.3/0.768	0.459/0.917

- Education level was significantly related to the use of face system: college graduated respondents found it more complicated (in term of easiness to use) than the others.

4.5 Comparative study of the studied systems

In this section, we present respondents' knowledge about biometric technology, and a comparative analysis between the studied systems based on a statistical analysis of their answers. A Kruskal-Wallis test was performed to identify the significant differences among this comparison. Table 2 shows the results for a confidence degree equal to 95%. Bold values indicate significant relationships based on the criterion defined in equation (3). From the respondents' answers and Table 2, we can put into obviousness some interesting points:

- Most of the respondents (74.8%) have already heard before our study of biometric authentication systems, and less than half of them (40.4%) have already used a biometric system.
- 38.4% of the respondents have expressed a good knowledge about biometric technology.
- Using Kruskal-Wallis test (p-value <0.01), respondents considered that biometric technology (90.9% agree) is much more appropriate than secret-based solutions (33.3% agree) against fraud.
- There were no significant differences on easiness to use, verification quickness, use in the future and trust factors: 24.2% of the respondents have found that face system is not easy to use and 14.1% for keystroke one. 10.1% found that the computation time during the verification phase of face system is not fast and 14.1% for keystroke one. 22.2% of the participants hesitate or refuse the use of face system in the future and 17.2% for keystroke one. For their perception about trust, 32.3% do not trust face system and 20.2% for keystroke one.
- There were significant differences on disturbed, threats to privacy, correct answer, system can be easily attacked and general appreciation factors: respondents were much more disturbed while using face system (25.3%) than keystroke one (16.2%). Respondents have expressed much more concerns about their privacy while using face system (47.5%) than keystroke one (13.1%). They found that keystroke performance is better than the face one. They found that keystroke system (53.5%) is more robust against attacks than face one (34.3%). For their general appreciation, they were more satisfied from the use of keystroke system (88.9%) than face one (75.8%).
- Finally, 23.2% prefer to use the face system and 62.6% for keystroke one for managing logical access, 41.4% prefer to use the face system and 14.1% for keystroke one for physical access, 31.3% prefer to use the face system and 21.2% for keystroke one for both kinds of access. This indicates that the keystroke system is more requested to be used for managing logical access, while the other system for physical access (which is of course an expected result).

Table 2 Comparative analysis of acceptance and satisfaction between the studied systems, Kruskal-Wallis analysis

	Mean ratings of the studied systems		
Satisfaction questions	*Face system*	*Keystroke system*	*p-value*
Disturbed	1.916	1.67	**0.034**
Threats to privacy	2.427	1.626	$<<$ **0.05**
Easiness to use	3.133	3.242	0.355
Verification quickness	3.329	3.357	0.611
Correct answer	3.176	3.51	**0.007**
System can be easily attacked	2.659	2.3	**0.005**
Use in the future	2.989	3.112	0.238
Trust	2.84	3	0.176
General appreciation	2.823	3.175	**0.0021**

4.6 Discussion

The results of this study and the statistical analysis of answers brought many interesting information. We found a surprising high rate (47.5%) concerning their concerns about privacy issues while using face system. The results also brought surprising rates concerning the perceived performance of the tested systems, and their general appreciation. Respondents found that keystroke performance (with an $EER = 17.51\%$) is better than the face one (with an $EER = 8.76\%$), and they were more satisfied from the keystroke system (88.9%) than the face one (75.8%).

Therefore, it would be important to explain respondents' answers to more understand these rates, and the significant differences among the studied systems. This is what we present in the next section.

4.7 Data-mining analysis

The purpose of this section is to study the dependences between satisfaction questions (Appendix, parts B and C except Q_{16}) to understand respondents' answers. We would also like to more understand the surprising rates provided by the previous section. Owing the nature of construction of Bayesian networks and decision trees (i.e., the target question should be a nominal attribute), we proceed as follows:

1 Since the target question should be nominal, we divide the satisfaction questions into two sets (S_{Cause} and S_{Effect}), according to the Cause and Effect relationship. We consider that the questions in $S_{Cause} = \{Q_i / i = 4, 5, 6, 7, 11, 12, 14\}$ set are numerical attributes, while questions in $S_{Effect} = \{Q_i / i = 9, 10, 13, 15, 17, 18\}$ set are nominal attributes. For example, we put concerns about privacy issues question (Q_{10}) in S_{Effect} set, since we would like to explain why we have a high rate (47.5%) concerning privacy issues while using face system.

