Évaluation de système biométrique

Mohamad El-Abed

UNIVERSITÉ de CAEN/BASSE-NORMANDIE
U.F.R. de Sciences
École doctorale S.I.M.E.M

THÈSE

présentée par

M. Mohamad El Abed

et soutenue

le 9 décembre 2011

en vue de l'obtention du

Doctorat de l'Université de Caen Basse-Normandie
Spécialité : Informatique et applications

Arrêté du 7 août 2006

Évaluation de systèmes biométriques

Membre du Jury

Liming Chen	Professeur à l'École Centrale de Lyon	(Rapporteur)
Mohamed Daoudi	Professeur à Télécom Lille 1	(Rapporteur)
Bernadette Dorizzi	Professeur à Télécom & Management Sud Paris	
Jean-Luc Dugelay	Professeur à Eurecom Sophia Antipolis	
Christophe Rosenberger	Professeur à l'ENSICAEN	(Directeur de thèse)
Christophe Charrier	Maître de conférences HDR à l'IUT de Saint-Lô	(Co-encadrant de thèse)

Pour Akram

Remerciements

Tous d'abord, je tiens à remercier le ministère d'Enseignement Supérieur et de la Recherche Français pour son soutien financier qui m'a permis de mener à bien ce travail au cours de ces trois années.

Je remercie également monsieur Liming Chen, professeur à l'École Centrale de Lyon, et monsieur Mohamed Daoudi, professeur à Télécom Lille 1, pour avoir accepté de rapporter ce manuscrit de thèse. Je tiens à remercier madame Bernadette Dorizzi, professeur à Télécom & Management Sud Paris, monsieur Yves Deswarte, directeur de recherche au CNRS, et monsieur Jean-Luc Dugelay, professeur à Eurecom Sophia Antipolis, pour avoir accepté d'examiner ce travail.

Au cours de ma thèse, j'ai travaillé au sein du laboratoire GREYC [1]. Je remercie donc Etienne Grandjean et Mohammed M'Saad, directeurs du GREYC qui m'ont accueilli pendant mes années de thèse. Je remercie également l'école doctorale SIMEM [2]. Je tiens à remercier Christophe Rosenberger, mon directeur de thèse, ainsi que Christophe Charrier qui a co-encadré ma thèse. Je tiens également à remercier Patrice Georget, Cécile Sénémeaud et Patrick Lacharme pour leur aide. Je les remercie pour leur professionnalisme ainsi que l'amitié qu'ils m'ont portée.

Enfin, je remercie tous ceux qui m'ont apporté leur amitié au cours de ces trois années, notamment mes collègues Romain Giot, Alexandre Ninassi et Baptiste Hemery. J'ai également apprécié l'accueil chaleureux de tous les membres du DRI [3] et de l'équipe Monétique & Biométrie du GREYC.

1. Groupe de REcherche en Informatique, Image, Automatique et Instrumentation de Caen
2. Structure, Information, Matière et Matériaux
3. Département des Relations Industrielles

Finalement, je remercie mes proches qui m'ont toujours encouragé et soutenu, notamment Samira, ma mère, ainsi que ma soeur Abir. Je dédie ce travail à Akram, mon père, parti trop tôt.

Résumé

Les systèmes biométriques sont de plus en plus utilisés pour vérifier ou déterminer l'identité d'un individu. Ces systèmes comportent un avantage primordial sur les systèmes d'authentification traditionnels, dans la mesure où la relation entre l'authentifiant et l'individu ne peut pas être plus étroite. Compte tenu des enjeux liés à leur utilisation, notamment pour des applications dans le domaine de commerce électronique ou le contrôle d'accès physique (contrôle aux frontières), il est particulièrement important de disposer d'une méthodologie d'évaluation de tels systèmes. Le problème traité dans cette thèse réside dans la conception d'une méthodologie générique (*i.e.*, indépendante de la modalité) visant à évaluer un système biométrique. Les défis sont nombreux comme la comparaison des systèmes biométriques pour une application particulière, l'évaluation d'un algorithme au cours de son développement ou son paramétrage optimal (seuil de décision). De nombreuses méthodes et métriques ont été proposées dans la littérature afin d'évaluer les systèmes biométriques dans le cadre de compétitions (BioSecure Multimodal Evaluation Campaign, *etc.*) et de plateformes (Fingerprint Verification Competition-onGoing, *etc.*). Cependant, ces travaux sont destinés à évaluer la performance d'un système biométrique sans prendre en considération la manière dont l'utilisateur interagit avec le système, ni la résistance de tels systèmes faces aux attaques. À noter qu'en biométrie, les travaux portant sur l'usage et la sécurité sont beaucoup moins nombreux que ceux liés à la performance.

Les métriques présentées dans la littérature pour évaluer la performance d'un système biométrique sont efficaces et complètes. Dans cette thèse, nous avons choisi d'axer cet aspect d'évaluation sur la qualité de données biométriques acquises. Nous proposons dans un premier temps une méthode d'évaluation de la qualité de données morphologiques reposant sur l'utilisation conjointe de deux types d'informations : la première est basée sur la qualité de l'image et l'autre sur la qualité de l'information contenue dans l'image en utilisant le descripteur Scale Invariant Feature Transform

(SIFT). La méthode proposée possède l'avantage d'être plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé.

Nous avons dans un second temps mis au point deux méthodes d'évaluation portant sur l'usage et la sécurité d'un système biométrique. La première consiste à évaluer l'acceptabilité et la satisfaction des usagers lors de l'utilisation des systèmes biométriques. La méthode proposée utilise un questionnaire et des outils statistiques pour analyser et expliquer leur acceptabilité et satisfaction. La seconde consiste à mesurer la robustesse d'un système biométrique (architecture et algorithmes) contre la fraude. La méthode proposée mesure quantitativement la robustesse d'un système contre la fraude en utilisant une base commune d'attaques et de vulnérabilités des systèmes biométriques, et la notion de facteurs de risque. La méthode proposée est inspirée des Critères Communs (ISO/CEI 15408) et de la méthode d'audit de sécurité EBIOS (Expression des Besoins et Identification des Objectifs de Sécurité). Les deux méthodes proposées sont génériques, ainsi elles peuvent être appliquées indépendamment de la modalité biométrique considérée.

Les différents études, ainsi que la validation de nos mesures de qualité ont été réalisées sur quatre bases de données publiques de visage (*FACES94*, *AR*, *FERET* et *ENSIB*) et une d'empreinte digitale (*FVC2002 DB*$_2$). Deux systèmes de vérification (GREYC-Keystroke et GREYC-Face) développés dans le laboratoire de recherche GREYC et un système commercial basé sur l'empreinte digitale (Fingerprint lock) ont été également utilisés. Ces trois aspects d'évaluation sont en cours d'implémentation dans une plate-forme EVABIO destinée aux industriels et chercheurs en biométrie.

Summary

Biometric systems are increasingly used to verify or identify the identity of an individual. These systems have a significant advantage comparing to traditional authentication systems, since the relationship between the authenticator and the individual can not be closer. Given the challenges related to their use in e-commerce applications or physical access control (border control), it is important to have an evaluation methodology of such systems. The problematic addressed in this thesis is the design of a modality-independent evaluation methodology of biometric systems. The challenges are many such as the comparison of biometric systems for a particular application, the evaluation of an algorithm during its development or in setting its optimal parameters (decision threshold). Many methods and metrics have been proposed in the literature to evaluate biometric systems within competitions (BioSecure Multimodal Evaluation Campaign, *etc.*) and platforms (Fingerprint Verification Competition-onGoing, *etc.*). However, these works focus on evaluating the performance of biometric systems without considering the user interaction way with the system neither the system robustness against attacks. Note that in biometrics, the evaluation works dedicated on the usability and security aspects are less than those dedicated on the performance aspect.

The presented performance metrics in the literature to evaluate biometric systems are efficient. In this thesis, we chose then to focus this evaluation aspect on the quality of biometric raw data. We propose a quality assessment method of morphological data based on the use of two types of informations : the first is based on the quality of the image and the second is a pattern-based quality using the Scale Invariant Feature Transform (SIFT) descriptor. The proposed method has the advantage of being plurimodal (Face, Fingerprint and hand veins), and independent from the used verification system.

We have also proposed two assessment methods dedicated to the usability and security aspects of a biometric system. The first one consists of evaluating the user acceptability and satisfaction when using biometric systems. The proposed method uses a questionnaire and statistical tools to analyze and explain their acceptability and satisfaction. The second consists of measuring the robustness of a biometric system (architecture and algorithms) against fraud. The proposed method quantitatively measures the robustness of the target system by using a database of common threats and vulnerabilities of biometric systems, and the notion of risk factors. The proposed method is inspired from the Common Criteria (ISO/IEC 15408) and the security audit method EBIOS (Expression of Needs and Identification of Security Objectives). The two proposed methods are modality-independent, hence they can be applied regardless of the considered biometric modality.

The experimental studies, as well as the validation of our quality measures were performed on four public face databases (*FACES94*, *ENSIB*, *AR* and *FERET*) and a fingerprint database (*FVC2002* DB_2). Two authentication systems (GREYC-Keystroke and GREYC-Face) developed in the GREYC research laboratory and a commercial fingerprint system (Fingerprint lock) were also used. These three evaluation aspects are being implemented in a platform called EVABIO for companies and researchers in biometrics.

Table des matières

Introduction

« If you can not measure it, you can not improve it. »

Sir William Thomas Kelvin

Positionnement de la problématique

La biométrie suscite une attention accrue depuis les attaques terroristes du 11 septembre 2001. L'usage de la biométrie s'est vite étendue dans de nombreuses applications destinées à gérer l'accès à des ressources physiques (aéroports, casinos, *etc.*) et logiques (ordinateurs, comptes bancaires, *etc.*). Traditionnellement, il existe deux manières d'authentifier un individu. La première méthode est basée sur une *connaissance* (code PIN, mot de passe, *etc.*), tandis que la seconde est basée sur une *possession* (badge, carte à puce, *etc.*). Ces deux méthodes d'authentification peuvent être utilisées de manière complémentaire afin d'obtenir une sécurité accrue. Cependant, chacune d'elle souffre de faiblesses qui peuvent dégrader considérablement leur utilité. En effet, les mots de passes peuvent être oubliés ou bien devinés par une autre entité (comme les attaques par dictionnaire [1]), et les badges peuvent être perdus voire volés. D'après *NTA Monitor Password*[4], une étude en 2002 sur 500 internautes a montré qu'ils avaient en moyenne 21 mots de passe dans leur vie courante : 81% parmi eux utilisent, dans la plupart du temps, des mots de passe communs et 30% écrivent leurs mots de passe dans un fichier. Cette étude montre que le nombre de mots de passe utilisés par les usagers est important, et que la probabilité de perdre un mot de passe est non négligeable.

L'authentification biométrique est une solution émergente permettant de pallier ce problème. Elle comporte un avantage primordial sur les solutions d'authentification

4. `http://www.nta-monitor.com/`

traditionnelles compte tenu de la relation forte entre l'authentifiant et l'utilisateur. Les systèmes biométriques se proposent de comparer deux (*vérification* «1 : 1») ou plusieurs (*identification* «1 : n») échantillons biométriques. Bien que les méthodes d'authentification biométrique promettent d'être très performantes, on ne peut pas garantir actuellement leur robustesse en pratique dans un contexte d'utilisation spécifique et une cible utilisateurs. De plus, il existe plusieurs facteurs affectant la fonctionnalité de ces systèmes [2] tels que :

- Manque de stabilité : en comparaison aux systèmes d'authentification basés sur une *connaissance* ou une *possession*, qui offrent une réponse binaire (oui ou non), les systèmes de vérification biométrique sont moins précis et donnent des réponses en terme de pourcentage de similarité (entre 0% et 100%, le 100% n'étant quasiment jamais atteint). Cette variation des résultats d'authentification d'un individu peut être due à une mauvaise interaction de l'utilisateur avec le capteur biométrique (cas d'un doigt mal positionné sur un capteur d'empreintes digitales), conditions d'acquisition différentes (cas de changements d'éclairage pour un système de reconnaissance faciale) ou utilisation de capteurs différents lors de la phase d'enrôlement et de reconnaissance. En raison de cette variation, la plupart des systèmes biométriques sont vulnérables. Par conséquent, des algorithmes efficaces sont requis pour prendre en compte les artefacts d'acquisition. Ce manque de stabilité peut ainsi augmenter le taux de faux rejets (FRR) d'un système biométrique ;

- Manque de précision : les données biométriques extraites d'individus différents peuvent être relativement similaires (comme le cas de vrais jumeaux, liens de parenté, *etc.*). Ce manque d'unicité peut ainsi augmenter le taux de fausse acceptation (FAR) de certaines modalités biométriques (comme le visage) ;

- Intrusivité et protection de la vie privée : le recours à la biométrie présente des risques en termes d'usage et de respect des droits et des libertés fondamentales. Le fait de capturer et de conserver des données biométriques brutes peut constituer une intrusion de la vie privée. Ces données sont sensibles. De plus, l'acquisition des données biométriques nécessite l'interaction d'un utilisateur avec un capteur biométrique. Cette interaction est effectuée avec ou sans contact (comme le cas des empreintes digitales et les veines de la main, respectivement). Ainsi, certaines modalités biométriques sont considérées plus intrusives que d'autres. Les japonais, par exemple, évitent d'utiliser les systèmes biométriques

avec contact, alors que les projets français de contrôle biométrique continuent de privilégier l'empreinte digitale. L'absence d'études qui prennent en compte l'acceptabilité des usagers lors de l'utilisation des systèmes biométriques peut ainsi limiter la prolifération de tels systèmes ;

– Sensibilité aux attaques : bien que pour certaines modalités, il semble difficile de voler les données biométriques d'une personne, il est toujours possible de contourner un système biométrique. Les travaux présentés par Ratha *et al.* [3] et Capppelli *et al.* [4] montrent la vulnérabilité des systèmes biométriques. Un exemple d'attaque sur ces systèmes est la présentation de fausses empreintes digitales en résine pour usurper l'identité d'un utilisateur légitime.

À cause de ces limitations, il est ainsi particulièrement important de disposer de méthodes d'évaluation pertinentes de ces systèmes. Les méthodes d'évaluation sont également utiles pour d'autres raisons. Premièrement, pour être utilisé dans le domaine industriel, la qualité d'un système biométrique doit être précisément quantifiée. Deuxièmement, le contexte d'utilisation, l'efficacité de la vérification et la robustesse de l'algorithme doivent être définis afin de déterminer s'il répond aux exigences d'une application opérationnelle donnée (accès physique, commerce électronique, *etc.*). Troisièmement, la comparaison des différents systèmes biométriques est indispensable pour qualifier leurs avantages et inconvénients. Enfin, l'évaluation est également nécessaire afin de faciliter la recherche dans ce domaine. Nous avons besoin d'une méthode d'évaluation fiable et efficace afin de mettre en évidence l'intérêt d'un nouveau système biométrique.

Objectifs

En biométrie, plusieurs compétitions et plateformes ont été créées pour évaluer et comparer les systèmes biométriques. Cependant, il n'existe pas de méthodologie générique d'évaluation d'un système biométrique intégrant les différents aspects opérationnels. L'objectif de cette thèse concerne la conception d'une méthodologie générique (*i.e.*, indépendante de la modalité) visant à évaluer un système biométrique (capteur et algorithmes). La méthodologie porte sur trois aspects d'évaluation :

1. Performance

La performance mesure l'efficacité et la fiabilité d'un système biométrique dans un contexte d'utilisation donné. Elle est quantifiée en terme de mesures statistiques (taux d'erreurs, temps de traitement, *etc.*). Les mesures proposées par l'Organisation Internationale de Normalisation ISO/IEC 19795-1 [5] pour

évaluer et comparer la performance des systèmes biométriques sont efficaces et complètes. Ainsi, nous avons choisi d'axer cet aspect d'évaluation sur la qualité de données biométriques acquises. Une telle évaluation est importante puisqu'elle impacte directement sur la performance d'un système biométrique [6]. En se basant sur la notion de qualité, les données biométriques de mauvaise qualité peuvent ainsi être supprimées lors de la phase d'enrôlement ou rejetées au cours de la phase de vérification. Cette information pourra être également utilisée dans les approches multimodales [7]. La qualité pourra être également utilisée pour évaluer les capteurs biométriques. Un algorithme de reconnaissance efficace avec un capteur biométrique de mauvaise qualité, n'est pas considéré comme intéressant ;

2. Usage

Cet aspect d'évaluation consiste à analyser la perception des utilisateurs vis-à-vis de l'utilisation du système pour quantifier leur acceptabilité et satisfaction. Les travaux présentés par Jain *et al.* [8], Kukula et Proctor [9] et Kukula *et al.* [10] montrent l'intérêt de cet aspect d'évaluation lors de la conception et la comparaison de systèmes biométriques. L'absence d'études qui prennent en compte l'acceptabilité des usagers lors de la conception des systèmes biométriques, entraînera une dégradation de la performance de tels systèmes [9]. L'acceptabilité est également un facteur important pour qu'un système soit utilisable. Un système efficace en terme de performance mais pas acceptable, n'est pas considéré comme intéressant (comme le cas des systèmes de vérification par ADN pour du contrôle d'accès physique) ;

3. Sécurité

Le dernier aspect d'évaluation consiste à mesurer la résistance d'un système face à divers types d'attaques. Les travaux présentés par Ratha *et al.* [3] et Maltoni *et al.* [11], montrent la vulnérabilité des systèmes biométriques qui peuvent dégrader considérablement leur fonctionnalité. Un exemple d'une attaque connue sur les systèmes biométriques est la présentation d'une fausse donnée biométrique (doigt prothétique) au capteur. L'évaluation de ces systèmes en terme de sécurité est donc considérée comme un enjeu majeur en biométrie, ainsi que leur impact sur la vie privée des utilisateurs.

Organisation du manuscrit

Ce manuscrit de thèse est composé de six parties :

- **Le chapitre 1** présente une introduction générale sur la technologie biométrique, les modalités biométriques les plus couramment utilisées, et les enjeux du domaine sont énoncés.

- **Le chapitre 2** aborde les différentes méthodologies d'évaluation existantes dans l'état de l'art, qu'elles soient basées sur des critères statistiques (performance), la qualité des données biométriques acquises (qualité), la robustesse contre la fraude (sécurité) ou qu'elles prennent en considération la perception des utilisateurs vis-à-vis de ces systèmes (usage).

- **Les chapitres 3, 4 et 5** présentent les méthodes d'évaluation que nous avons proposées en termes de qualité des données biométriques, de l'usage lors de l'utilisation des systèmes biométriques et leur sécurité, respectivement.

- Enfin, une conclusion et des perspectives de ces travaux viennent clore ce manuscrit de thèse.

Chapitre 1

Les systèmes biométriques

Ce chapitre présente la technologie biométrique et la problématique traitée dans cette thèse. Dans un premier temps, nous présentons la définition et l'utilisation de la biométrie. Par la suite, nous abordons le fonctionnement ainsi que l'architecture générale d'un système biométrique. Enfin, nous présentons les limitations de ces systèmes ainsi que les enjeux liés à leur évaluation.

Sommaire

1.1 Définitions et usage

1.1.1 La biométrie

Il existe trois façons génériques pour vérifier ou déterminer l'identité d'un individu : i) *ce que l'on sait* (code PIN, mot de passe, *etc.*), ii) *ce que l'on possède* (badge, carte à puce, *etc.*) et iii) *ce que l'on est ou ce que l'on sait faire* (empreinte digitale, dynamique de frappe au clavier, *etc.*). Ce dernier point fait référence à la biométrie. La biométrie consiste à vérifier ou déterminer l'identité d'un individu à partir de ses

caractéristiques biologiques (comme l'ADN), comportementales (comme la voix) ou morphologiques (comme l'empreinte digitale).

1.1.2 Les caractéristiques biométriques

Les caractéristiques biométriques par lesquelles il est possible de vérifier l'identité d'un individu sont appelées modalités biométriques. La figure 1.1 illustre un exemple de quelques modalités biométriques. Ces modalités sont basées sur l'analyse des données liées à l'individu et sont généralement classées en trois catégories : biométrie biologique, comportementale et morphologique. La biométrie biologique se base sur l'analyse des données biologiques liées à l'individu (salive, ADN, *etc.*). La biométrie comportementale se base sur l'analyse de comportements d'un individu (manière de marcher, dynamique de frappe au clavier, *etc.*). La biométrie morphologique se base sur les traits physiques particuliers qui, pour toutes personnes, sont permanents et uniques (empreinte digitale, visage, *etc.*). Pratiquement, n'importe quelle caractéristique morphologique ou comportementale peut être considérée comme une caractéristique biométrique, dans la mesure où elle satisfait les propriétés suivantes [12] :

- Universalité : toutes les personnes à identifier doivent la posséder ;
- Unicité : l'information doit être aussi dissimilaire que possible entre les différentes personnes ;
- Permanence : l'information collectée doit être présente pendant toute la vie d'un individu ;
- Collectabilité : l'information doit être collectable et mesurable afin d'être utilisée pour les comparaisons ;
- Acceptabilité : le système doit respecter certains critères (facilité d'acquisition, rapidité, *etc.*) afin d'être employé.