2 To generate the Bayesian networks and decision trees models, missing values (i.e., questions without answers) are handled for both kinds of attributes.

For nominal attributes, they are replaced by the most frequent choice. While for numerical attributes, they are replaced by the average value.

Using Bayesian networks and Decision Trees, several points can be concluded:

- Using Table 3, respondents' concerns about their privacy while using the face system can be explained by their perception about its robustness against attacks. From the respondents who have expressed concerns about their privacy while using face system, more than half of them (66%) found that it can be easily attacked. Since more than half of the respondents (51.5%) do not found the face system robust against attacks, this explains why a lot of respondents (47.5%) have expressed such concerns. Using both clauses C_1 and C_2 (Table 3), we can also deduce that the easiness to use face system is an another possible candidate having impact on respondents' concerns about their privacy.

Table 3 Excerpt from decision tree explaining respondents' answers for privacy issues of the face verification system. Bold rules indicate important ones

if (System can be easily attacked <=2) then
... ...
if (System can be easily attacked >2) then
... ...
if (System can be easily attacked <=3) then
... ...
if (Verification fast <=3) then
if(Easy to use <=3) then intrusive (20.0/3.0) (C_1)
if (Easy to use >3) then
... ...
if (Password appropriate solution >2) then intrusive (5.0/2.0)
if (Verification fast >3) then
if (Fraud awareness <=2) then
... ...
if (Easy to use >3) then intrusive (6.0/3.0)
if (Fraud awareness >2) then intrusive (6.0/2.0)
if (System can be easily attacked >3) then
if (Easy to use <=3) then quite intrusive (5.0/1.0) (C_2)
... ...
Accuracy: 76.77%
AUC:
not at all intrusive: 0.94, not intrusive:0.938, intrusive:0.916 and quite intrusive:0.919

- Respondents' perception about keystroke performance (Table 4) was related to the robustness of the system against attacks, and its easiness to use. From the respondents who found that keystroke performance is important, 23.5% of them considered that the system is quite easy to use, and 24.7% found that the system is robust against attacks. For the face system (Table 5), it was related to their perception about its robustness against attacks, and their awareness about fraud identity. From the respondents who do not found that face performance important, 15% of them found that their awareness about fraud identity is not at all important, and 20% considered that the system is not robust against attacks. From these results, we find that the robustness of the system against attacks is a possible candidate having impact on their perception about the performance of both systems. Since respondents significantly (with a p-value $= 0.005$) considered that the robustness of

keystroke system against attacks is better than the face system, and most of them (85.9%) considered that keystroke system is easy to use, this clearly explains why respondents found the performance of the keystroke system is better than the face one.

Table 4 Excerpt from decision tree explaining respondents' answers of the keystroke performance. Bold rules indicate important ones

if(Fraud awareness <=2) then
... ...
if(Easy to use >3) then always (25.0/5.0)
if(Fraud awareness >2) then
if(System can be easily attacked <=2) then
if(Secret-based against fraud <=1) then
... ...
if(Secret-based against fraud >1) then always (24.0/3.0)
... ...
Accuracy:81.82%
AUC:
rarely:0.894, sometimes:0.887 and always:0.867

Table 5 Excerpt from decision tree explaining respondents' answers of the face verification performance. Bold rules indicate important ones

if (Easy to use <=3) then
if(Fraud awareness <= 2) then
if(Biometric technology knowledge <= 1) then
if(System can be easily attacked <= 2) then always (6.0/2.0)
if(System can be easily attacked > 2) then rarely (2.0)
if(Biometric technology knowledge > 1) then
... ...
if(Fraud awareness <= 1) then rarely (4.0/1.0)
... ...
if(Fraud awareness > 2) then
... ...
if(Easy to use > 3) then
if(System can be easily attacked <= 2) then always (19.0/2.0)
if(System can be easily attacked > 2) then
... ...
if(Secret-based against fraud > 2 then rarely (3.0/1.0)
... ...
Accuracy: 73.74%
AUC:
never:0.962, rarely:0.869, sometimes:0.877 and always:0.901