Les caractéristiques biométriques ne possèdent pas toutes ces propriétés, ou les possèdent mais à des degrés différents. Le tableau 1.1, extrait de [13], compare les principales modalités biométriques selon les propriétés suivantes : universalité, unicité, permanence, collectabilité, acceptabilité et performance. Ce tableau montre qu'aucune caractéristique n'est donc idéale et qu'elles peuvent être plus ou moins adaptées à des applications particulières. Par exemple, l'analyse basée sur l'ADN est une des techniques les plus efficaces pour vérifier l'identité d'un individu ou l'identifier [14]. Néanmoins, elle ne peut pas être utilisée pour le contrôle d'accès logique ou physique pour des raisons de temps de calcul, mais aussi, parce que personne ne serait prêt à donner un peu de sang pour faire la vérification. Le choix de la modalité

est ainsi effectué selon un compromis entre la présence ou l'absence de certaines de ces propriétés selon les besoins de chaque application. À noter que le choix de la modalité biométrique peut aussi dépendre de la culture locale des usagers. En Asie, les méthodes nécessitant un contact physique comme les empreintes digitales sont rejetées pour des raisons d'hygiène alors que les méthodes sans contact sont plus répandues et acceptées.

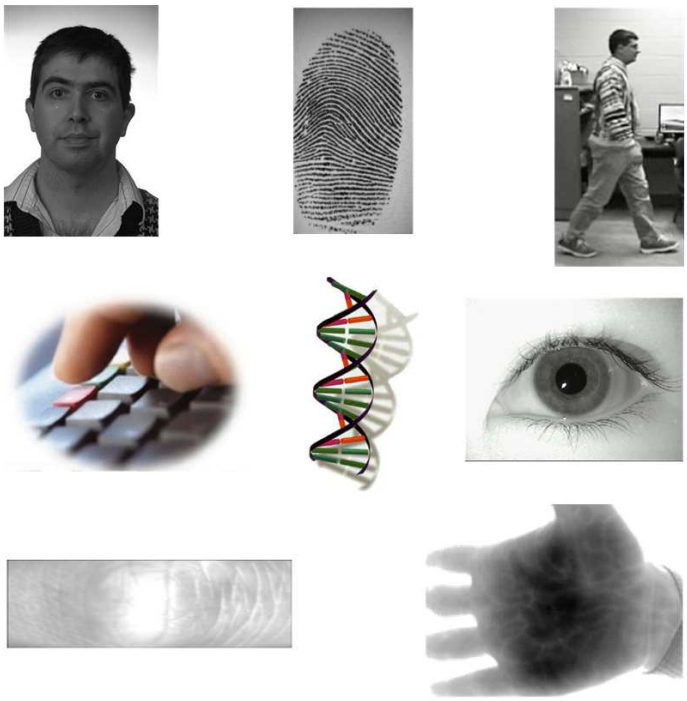


FIG. 1.1 – Quelques modalités biométriques.

Information	**U**	**N**	**P**	**C**	**A**	**E**
ADN	Oui	Oui	Oui	Faible	Faible	*****
Sang	Oui	Non	Oui	Faible	Non	*
Démarche	Oui	Non	Faible	Oui	Oui	***
Dynamique de frappe	Oui	Oui	Faible	Oui	Oui	****
Voix	Oui	Oui	Faible	Oui	Oui	****
Iris	Oui	Oui	Oui	Oui	Faible	*****
Rétine	Oui	Oui	Oui	Oui	Faible	*****
Visage	Oui	Non	Faible	Oui	Oui	****
Géométrie de la main	Oui	Non	Oui	Oui	Oui	****
Oreille	Oui	Oui	Oui	Oui	Oui	*****
Empreinte digitale	Oui	Oui	Oui	Oui	Moyenne	****

TAB. 1.1: Comparaison des modalités biométriques selon les propriétés suivantes : (U) Universalité, (N) Unicité, (P) Permanence, (C) Collectabilité, (A) Acceptabilité et (E) Performance. Pour la performance, le nombre d'étoiles est relié à la valeur du taux d'égale erreur (EER) obtenue dans l'état de l'art (extrait de [13]).

1.1.3 Les modèles biométriques

Un modèle biométrique (appelé aussi gabarit ou template) est l'ensemble des données utilisées pour représenter un utilisateur. Les caractéristiques biométriques acquises ne sont pas enregistrées et utilisées telles quelles. Une phase de traitement est effectuée pour réduire les données biométriques brutes et produire ainsi le modèle biométrique. La figure 1.2 illustre quelques exemples de modèles biométriques. Pour le stockage de ces modèles, il existe quatre emplacements principaux que sont le clé USB, la base centralisée, la machine individuelle de travail et le capteur biométrique. Chacun de ces emplacements présente des avantages et faiblesses en termes de temps de traitement, confidentialité et respect de la vie privée. En France, l'utilisation de la base centralisée est proscrite pour un nombre d'individus élevé par la Commission Nationale Informatique et Libertés (CNIL).

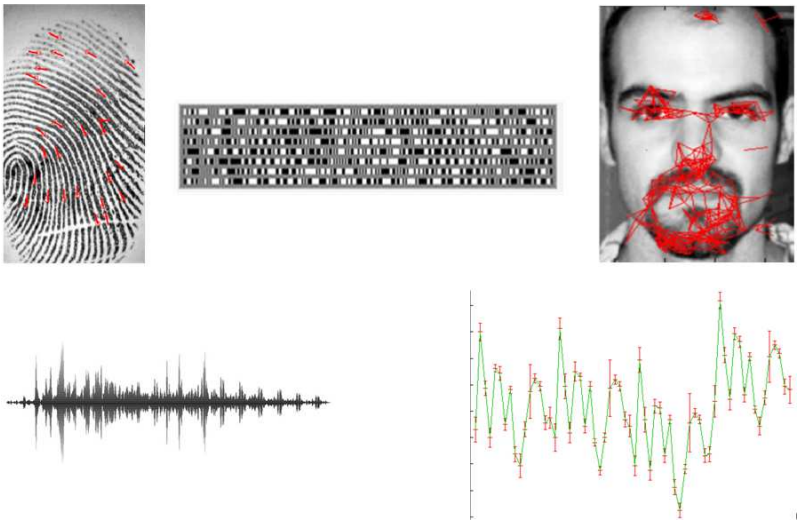


FIG. 1.2 – Quelques exemples de modèles biométriques. De gauche à droite, de haut en bas : minuties extraites d'une empreinte, Iris code, graphe d'un visage utilisant les points d'intérêt, signal vocal et signal de dynamique de frappe au clavier.

1.1.4 Utilisation de la biométrie

Le champ d'application de la biométrie est très vaste. En effet, tous les domaines qui nécessitent de vérifier ou déterminer l'identité de personnes sont concernés. On retrouve ainsi des applications de la biométrie pour gérer l'accès à des ressources physiques (comme l'accès à des lieux sécurisés) et logiques (comme le commerce électronique). La biométrie intéresse aussi plusieurs pays (l'Europe, les États-Unis, *etc.*) afin de produire des titres d'identité plus sûrs, telle que la carte nationale d'identité ou le passeport biométrique. À noter qu'en France, le passeport biométrique est

désormais déployé. Il intègre une puce *RFID* qui contient au moins deux informations biométriques : une empreinte digitale et une image du visage numérisée. Enfin, la biométrie n'a pas que des applications à vocation sécuritaire, mais également des applications qui facilitent le quotidien des usagers. Ainsi, la biométrie est utilisée dans certains aéroports permettant aux clients réguliers de ne pas perdre de temps lors de l'embarquement. La figure 1.3, réalisée d'après les chiffres de *International Biometric Group* [15], montre les parts de marché des principales méthodes biométriques en 2009. Les empreintes digitales sont toujours les plus utilisées, suivies par la reconnaissance faciale. Ces deux modalités représentent les trois quarts du marché de la biométrie.

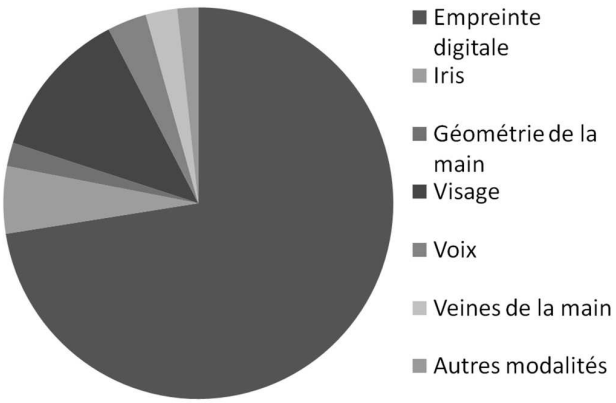


FIG. 1.3 – Parts de marché des techniques biométriques en 2009 (extrait de [15]).

1.1.5 La biométrie et les méthodes d'authentification traditionnelles

Puisque la biométrie fait appel à *ce que l'on est*, elle comporte un avantage primordial sur les méthodes traditionnelles, dans le sens où elle évite l'usage d'un grand nombre de mots de passe complexes, de badges, *etc.* Le tableau 1.2 présente un parallèle entre la biométrie et les méthodes d'authentification traditionnelles. Ce tableau montre que les systèmes biométriques facilitent le processus d'authentification et résiste aux différentes attaques existantes sur les systèmes basés sur *un secret* ou *une possession*. Cependant, ces systèmes présentent plusieurs inconvénients concernant le respect de la vie privée et l'incertitude de l'information biométrique. Une comparaison de ces techniques est détaillée par O'Gorman [16].