- Respondents' general appreciation on both systems (Tables 6 and 7) was related to the robustness of the system against attacks, and its easiness to use. From the respondents who were not satisfied while using face verification system, 57.1% of them do not found the system is robust against attacks, and 14.3% considered that it is not easy to use. From the respondents who were satisfied from keystroke system, at least 12.5% of them found that keystroke system is robust against attacks, and almost 40.9% found that the system is

easy to use. In addition, the computation time during the verification phase of keystroke system is also considered as a possible candidate having impact on respondents' general appreciation. Therefore, we find that respondents' perception about the robustness of the system against attacks is the major reason explaining why they were more satisfied from keystroke system than face one.

Table 6 Excerpt from decision tree explaining respondents' answers for their general appreciation of the face verification system. Bold rules indicate important ones

if(System can be easily attacked <= 2) then
... ...
if(System can be easily attacked > 2) then
if(Easy to use <= 3) then
if(Biometric-based against fraud <= 2 then not satisfied (2.0)
if(Biometric-based against fraud > 2) then
... ...
if(Easy to use <= 2) then
if (Heard before =no) then not at all satisfied (3.0)
... ...
if(Easy to use > 2) then
... ...
if(Secret-based against fraud <= 1) then not satisfied (2.0)
if(Secret-based against fraud > 1) then not at all satisfied (6.0/1.0)
... ...
Accuracy: 77.78%
AUC:
not at all satisfied:0.958, not satisfied:0.872, satisfied:0.761, and quite satisfied:0.738

Table 7 Excerpt from decision tree explaining respondents' answers for their general appreciation of the keystroke system. Bold rules indicate important ones

if(System can be easily attacked <=1) then
if(Biometric technology knowledge <=2) then quite satisfied (9.0/2.0)
if(Biometric technology knowledge >2) then satisfied (4.0)
if(System can be easily attacked >1) then
... ...
if(Secret-based against fraud <=2) then
if(Easy to use <=1) then quite satisfied (2.0/1.0)
if(Easy to use >1) then satisfied (48.0/12.0)
if(Secret-based against fraud >2) then
if(Verification fast <=2) then not satisfied (4.0/2.0)
if(Verification fast >2) then
... ...
Accuracy: 76.77%
AUC:
not at all satisfied:0.964, not satisfied:0.953, satisfied:0.817, and quite satisfied:0.814

5 Conclusion and perspectives

The study of users' acceptance and satisfaction of biometric systems is considered as an important factor to take into account when designing and evaluating such systems (Theofanos et al., 2008b). Despite this, existing studies in the literature are

very few in comparison with performance ones. However, most of the studies in this area are modality-dependent (such as the usability study presented by Theofanos et al. (2008a)). In addition, these studies are based on statistical answers to a questionnaire, but no data analysis is conducted for understanding respondents' answers to identify reasons. To contribute in more taking into account users' point of view when designing biometric systems, we proposed in this paper a modality-independent evaluation methodology to study users' acceptance and satisfaction of biometric systems. The methodology is based on a survey questionnaire for data collection. It uses

- the Kruskal-Wallis test to extract significant relationships between demographical characteristics and satisfaction questions
- two data-mining tools, Bayesian networks and decision trees, that illustrate dependencies to analyse respondents' answers and behaviours.

The main advantage of the proposed methodology is twofold. First, it is independent from the tested modality. Hence, it could be applied to any kind of biometric systems. Second, it uses data-mining tools to explain respondents' answers. Such kind of analysis would determine the possible candidates that may influence their acceptance and satisfaction in a specified context of use and a target population. We have illustrated the proposed methodology on a crew of 100 volunteers, using two biometric systems developed in our research laboratory for clarifying its benefits. The first system is a morphological-based analysis (face verification with an EER equal to 8.76%), the second one is a behavioural-based analysis (keystroke dynamics with an EER equal to 17.51%). The main results outlined from this study are:

- Respondents considered that biometric-based technology is more appropriate than secret-based solutions against fraud.
- Socio-demographic characteristics have influenced their answers on some satisfaction questions.
- Both systems are acceptable and respondents were more satisfied with the keystroke system (88.9%) than the other one (75.8%).
- The robustness of a system against attacks, its easiness to use and the computation time during the verification phase are identified as important factors influencing respondents acceptance and satisfaction.
- Finally, from the volunteers who have willingness to use the studied systems in the future, the keystroke system was more requested to be used to manage logical access and the other system for physical access.

These findings clearly show that users' acceptance and satisfaction should be taken into account when developing and evaluating biometric systems. Even if the performance of a biometric system outperformed another one, this will not necessarily mean that it will be more operational or acceptable. In our opinion, there is a lack of an evaluation methodology that more take into account users' point of view when designing and evaluating biometric systems, which constitutes one of the main drawbacks of biometric systems proliferation.

For the perspectives, we intend to develop a web-based software to compare different types of biometric systems (fingerprint, iris, keystroke dynamics, signature dynamics and face verification), and to have a much larger population.

Terms and definitions

Authentication: Security measure designed to establish the validity of a transmission, message, or originator, or a means of verifying an individual's authorisation to receive specific categories of information.

Biometric: Any specific and uniquely identifiable physical human characteristic (e.g., of the retina that may be used to validate the identity of an individual).

enrolment: The process of collecting biometric samples from a person and the subsequent preparation and storage of biometric reference templates representing that person's identity.

False Acceptance Rate (FAR): Rate at which an impostor is accepted by an authentication system.

False Rejection Rate (FRR): Rate at which the authorised user is rejected from the system.

Equal Error Rate (EER): This error rate equates to the point at which the FAR and FRR cross (compromise between FAR and FRR).

Vulnerability: A weakness in the system that can be exploited to violate its intended behaviour.

Threat: A potential event that could compromise the security integrity of the system.

Acknowledgement

The authors thank the Basse-Normandie Region and the French Research Ministry for their financial support of this work.

References

Baldwin, J.F. and Xie, D. (2005) *Intelligent Information Processing II*, Chapter Simple fuzzy logic rules based on fuzzy decision tree for classification and prediction problem, Springer-Verlag, pp.175–184.

Bouckaert, R.R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A. and Scuse, D. (2009) *Weka Manual*, Technical report, Department of Computing Science, University of Waikato, New Zealand.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984) *Classification and Regression Trees*, Wadsworth International Group, 359 pages.

Coventry, L., De Angeli, A. and Johnson, G. (2003) 'Honest it's me! self service verification', *The ACM Conference on Human Factors in Computing Systems (CHI)*, pp.1–4.

Deane, F., Barrelle, K., Henderson, R. and Mahar, D. (1995) 'Perceived acceptability of biometric security systems', *Computers & Security*, Vol. 14, pp.225–231.

Edwards, B., Zatorsky, M. and Nayak, R. (2008) 'Clustering and classification of maintenance logs using text data mining', *Seventh Australasian Data Mining Conference (AusDM 2008)*.

El Abed, M., Giot, R., Hemery, B. and Rosenberger, C. (2010) 'A study of users' acceptance and satisfaction of biometric systems', *International Carnahan Conference on Security Technology (ICCST)*, pp.170–178.

Elliott, S.J., Massie, S.A. and Sutton, M.J. (2007) 'The perception of biometric technology: A survey', *Automatic Identification Advanced Technologies*, pp.259–264.

Friedman, N., Geiger, D. and Goldszmidt, M. (1997) 'Bayesian network classifiers', *Machine Learning*, Vol. 29, pp.131–163.

Giot, R., El Abed, M. and Rosenberger, C. (2009a) 'Greyc keystroke: a benchmark for keystroke dynamics biometric systems', *IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp.1–6.

Giot, R., El Abed, M. and Rosenberger, C. (2009b) 'Keystroke dynamics authentication for collaborative systems', *Collaborative Technologies and Systems, International Symposium*, Washington, DC, USA, pp.172–179.