Authentification biométrique	**Authentification par mot de passe/clé**
- basée sur des mesures morphologiques, comportementales ou biologiques	- basée sur que l'*on sait* ou *possède*
- utilisation facile (pas de secret à retenir)	- pouvant être plus compliquée (mots de passe complexe)
- authentifie l'individu	- authentifie la clé
- l'information est en relation étroite à l'utilisateur de façon permanente	- il peut être perdu, volé ou oublié
- utilise une comparaison probabiliste	- utilise une comparaison exacte
- l'information biométrique peut être modifiée et/ou altérée avec le temps, il est incertain	- l'information ne varie pas, elle est sûre
- problème de respect de la vie privée	- moindre impact sur la vie privée
- difficile de révoquer l'information	- changement aisé

TAB. 1.2: Comparaison entre l'authentification biométrique et par mot de passe/clé.

1.2 Technologie biométrique

1.2.1 Enrôlement, vérification et identification

Les systèmes biométriques fonctionnent selon trois modes que sont l'enrôlement, la vérification d'identité et l'identification :

– **Enrôlement**
L'enrôlement est la première phase de tout système biométrique. Il s'agit de l'étape pendant laquelle un utilisateur est enregistré dans le système pour la première fois. Elle est commune à la vérification et l'identification. Pendant l'enrôlement, la caractéristique biométrique est mesurée en utilisant un capteur biométrique afin d'extraire une représentation numérique. Cette représentation est ensuite réduite, en utilisant un algorithme d'extraction bien défini, afin de réduire la quantité de données à stocker pour ainsi faciliter la vérification et l'identification. Dépendant de l'application et du niveau de sécurité souhaité, le modèle biométrique retenu, est stocké soit dans une base de données centrale soit sur un élément personnel propre à chaque personne ;

– **Vérification**
La vérification d'identité consiste à contrôler si l'individu utilisant le système est bien la personne qu'il prétend être. Le système compare l'information biométrique acquise avec le modèle biométrique correspondant stocké dans la base de données, on parle de test *1 : 1*. Dans ce cas, le système renvoie uniquement une décision binaire (oui ou non) pouvant être pondérée. Le processus de vérification peut être formalisé comme suit :

Soit le vecteur d'entrée C_U définissant les caractéristiques biométriques de l'utilisateur U extraites par le système, et M_U son modèle biométrique stocké dans la base de données, le système retourne une valeur booléenne suite au calcul de la fonction f définie par :

$$f(C_U, M_U) = \begin{cases} 1 & \text{si } S(C_U, M_U) \geq \tau \\ 0 & \text{sinon} \end{cases} \tag{1.1}$$

où S est la fonction de similarité définissant la correspondance entre les deux vecteurs biométriques, et τ le seuil de décision à partir duquel les deux vecteurs sont considérés comme identiques ;

– **Identification**
En mode identification, le système biométrique détermine l'identité d'un individu inconnu à partir d'une base de données d'identités, on parle de test *1 : N*. Dans ce cas, le système peut alors soit attribuer à l'individu inconnu l'identité correspondant au profil le plus proche retrouvé dans la base (ou une liste des profils proches), soit rejeter l'individu. Le processus d'identification peut être formalisé ainsi :
Soit le vecteur d'entrée C_U définissant les caractéristiques biométriques extraites par le système lorsqu'un utilisateur U se présente devant celui-ci, l'identification revient à déterminer l'identité de I_t, $t \in \{0, 1, \cdots, N\}$ où $I_1, \cdots, I_N$ sont les identités des utilisateurs préalablement enrôlés dans le système, et I_0 indique une identité inconnue. La fonction d'identification f peut ainsi être définie par :

$$f(C_U) = \begin{cases} I_k & \text{si } \max_{1 \leq k \leq N} S(C_U, M_k) \geq \tau \\ I_0 & \text{sinon} \end{cases} \tag{1.2}$$

où M_k est le modèle biométrique correspondant à l'identité I_k, S est la fonction de similarité, et τ le seuil de décision.

1.2.2 Architecture d'un système biométrique

L'architecture d'un système biométrique contient cinq modules comme le montre la figure 1.4 :

– Le **module de capture** qui consiste à acquérir les données biométriques afin d'extraire une représentation numérique. Cette représentation est ensuite utilisée pour l'enrôlement, la vérification ou l'identification. Il s'agit d'un capteur biométrique qui peut être de type sans ou avec contact ;

– Le **module de traitement du signal** qui permet de réduire la représentation numérique extraite afin d'optimiser la quantité de données à stocker lors de la phase d'enrôlement, ou pour faciliter le temps de traitement pendant la phase de vérification et l'identification. Ce module peut avoir un test de qualité pour contrôler les données biométriques acquises ;

– Le **module du stockage** qui contient les modèles biométriques des utilisateurs enrôlés du système ;

– Le **module de similarité** qui compare les données biométriques extraites par le module d'extraction de caractéristiques à un ou plusieurs modèles préalablement enregistrés. Ce module détermine ainsi le degré de similarité (ou de divergence) entre deux vecteurs biométriques ;

– Le **module de décision** qui détermine si l'indice de similarité retourné est suffisant pour déterminer l'identité d'un individu.

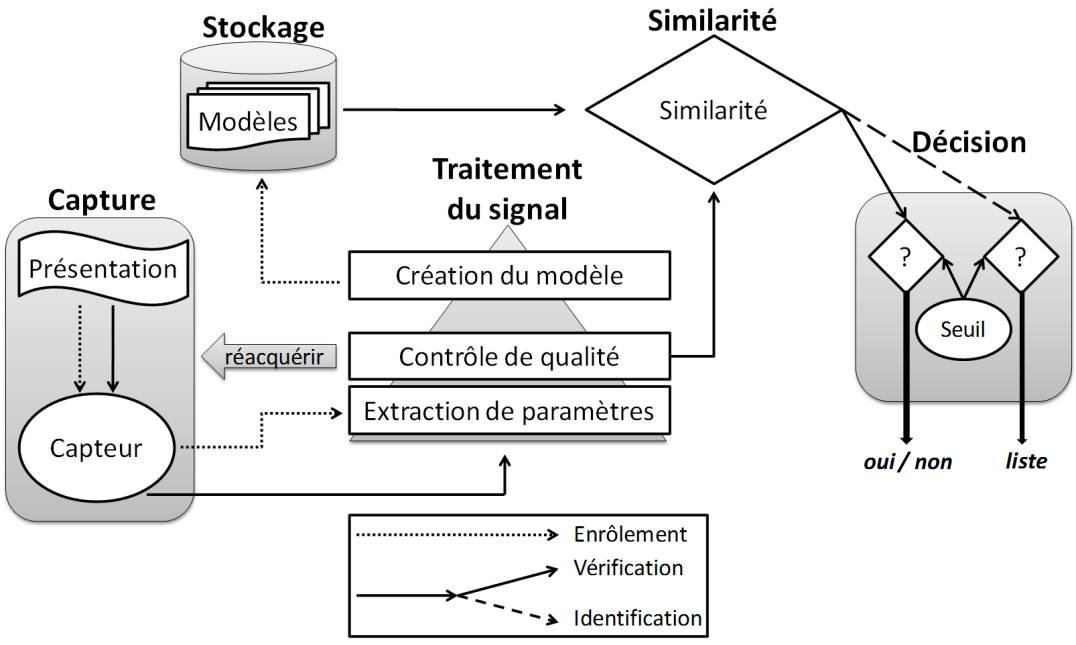


FIG. 1.4 – Architecture générique d'un système biométrique (extrait de l'Organisation Internationale de Normalisation ISO/IEC 19795-1 [5]).

1.3 Les modalités biométriques

Cette section aborde quelques exemples de différentes modalités biométriques qui sont basées soit sur l'analyse biologique, comportementale ou morphologique. Nous présentons notamment les techniques biométriques basées sur la reconnaissance de visage, d'empreinte digitale et de dynamique de frappe au clavier. Nous avons choisi de détailler un peu plus ces modalités biométriques pour plusieurs raisons : tout d'abord, la reconnaissance de visage et d'empreinte digitale sont parmi les méthodes biométriques les plus utilisées [17, 15]. D'autre part, ces modalités seront étudiées et comparées dans le cadre de cette thèse pour illustration.

1.3.1 Biologie

Cette catégorie s'appuie sur l'analyse de caractéristiques biologiques de l'individu. La prémisse à ce type d'analyse est que la donnée biologique de chaque individu constitue une signature personnelle. L'analyse biologique comprend : l'odeur [18], l'ADN [19], et les signaux physiologiques [20]. Cette modalité n'est pas beaucoup utilisée pour du contrôle d'accès logique et physique. Nous ne détaillons pas plus ce type de biométrie.

1.3.2 Comportement

Cette catégorie se base sur l'analyse de comportement d'un individu comme la dynamique de signature [21, 22], sa démarche [23, 24], sa façon de taper au clavier [25, 26, 27] et sa voix [28]. Dans la suite, nous présentons uniquement un exemple de la modalité basée sur la dynamique de frappe au clavier. Le système biométrique présenté est étudié au laboratoire GREYC dans le cadre de cette thèse.

1.3.2.1 GREYC-Keystroke

C'est un système d'authentification biométrique basée sur la dynamique de frappe au clavier développé dans le laboratoire de recherche GREYC dans le cadre de la thèse de Romain Giot [29]. Une illustration de ce logiciel est donnée à la figure 1.5. Les principaux objectifs de ce logiciel sont de permettre la création d'une base de données afin de comparer les différents algorithmes de l'état de l'art sous les mêmes conditions d'acquisition. Le logiciel est librement téléchargeable sur `http://www.epaymentbiometrics.ensicaen.fr/index.php/research-activities/resources/65`. Les principales caractéristiques du système sont :

– Le système utilise plusieurs temps pour un même couple de touches successives comme le montre la figure 1.6 :

  - P-P (Pression - Pression) : temps entre deux pressions de touches $(T_2 - T_1)$;
  - P-R (Pression - Relâchement) : temps entre l'appui sur la touche et le moment où elle est relâchée $(T_3 - T_1$ et $T_4 - T_2)$;
  - R-P (Relâchement - Pression) : temps entre le relâchement d'une touche et l'appui sur la suivante $(T_3 - T_2)$;
  - R-R (Relâchement - Relâchement) : temps entre le relâchement de deux touches successives $(T_4 - T_3)$.
  - la vitesse de frappe de l'intégralité du mot de passe ;
  - la fréquence des fautes de frappes.

– Le système implémente une méthode présentée dans l'article [30]. Il s'agit d'une méthode statistique qui permet de calculer un score dépendant du vecteur moyen et de l'écart type. Pour un vecteur de taille n, avec v_i la i$^{\text{ème}}$ donnée et μ_i et σ_i respectivement la valeur moyenne et son écart type, le score est donné par :

$$score \;=\; 1 \;-\; \frac{1}{n}\sum_{i=1}^{n} exp\left(-\frac{|v_i - \mu_i|}{\sigma_i}\right) \tag{1.3}$$

– Le système possède un taux d'égale erreur (EER) égal à $17,51\%$ sur une base composée de 70 individus, avec 3 vecteurs d'enrôlement et 2 pour le test ;

– L'architecture du système n'est pas distribuée (*i.e.*, tous les composants du système, y compris la base des modèles biométriques sont implémentés dans le même PC) ;

– Il n'y a aucun mécanisme de protection des données ni de chiffrement sur la base des modèles biométriques ;

– Aucun test de qualité n'est utilisé pour contrôler la qualité des données biométriques acquises pendant la phase d'enrôlement ;

– Le PC utilisé est connecté sur Internet.

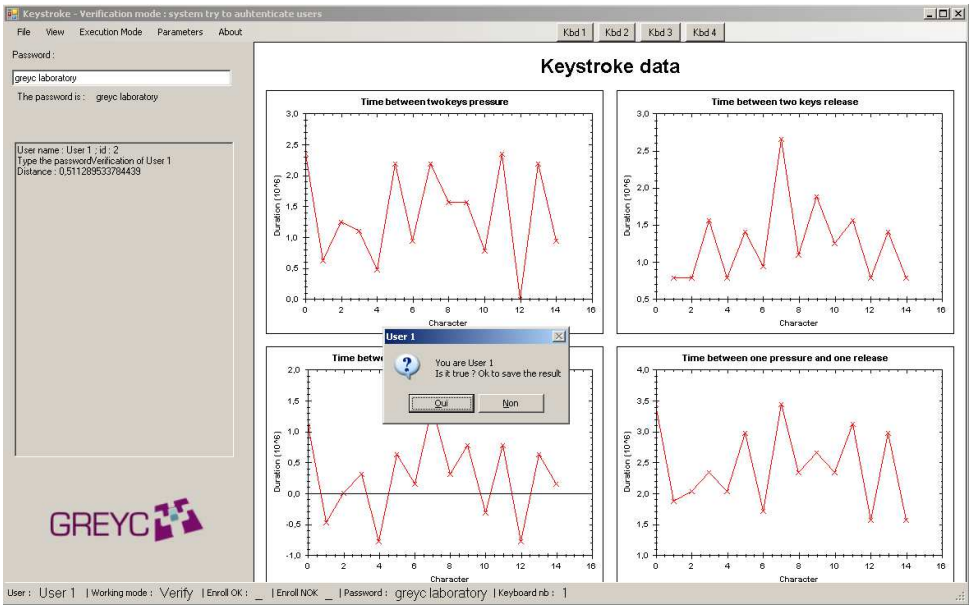


FIG. 1.5 – Le logiciel GREYC-Keystroke. Exemple de vérification résultant d'une tentative d'un utilisateur légitime.

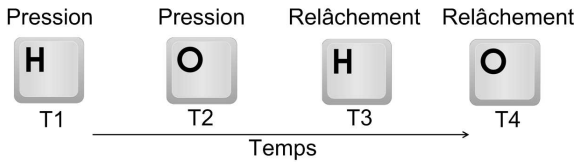


FIG. 1.6 – Temps extrait au cours de la frappe de «HO».

1.3.3 Morphologie

Cette catégorie s'appuie sur l'utilisation de traits physiques qui sont uniques et permanents chez l'individu. Plusieurs modalités ont été employées pour extraire cette information comme le visage 2D [31, 32] et 3D [33, 34], l'empreinte digitale [35, 36], la géométrie de la main [37, 38], l'iris [39, 40], *etc.* Dans la suite, nous présentons uniquement un exemple des deux modalités basées sur le visage et l'empreinte digitale. Les deux systèmes présentés sont étudiés dans le cadre de cette thèse.

1.3.3.1 GREYC-Face

C'est un système de reconnaissance faciale développé dans le laboratoire de recherche GREYC [41]. Les principales caractéristiques du système sont :

– Le système est basé sur le descripteur Scale Invariant Feature Transform (SIFT) proposé par Lowe [42]. L'image I est ainsi caractérisée par l'ensemble

$Y(I) = \{k_i = (x_i, y_i, \sigma_i, \theta_i, v_i) \mid i = 1 : N(I)\}$ avec : $N(I)$ le nombre de points d'intérêt détectés dans I ; (x_i, y_i) la position du point d'intérêt i dans I ; (σ_i, θ_i) l'échelle et l'orientation du point d'intérêt i ; et v_i le vecteur de descripteurs du point d'intérêt i. La vérification entre deux images I_1 et I_2 correspond au calcul du nombre d'associations entre les deux ensembles $Y(I_1)$ et $Y(I_2)$. Une association est définie par une double mise en correspondance entre deux points d'intérêt. La méthode de mise en correspondance utilisée est celle présentée par Ladoux *et al.* [43] (version modifiée de la méthode proposée par Lowe [42]). Pour le point d'intérêt x de l'image I_1, nous recherchons le point d'intérêt y de I_2 le plus proche parmi l'ensemble de points d'intérêt de I_2. Nous regardons également si le second point d'intérêt $y\prime$ le plus proche est suffisamment loin de x au moyen d'une valeur seuil C :

$$d(x, y) = min_{\{z \in Y(I_2)\}} d(x, z) \tag{1.4}$$

et

$$d(x, y) \leq C\ d(x, y\prime) \tag{1.5}$$

avec

$$d(x, y\prime) = min_{\{z \in Y(I_2),\ d(x,z) > d(x,y)\}} d(x, z) \tag{1.6}$$

La distance $d(.,.)$ est une distance euclidienne calculée entre les deux descripteurs normalisés correspondant aux points d'intérêt. Si ces deux conditions ne sont pas remplies, alors le point x n'est pas mis en correspondance avec le point y. Nous disons ainsi dans la suite qu'il y a une association entre les deux points d'intérêt x et y si est seulement si : le point x est mis en correspondance avec le point y, le point y est mis en correspondance avec le point x. La similarité entre les deux ensembles de point est le nombre de points d'intérêt mis en correspondance. La figure 1.7 illustre un exemple de vérification résultant d'une tentative d'un utilisateur légitime. Le nombre d'associations est utilisé comme mesure de similarité ;

– Le système possède un EER égal à $8,76\%$ sur une base composée de 70 individus, avec 1 image d'enrôlement et 2 pour le test ;

– L'architecture du système n'est pas distribuée (*i.e.*, tous les composants du système, y compris la base des modèles biométriques sont implémentés dans le

même PC) ;

– Il n'y a aucun mécanisme de protection des données ni de chiffrement sur la base des modèles biométriques ;

– Aucun test de qualité n'est utilisé pour contrôler la qualité des données biométriques acquises pendant la phase d'enrôlement.

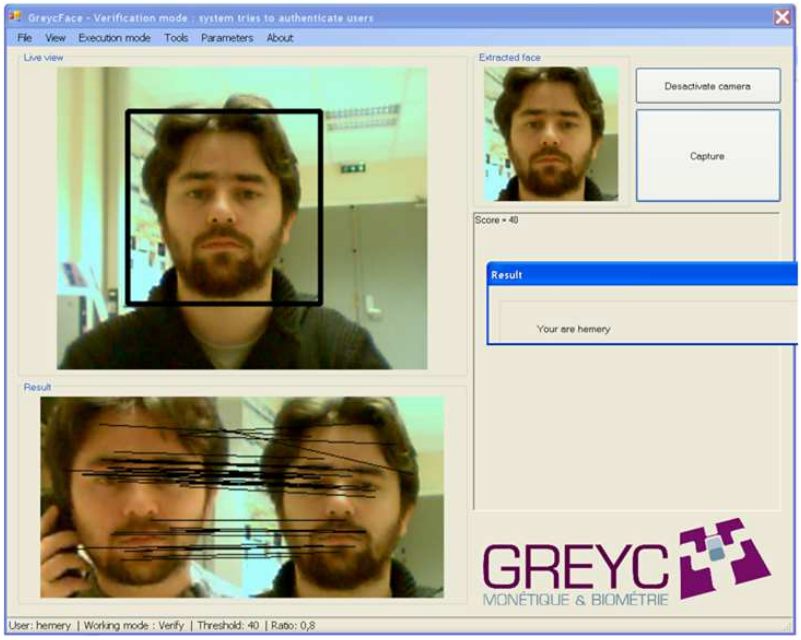


FIG. 1.7 – Le logiciel GREYC-Face. Exemple de vérification résultant d'une tentative d'un utilisateur légitime.

1.3.3.2 Fingerprint lock

C'est un système d'authentification commercial basé sur la reconnaissance d'empreintes digitales. Nous utilisons ce système pour contrôler l'accès à la salle de développement dans notre laboratoire. Une illustration de ce système est donnée à la figure 1.8. Les principales caractéristiques du système sont :

– Le système fournit un taux de fausses acceptations (FAR) égal à 0.0001% et un taux de faux rejets (FRR) égal à 0.1%. Le taux d'erreur moyen (HTER) est ainsi égal à 0.05% ;

– L'architecture du système n'est pas distribuée (*i.e.*, tous les composants du système, y compris la base des modèles biométriques sont implémentés dans le

matériel) ;

– Il n'y a aucun mécanisme de protection des données ni de chiffrement sur la base des modèles biométriques, mais le système est physiquement protégé ;

– Le matériel n'est pas connecté sur Internet, et ne possède pas de port USB ;

– Aucun test de qualité n'est utilisé pour contrôler la qualité des données biométriques acquises pendant la phase d'enrôlement.

FIG. 1.8 – La serrure d'empreinte digitale Fingerprint lock.

1.4 Les standards en biométrie

Les standards en biométrie sont nombreux pour répondre aux exigences en terme du niveau de sécurité demandé par les applications industrielles. L'objectif principal de ces standards est de maintenir l'interopérabilité et l'échange de données entre les systèmes biométriques. Un exemple de l'interopérabilité est la capacité des passeports biométriques délivrés afin qu'ils soient lisibles par les lecteurs placés aux frontières. En biométrie, il existe plusieurs travaux réalisés par le consortium *BioAPI*[1] [44, 45] et les standards *Common Biometric Exchange File Format* (CBEFF) [46], ANSI/NIST-ITL 1 [47], ISO/IEC 19795-1 [5], ISO/IEC FCD 19792 [48], ISO/IEC JRC 1/SC 37 [49, 50, 51, 52, 53]. Dessimoz *et al.* [54] présente un aperçu plus détaillé des standards en biométrie.