Hemery, B., Rosenberger, C. and Laurent, H. (2007) 'The ENSIB database: a benchmark for face recognition', *International Symposium on Signal Processing and its Applications (ISSPA), special session "Performance Evaluation and Benchmarking of Image and Video Processing"*, United Arab Emirates (U.A.E.), Sharjah, pp.459–464.

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

Higgins, J.J. (2003) *An Introduction to Modern Nonparametric Statistics*, The American Statistician, 500 pages.

Hocquet, S., Ramel, J.Y. and Cardot, H. (2007) 'User classification for keystroke dynamics authentication', *International Conference on Biometrics (ICB)*, pp.531–539.

ISO 13407:1999 (1999) *Human Centred Design Process for Interactive Systems.*

ISO/IEC 19795-1 (2006) *Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*.

Jain, A.K., Pankanti, S., Prabhakar, S., Hong, L. and Ross, A. (2004) 'Biometrics: a grand challenge', *International Conference on Pattern Recognition (ICPR)*, Vol. 2, pp.935–942.

Jones, L.A., Antón, A.I. and Earp, J.B. (2007) 'Towards understanding user perceptions of authentication technologies', *ACM Workshop on Privacy in the Electronic Society*, Alexandria, Virginia, USA, pp.91–98.

Kryszczuk, K., Richiardi, J. and Drygajlo, A. (2009) 'Impact of combining quality measures on biometric sample matching', *Proceedings of the 3rd IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Washington, DC, pp.1–6.

Kukula, E.P., Blomeke, C.R., Modi, S.K. and Elliott, S.J. (2009) 'Effect of human-biometric sensor interaction on fingerprint matching performance, image quality and minutiae count', *International Journal of Computer Applications in Technology*, Vol. 34, No. 4, pp.270–277.

Ling, C.X., Huang, J. and Zhang, H. (2003) 'AUC: a better measure than accuracy in comparing learning algorithms', *Canadian Conference on Artificial Intelligence*, Vol. 2671, pp.329–341.

Lowe, D.G. (2004) 'Distinctive image features from scale-invariant keypoints', *International Journal of Computer Vision (IJCV)*, Vol. 60, pp.91–110.

Mahier, J., Pasquet, M., Rosenberger, C. and Cuozzo, F. (2008) 'Biometric authentication', *Encyclopedia of Information Science and Technology*, pp.346–354.

Maio, D., Maltoni, D., Wayman, J.L. and Jain, A.K. (2004) 1Fvc2004: third fingerprint verification competition', *Proceedings of the 1st International Conference on Biometric Authentication*, Hong Kong, pp.1–7.

Moody, J. (2004) 'Public perceptions of biometric devices: the effect of misinformation on acceptance and use', *The Informing Science and Information Technology Education*, Vol. 1, pp.753–761.

ORC (2002) *Public Attitudes Toward the Uses of Biometric Identification Technologiesby Government and the Private Sector*, Technical Report, Opinion Research Corporation International (ORC).

Petrovska, D. and Mayoue, A. (2007) *Description and Documentation of the Biosecure Software Library*, Technical Report, BioSecure.

Pons, A.P. and Polak, P. (2008) 'Understanding user perspectives on biometric technology', *Communications of the Association for Computing Machinery (ACM)*, Vol. 51, No. 9, pp.115–118.

Quinlan, J.R. (1986) 'Induction of decision trees', *Machine Learning*, Kluwer Academic Publishers, Vol. 1, pp.81–106.

Quinlan, J.R. (1993) *C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning)*, Vol. 16, Morgan Kaufmann, 302 pages.

Rakotomalala, R. (1997) *Graphes d'induction*, PhD Thesis, Université Claude Bernard – Lyon 1.

Ratha, N.K., Connell, J.H. and Bolle, R.M. (2001) 'An analysis of minutiae matching strength', *Audio- and Video-based Biometric Person Authentication*, pp.223–228.

Smith, S. (2003) 'Humans in the loop: human computer interaction and security', *IEEE Security and Privacy*, Vol. 1, No. 3, pp.75–79.

Tabassi, E. and Wilson, C.L. (2005) 'A novel approach to fingerprint image quality', *International Conference on Image Processing (ICIP)*, pp. 37–40.