La spécification *BioAPI* permet de définir une interface de programmation d'applications biométriques standardisées (*API*). Cette *API* permet aux applications de

1. http ://www.bioapi.org/

communiquer avec un ensemble de technologies biométriques dans un environnement commun. Les principaux objectifs de cette spécification sont divers :

- Définir un framework permettant d'intégrer différents systèmes biométriques (visage, dynamique de frappe, *etc.*) provenant de différents fournisseurs ;
- Offrir aux chercheurs et aux développeurs un outil significatif et fiable pour évaluer leurs systèmes avec ceux qui existent dans l'état de l'art ;
- Faciliter l'implémentation de l'approche multimodale qui, d'une part, améliore la performance de ces systèmes, et d'autre part, tolère les problèmes d'acquisition de l'information biométrique. Dans les faits, la spécification BIOAPI n'est pas très utilisée par les industriels car trop contraignante.

Le standard ISO/IEC 19795-1 [5] porte sur l'évaluation des systèmes biométriques en terme de performance. Le standard propose des mesures statistiques destinées à quantifier et évaluer les systèmes biométriques. Une des mesures proposées est l'EER utilisée pour tester et comparer les systèmes biométriques.

Le standard ISO/IEC FCD 19792 [48] porte sur les aspects de l'évaluation de la sécurité des systèmes biométriques. La norme présente une liste des menaces et des vulnérabilités des systèmes biométriques. La norme souligne également d'autres facteurs principaux à prendre en compte lors de l'évaluation de ces systèmes. La norme souligne également des facteurs liés à la protection de la vie privée (gestion de stockage et accès aux modèles biométriques) qui doivent également être pris en compte lors de l'évaluation.

1.5 Les limitations des systèmes biométriques

Malgré les avantages des systèmes biométriques par rapport aux systèmes traditionnels, leur utilisation est toujours limitée à des applications spécifiques (comme le passeport biométrique). Ces systèmes souffrent de plusieurs limitations qui peuvent dégrader considérablement leur intérêt.

La première limitation se situe dans la performance. Contrairement aux systèmes d'authentification traditionnels, les systèmes d'authentification basés sur la biométrie sont moins précis (*i.e.*, pourcentage de similarité entre 0% et 100%, le 100% n'étant quasiment jamais atteint). Ce manque de précision est dû à plusieurs facteurs : la variabilité lors de la capture (*i.e.*, bruits d'acquisition, utilisation de plusieurs capteurs

d'acquisition, *etc.*), la variabilité intra-classe (variabilité des données biométriques pour un individu) et la similarité inter-classe (*i.e.*, similarité des données biométriques de plusieurs individus).

Une autre limitation de la biométrie est la limitation d'usage ou culturelle. La biométrie et particulièrement les empreintes digitales ont une mauvaise réputation et sont associées à la surveillance des personnes et à l'identification de criminels. Dépendant de la modalité utilisée, l'acquisition de données biométriques est effectuée sans ou avec contact avec le capteur biométrique. Ce contact est une source d'inquiétudes pour certains utilisateurs pour des raisons d'hygiène et d'intrusion physique. Prenons le cas de reconnaissance par la rétine : cette technologie assure une bonne fiabilité et une haute barrière contre la fraude. Malgré l'efficacité de cette technologie, elle est considérée comme intrusive et elle est très peu utilisée dans les milieux de la sécurité privée. Le recours à la biométrie présente également des risques en termes de respect des droits et des libertés fondamentales. En France, la Commission Nationale de l'Informatique et des Libertés (CNIL), n'autorisent les applications qui font de l'usage de biométrie de trace (*ex.*, empreinte digitale) que dans la mesure où le besoin de la sécurité est important. Ces contraintes d'utilisation limitent de plus en plus la prolifération de certaines modalités pour des applications moins sécuritaires (comme le contrôle d'accès à des bâtiments).

Enfin, les systèmes biométriques sont vulnérables à des attaques spécifiques. Ratha *et al.* [3] présente huit emplacements de points de compromission d'un système biométrique. Même s'il est plus difficile de falsifier un iris que de décrypter un mot de passe, il est toutefois possible de reproduire d'autres types de modalités. Les travaux présentés dans [55] montrent la facilité de reproduire des empreintes digitales en utilisant des images résiduelles sur le capteur.

1.6 Évaluation des systèmes biométriques

L'évaluation des systèmes biométriques a pour objectif d'en diminuer les limitations vues dans la section 1.5. L'évaluation de ces systèmes est généralement réalisée selon trois aspects d'évaluation comme le montre la figure 1.9 :

1. **La performance** qui mesure l'efficacité d'un système biométrique en terme d'erreur tel que le taux d'échec à l'acquisition (FTA) [5] ;

2. **L'usage** qui mesure l'acceptabilité et la satisfaction des utilisateurs lors de l'utilisation de systèmes biométriques [56] ;

3. **La sécurité** qui mesure la robustesse d'un système biométrique (capteur et algorithmes) contre la fraude [48].

L'évaluation des systèmes biométriques est un enjeu majeur en biométrie pour plusieurs raisons. Premièrement, elle permet d'offrir aux chercheurs et aux développeurs un outil pour mieux tester et évaluer leurs systèmes avec ceux qui existent dans l'état de l'art. Deuxièmement, elle permet de prendre en considération le comportement des utilisateurs durant le processus d'évaluation, ce qui permet de mieux comprendre leur besoin et mieux déployer cette technologie dans notre vie quotidienne. Enfin, elle permet d'identifier, pour chaque système, les applications industrielles en se basant sur divers critères que sont la performance, l'usage, la sécurité et le coût de déploiement de la technologie.

En biométrie, les travaux liés à la qualité, l'usage et la sécurité sont beaucoup moins nombreux que ceux liés à la performance. L'Organisation Internationale de Normalisation ISO/IEC 19795-1 [5] a proposé plusieurs métriques pour évaluer et comparer la performance des systèmes biométriques. Ainsi, nous avons choisi d'axer cet aspect d'évaluation sur l'évaluation de la qualité de données biométriques. Une telle évaluation est utile pour améliorer la performance globale des systèmes biométriques. Ce type d'information pourra également être exploité dans les approches multimodales [7] ou pour évaluer les capteurs biométriques. Dans le cadre de cette thèse, nous traitons les aspects liés à la qualité, l'usage et la sécurité des systèmes biométriques.

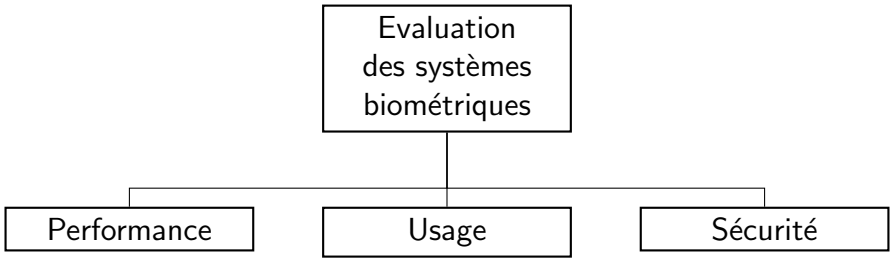


FIG. 1.9 – Aspects d'évaluation des systèmes biométriques.

1.7 Conclusion

Dans ce premier chapitre, nous avons présenté la problématique traitée dans cette thèse : l'évaluation des systèmes biométriques. Après une introduction générale sur la biométrie, ses propriétés et son utilisation, nous avons présenté trois systèmes biométriques (deux développés dans le laboratoire de recherche GREYC et un système commercial) étudiée dans le cadre de cette thèse. Nous avons ensuite présenté l'architecture d'un système biométrique générique, les limitations de ces systèmes ainsi que ses aspects d'évaluation.

Nous présentons dans le chapitre suivant un état de l'art sur les différentes méthodologies d'évaluation des systèmes biométriques.

Chapitre 2

État de l'art de l'évaluation de systèmes biométriques

Ce chapitre présente un état de l'art des méthodes d'évaluation des systèmes biométriques. Nous nous sommes intéressés à l'évaluation réalisée sous les quatre aspects : performance en terme d'erreurs, qualité de données biométriques acquises, usage en terme d'acceptabilité et sécurité en terme de robustesse contre la fraude.

Sommaire

2.1 Introduction

Les technologies biométriques constituent un assemblage de composants matériels (capteurs, ordinateurs, *etc.*) et logiciels (algorithmes, drivers des capteurs, *etc.*). Chacun de ces composants a des faiblesses et des limites. Certaines technologies sont plus fiables que d'autres et certains manufacturiers offrent des produits plus ou moins performants pour une technologie donnée. De plus, ces systèmes nécessitent

l'acquisition des données biométriques de personnes, ceci à l'aide d'un capteur et dans un environnement qui est, dans la plupart de temps, non contrôlé. En conséquence, ces systèmes ne donnent pas une réponse précise sur l'identité d'une personne mais une réponse relative qui s'exprime par un indice de similarité qui n'atteint quasiment jamais 100%. L'objectif de l'état de l'art est de pouvoir analyser les méthodes d'évaluation existantes afin de sélectionner celles qui vont être intégrées dans la méthodologie d'analyse de systèmes biométriques que nous cherchons à concevoir.

Selon Phillips *et al.* [57], l'évaluation des systèmes biométriques est traitée selon trois types d'évaluation : l'évaluation technologique, l'évaluation par scénario et l'évaluation opérationnelle. L'**évaluation technologique** consiste à tester un ou plusieurs algorithmes de la même modalité en utilisant une base de données pré-acquise. Les tests sont effectués de manière *offline* et les résultats obtenus sont reproductibles. L'**évaluation par scénario** consiste à tester tout le système (capteur et algorithmes). Les résultats obtenus ne sont pas forcément reproductibles. L'**évaluation opérationnelle** consiste à tester un système biométrique en condition réelle d'utilisation. Généralement, les résultats obtenus ne sont pas reproductibles.

Dans ce chapitre, nous abordons les méthodologies d'évaluation existantes dans l'état de l'art qui sont généralement divisées en quatre aspects que sont la performance, la qualité, l'usage et la sécurité.

2.2 Performance des systèmes biométriques

La performance mesure l'efficacité et la fiabilité d'un système biométrique dans un contexte d'utilisation donné. Dans cette section, nous commençons par définir les différentes mesures des taux d'erreur utilisées pour quantifier la performance d'un système biométrique. Ensuite, nous présentons les principales bases de données collectées, les compétitions de comparaison de systèmes biométriques et les plateformes d'évaluation existantes.

2.2.1 Métriques

Il existe dans la littérature plusieurs métriques [58, 59, 5, 60] de diverses natures que sont les mesures des taux d'erreur, les mesures liées au temps de traitement et occupation mémoire, les courbes de performance ainsi que les points de fonctionnement associés.

2.2.1.1 Les mesures des taux d'erreur

Selon l'Organisation Internationale de Normalisation ISO/IEC 19795-1 [5], les mesures des taux d'erreur sont divisées en trois classes que sont les taux d'erreur fondamentale, taux d'erreur de systèmes d'authentification et taux d'erreur de systèmes d'identification.

A. Taux d'erreur fondamentale

- Taux d'échec à l'acquisition (*failure-to-acquire rate*, FTA) : proportion des tentatives de vérification ou d'identification pour lesquels le système biométrique n'a pas pu acquérir l'information biométrique requise ;

- Taux d'échec à l'enrôlement (*failure-to-enroll rate*, FTE) : proportion des individus pour lesquels le système n'a pas pu générer le modèle biométrique durant la phase d'enrôlement. Prenons par exemple le cas des empreintes, il existe certaines personnes qui n'ont pas d'empreintes pour des raisons génétiques, ou des empreintes quasi-inexistantes pour des raisons médicales ;

- Taux de fausse non-correspondance (*false non-matche rate*, FNMR) : proportion de fausse non-correspondance, par l'algorithme de comparaison, entre la donnée biométrique acquise et le modèle correspondant ;

- Taux de fausse correspondance (*false matche rate*, FMR) : proportion de fausse correspondance, par l'algorithme de comparaison, entre la donnée biométrique acquise et le modèle correspondant à un autre individu.

B. Taux d'erreur de systèmes d'authentification

- Taux de faux rejets (*false rejection rate*, FRR) : proportion des transactions des utilisateurs légitimes rejetées par erreur. Ces transactions sont rejetées, par l'algorithme de correspondance, en raison de non-correspondance à tort ainsi que ceux rejetées en raison d'un échec à l'acquisition.

  Exemple : pour une transaction de vérification à une seule tentative et un seuil fixé τ, le taux de faux rejets est calculé par :

$$FRR(\tau) = FTA + FNMR(\tau) * (1 - FTA) \tag{2.1}$$

– Taux de fausses acceptations (*false acceptance rate*, FAR) : proportion des transactions des imposteurs acceptées par erreur.

Exemple : pour une transaction de vérification à une seule tentative et un seuil fixé τ, le taux de fausses acceptations est calculé par :

$$FAR(\tau) = FMR(\tau) * (1 - FTA) \tag{2.2}$$

La figure 2.1 représente la distribution théorique des taux de vraisemblance des utilisateurs légitimes et des imposteurs. Les deux taux d'erreurs, FAR et FRR, sont liés et dépendent d'un seuil de décision qui doit être ajusté en fonction de la caractéristique ciblée du système biométrique haute ou basse sécurité. En effet, plus le seuil est bas, plus le taux de fausses acceptions est élevé. Dans ce cas, le système biométrique acceptera des imposteurs. À l'inverse, plus le seuil est élevé, plus le taux de fausses acceptions est bas. Le système biométrique sera alors robuste aux imposteurs mais rejettera de vrais utilisateurs.

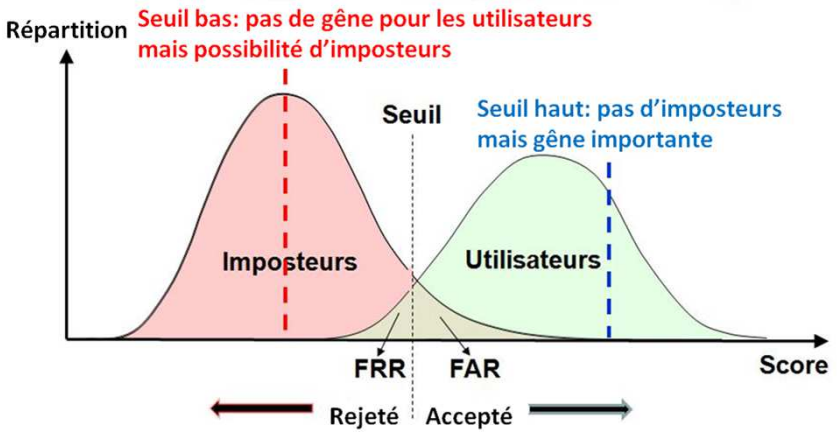


FIG. 2.1 – Taux de vraisemblance des utilisateurs légitimes et des imposteurs d'un système d'authentification biométrique (dont la comparaison est basée sur le calcul d'une similarité).

C. Taux d'erreur de systèmes d'identification

– Taux d'identification (*identification rate*, IR) : le taux d'identification au rang r est la proportion de transactions d'identification, par des utilisateurs enrôlés dans le système, pour lesquels l'identifiant de l'utilisateur est dans les r identifiants retournés ;

– Taux de faux-négatif d'identification (*false-negative identification-error rate*, FNIR) : proportion de transactions d'identification, par des utilisateurs enrôlés dans le système, pour lesquels l'identifiant de l'utilisateur ne figure pas dans la liste des identifiants retournée.

Exemple : Pour une transaction d'identification à une seule tentative contre une base de données contenant N modèles, le taux de faux-négatif d'identification est calculé par :

$$FNIR(\tau) = FTA + (1 - FTA) * FNMR(\tau) \tag{2.3}$$

– Taux de faux-positif d'identification (*false-positive identification-error rate*, abrégé par FPIR) : proportion de transactions d'identification, par des utilisateurs non enrôlés dans le système, pour lesquels la liste des identifiants retournée est non vide.

Exemple : Pour une transaction d'identification à une seule tentative contre une base de données contenant N modèles, le taux de faux-positif d'identification est calculé par :

$$FPIR = (1 - FTA) * (1 - (1 - FMR)^N) \tag{2.4}$$

– Erreur de l'algorithme de présélection (*pre-selection error*) : l'algorithme de présélection permet de réduire le nombre de modèles biométriques à comparer avec l'image acquise pendant la phase d'identification. L'erreur de l'algorithme de présélection est l'erreur qui se produit quand le modèle correspondant à la donnée biométrique acquise ne figure pas dans la liste des modèles retournée ;

– Taux de pénétration (*penetration rate*, PR) : mesure, en moyenne, le nombre de modèles biométriques présélectionnés par rapport au nombre total de modèles.

2.2.1.2 Les mesures de temps de traitements et occupation mémoire

Le temps de traitement de l'information, par le système, est un facteur très important pour l'évaluation de systèmes biométriques. Il est généralement mesuré en :

– Temps moyen d'enrôlement : désigne le temps moyen pour générer les modèles biométriques des individus ;

– Temps moyen de vérification : désigne le temps moyen pour l'acquisition des données biométriques requises et la comparaison de ces données avec le modèle correspondant. Ce temps ne dépend pas du nombre de personnes dans la base de données ;

– Temps moyen d'identification : désigne le temps moyen pour l'acquisition des données biométriques requises et la comparaison de ces données avec les modèles existants dans la base. Le nombre d'utilisateurs du système a un impact très important sur cette information. Il peut être conséquent pour de grandes bases, comme cela peut être le cas lors de contrôle douanier.

L'espace mémoire, requis par le système, est également un facteur important à prendre en considération lors de l'évaluation de systèmes biométriques. Il est généralement mesuré en *taille moyenne et maximale d'un modèle biométrique* et en *espace mémoire maximal alloué* pendant les phases d'enrôlement, de vérification et d'identification.

2.2.1.3 Les courbes de performance

La performance d'un système biométrique pour différents paramétrages (seuil de décision) est illustrée graphiquement en utilisant des courbes spécifiques. L'échelle logarithmique est parfois utilisée, pour les rendre plus lisible et plus exploitable, surtout dans le cas de comparaison des systèmes biométriques qui ont des performances similaires. Nous trouvons ainsi :

– Courbe ROC (*Receiver operating characteristic curve*) [58] : cette courbe constitue l'une des méthodes les plus couramment utilisées afin d'évaluer la performance globale d'un système d'authentification biométrique. La courbe ROC représente la relation entre le taux de fausses acceptations (FAR) et le taux de faux rejets (FRR) pour les différentes valeurs du seuil de décision, respectivement en abscisses et en ordonnées. Au lieu de la courbe ROC, parfois le terme DET (*détection d'erreur Tradeoff*) est utilisé. Dans ce cas, le terme ROC est réservé pour représenter le taux de vrais rejets (1-FRR) contre le taux de fausses acceptations (FAR). Une illustration de la courbe ROC est donnée à la figure 2.2. L'avantage de cette méthode est qu'on obtient une

représentation compacte de la performance d'un système biométrique pour ses différents paramétrages au travers d'une seule courbe, qui permet de comparer objectivement différents systèmes biométriques ;

– Courbe CMC (*Cumulative match characteristic curve*) : cette courbe présente les valeurs du rang d'identification et les probabilités d'une identification correcte inférieure ou égale à ces valeurs, respectivement en abscisses et en ordonnées. Cette courbe est utilisée pour comparer la performance de systèmes d'identification biométrique. Des exemples de cette courbe sont donnés à la figure 2.3 ;

– Courbe RC (*Robustness curve*) : cette courbe illustre la robustesse du système en terme de performance contre les divers types d'altérations (*i.e.*, altérations dues au bruit pendant l'acquisition de données biométriques). La performance du système est illustrée par le point de fonction taux d'égale erreur (EER) que l'on présente dans la section 2.2.1.4. Un exemple de cette courbe présenté par Cherifi *et al.* [61] est donné à la figure 2.4. Les auteurs ont généré des données biométriques synthétiques (*cf.*, figure 2.5) à partir de modèles stockés dans la base de données afin de tester l'efficacité de leur système. Le système testé est GREYC-Keystroke présenté dans la section 1.3.2.1.