Theofanos, M., Stanton, B., Orandi, S., Micheals, R. and Zhang, N.F. (2007) *Usability Testing of Ten-print Fingerprint Capture*, Technical Report, National Institute of Standards and Technology (NIST).

Theofanos, M., Stanton, B., Sheppard, C., Micheals, R., Zhang, N., Wydler, J., Nadel, L. and Rubin, W. (2008a) *Usability Testing of Height and Angles of Ten-print Fingerprint Capture*, Technical Report, National Institute of Standards and Technology (NIST).

Theofanos, M., Stanton, B. and Wolfson, C.A. (2008b) *Usability & Biometrics: Ensuring Successful Biometric Systems*, National Institute of Standards and Technology (NIST).

Uzoka, F-M.E. and Ndzinge, T. (2009) 'An investigation of factors affecting biometric technology adoption in a developing country context', *International Journal of Biometrics (IJBM)*, Vol. 1, No. 3, pp.307–328.

Yampolskiy, R.V. and Govindaraju, V. (2008) 'Behavioural biometrics: a survey and classification', *International Journal of Biometrics (IJBM)*, Vol. 1, No. 1, pp.81–113.

Appendix

Survey questionnaire

Part A. Socio-demographic characteristics	
Date of birthday	...
Gender	☐ male ☐ female
In which continent do you live?	☐ asia ☐ europe ☐ north America ☐ south America ☐ other
Highest education level	☐ high school graduate ☐ college graduate ☐ other
Profession	☐ student ☐ worker ☐ retired ☐ other

Part B. General perception of biometric systems	
Q_1. Have you ever heard before about biometric authentication systems **(before our study)**?	☐ yes ☐ no
Q_2. Have you ever tried a biometric system **(before our study)**?	☐ yes ☐ no
Q_3. Have you ever been personally the victim of identity fraud?	☐ yes ☐ no
Q_4. How would you rate your knowledge about biometric technology?	☐ not at all important ☐ not important ☐ almost important ☐ quite important ☐ I do not know
Q_5. How would you rate your awareness about fraud identity?	☐ not at all important ☐ not important ☐ almost important ☐ quite important ☐ I do not know
Q_6. In your opinion, are secret-based solutions (eg. password) an appropriate solution against fraud (eg. e-commerce)?	☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree ☐ I do not know
Q_7. In your opinion, are biometric-based solutions an appropriate solution against fraud (eg. e-commerce)?	☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree ☐ I do not know

Part C. Perception of the tested system	
Q_8. Have you ever tried this biometric modality **(before our study)**?	☐ yes ☐ no
Q_9. were you disturbed while using this system?	☐ not at all disturbed ☐ not disturbed ☐ disturbed ☐ quite disturbed ☐ I do not know
Q_{10}. does this technology threats your privacy?	☐ not at all intrusive ☐ not intrusive ☐ intrusive ☐ quite intrusive ☐ I do not know
Q_{11}. is it easy to use this system?	☐ not at all easy ☐ not easy ☐ easy ☐ quite easy ☐ I do not know
Q_{12}. Do you find the verification fast?	☐ not at all fast ☐ not fast ☐ fast ☐ quite fast ☐ I do not know
Q_{13}. Is the answer of the biometric system is correct?	☐ never ☐ rarely ☐ sometimes ☐ always ☐ I do not know
Q_{14}. In your opinion, is the system used can be easily attacked?	☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree ☐ I do not know
Q_{15}. Are you ready to use this biometric system in the future?	☐ strongly disagree ☐ disagree ☐ agree ☐ strongly agree ☐ I do not know
Q_{16}. If you are ready to use this system in the future, would you like to use it for physical or logical access?	☐ physical (*e.g.*, access a building) ☐ logical (*e.g.*, log on to a computer)
Q_{17}. do you trust this system?	☐ no at all ☐ not really ☐ rather ☐ yes ☐ I do not know
Q_{18}. What is your general appreciation of this system?	☐ not at all satisfied ☐ not satisfied ☐ satisfied ☐ quite satisfied ☐ I do not know