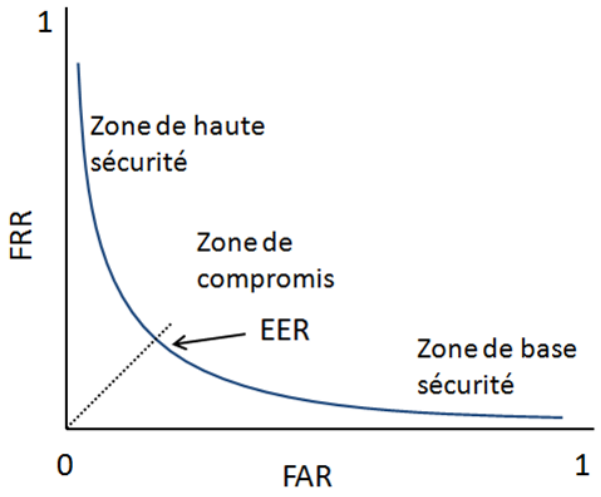


FIG. 2.2 – Exemple de la courbe ROC : Variation du FRR en fonction de FAR lorsque le seuil de décision varie.

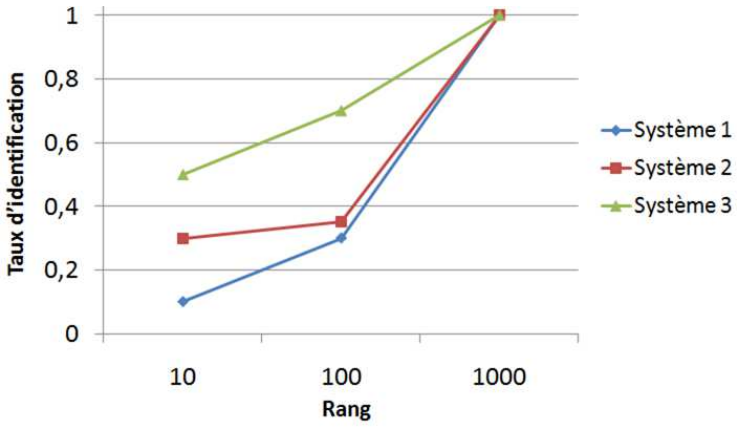


FIG. 2.3 – Exemple de courbes *CMC* pour différents systèmes biométriques.

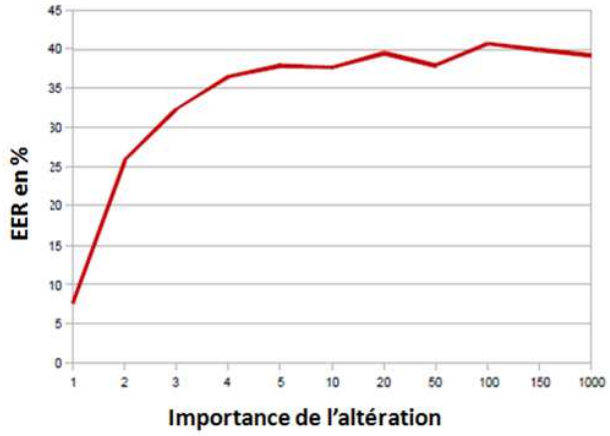


FIG. 2.4 – Evolution des valeurs de l'EER en fonction de la quantité des altérations.

FIG. 2.5 – Exemples des données générées synthétiquement par rapport à la moyenne des données biométriques acquises.

2.2.1.4 Les points de performance

Les points de performance sont utilisés pour illustrer la performance des systèmes biométriques. Il existe dans la littérature plusieurs métriques [5, 60] que sont le taux d'égale erreur (EER), le taux d'erreur pondéré (WER), le taux d'erreur moyenne (HTER), l'aire sous la courbe ROC (AUC) et la capacité.

- Taux d'égale erreur (*Equal Error Rate*, EER) : l'EER est obtenu à l'intersection de la courbe ROC et de la droite d : FAR = FRR (*cf.*, figure 2.2). Cette valeur n'a presque pas d'utilité pratique car on ne souhaite généralement pas que le FAR et le FRR soient les mêmes, mais elle constitue un indicateur de la précision du dispositif biométrique. En d'autres termes, plus l'EER est faible, plus le système est performant. À noter que ce taux d'erreur est le plus couramment utilisé dans la littérature pour illustrer la performance des systèmes biométriques.

- Taux d'erreur pondéré (*Weighted Error Rate*, WER) : ce taux d'erreur correspond au seuil tel que le FRR soit proportionnel au FAR avec un coefficient qui dépend de l'application. Pour un coefficient égal à 1, le seuil du WER est égal au seuil de l'EER ;

- Taux d'erreur moyenne (*Half Total Error Rate*, HTER) : c'est une métrique qui correspond à la moyenne entre le FAR et FRR pour un seuil fixé τ :

$$HTER(\tau) = \frac{FAR(\tau) + FRR(\tau)}{2} \tag{2.5}$$

  Théoriquement, l'HTER est utilisée pour approximer l'EER dans le cas ou les deux taux d'erreur FAR et FRR sont du même ordre de grandeur. D'une manière générale, l'HTER est utilisée pour quantifier la performance du système dans le cas où la distribution de scores des utilisateurs légitimes et d'imposteurs n'est pas disponible (comme le cas des systèmes biométriques commerciaux). Elle est estimée en utilisant le seuil de décision opérationnel τ du système.

- Aire sous la courbe ROC (*Area Under ROC Curve*, AUC) : c'est une métrique qui permet de quantifier la diversification de la distribution de scores des utilisateurs légitimes et d'imposteurs. En d'autres termes, étant donnés deux utilisateurs choisis au hasard, un parmi les utilisateurs légitimes et l'autre parmi les imposteurs, l'AUC représente la probabilité $P(S^{gen} > S^{imp})$ (*i.e.*,

probabilité de bon classement). Plusieurs méthodes sont proposées dans [62] pour estimer l'AUC. Tronci *et al.* [63] suggèrent une estimation de l'AUC basée sur le test statistique de Wilcoxon-Mann-Whitney (WMW) [64]. L'AUC est ainsi définie par :

$$AUC = \frac{\sum_{p=1}^{n_g} \sum_{q=1}^{n_i} I(S_p^{gen}, S_q^{imp})}{n_g n_i} \tag{2.6}$$

où n_g et n_i représentent le nombre des utilisateurs légitimes et d'imposteurs respectivement, $\{S_p^{gen}\}$ et $\{S_q^{imp}\}$ correspondent aux scores des utilisateurs légitimes et d'imposteurs respectivement, et la fonction $I(S_p^{gen}, S_q^{imp})$ est définie par :

$$I(S_p^{gen}, S_q^{imp}) = \begin{cases} 1 & si\ S_p^{gen} > S_q^{imp} \\ 0 & sinon \end{cases} \tag{2.7}$$

L'AUC constitue également un bon indicateur pour évaluer et comparer les systèmes biométriques. Plus l'AUC est grande, plus l'algorithme est performant.

– FAR ou FRR fixé : l'EER estime un taux d'erreur sur la courbe où on a autant de faux rejets que de fausses acceptations. Cependant, il ne donne pas beaucoup d'informations sur l'interaction entre les données biométriques de différents utilisateurs (*i.e.*, chevauchement des scores intraclasses et interclasses). Ainsi, il est judicieux pour certaines applications de fixer un de ces deux taux d'erreur (FAR ou FRR). Dans ce cas, la performance du système est donnée par le taux FAR (ou FRR) pour un FRR (ou FAR) fixé.

– Capacité d'un système biométrique (*Constrained capacity of biometric system*) : la capacité permet de quantifier la performance de systèmes biométriques, en utilisant la base de données des utilisateurs et la fonction de similarité. Bhatnagar et Kumar [60] ont caractérisé la distribution de scores des utilisateurs légitimes et d'imposteurs par la distribution gaussienne comme le montre la figure 2.6. Les indices de performances proposés sont :

  • La capacité d'un utilisateur m, notée par C_m, permet d'illustrer le degré de distinction de l'utilisateur m par rapport aux autres utilisateurs de la base. Elle est donnée par :

$$C_m = \frac{1}{2} \log_2 \left(1 + \frac{d_m^2}{4 \max(\sigma_{g(m)}^2, \sigma_{i(m)}^2)} \right) \tag{2.8}$$

où d_m est la distance entre les médianes $\hat{g}_m$ et $\hat{i}_m$, $\sigma^2_{g(m)}$ et $\sigma^2_{i(m)}$ sont les variances de la distribution de scores intraclasses (échantillons du même utilisateur) et interclasses (échantillons d'utilisateurs différents), respectivement.

- La capacité d'un système biométrique, notée par C_s, permet d'illustrer la fiabilité du système en terme de nombre d'utilisateurs correctement authentifiés (*i.e.*, peuvent être authentifiés d'une manière sûre). Elle est donnée par :

$$C_s = \frac{1}{2}\log_2\left(1 + \frac{\overline{d_m^2}}{4\max(\overline{\sigma_g^2}, \overline{\sigma_i^2})}\right) \tag{2.9}$$

où $\overline{d_m}$ est la moyenne des distances d_m pour chaque utilisateur de la base, $\overline{\sigma_g^2}$ et $\overline{\sigma_i^2}$ sont les moyennes des variances des distributions intraclasses et interclasses, respectivement.

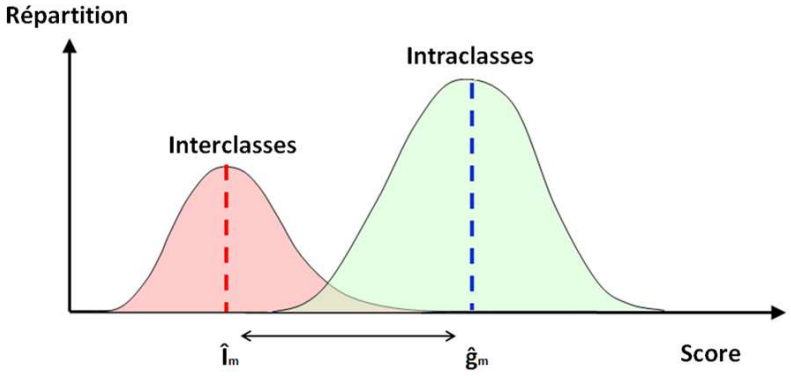


FIG. 2.6 – Distribution de scores pour l'utilisateur m.

2.2.1.5 Intervalle de confiance

En biométrie, les bases de données collectées sont utilisées pour évaluer la performance des systèmes biométriques. Cependant, ces bases ne sont pas représentatives de la population globale pour deux raisons principales. Premièrement, ces bases ne contiennent pas assez de personnes, et on a en général peu de données par personne. Deuxièmement, il y a souvent une différence entre le nombre de scores des utilisateurs légitimes et d'imposteurs, ce qui n'est également pas représentatif de la réalité. Enfin, les taux d'erreur (EER, WER, HTER et AUC) utilisés pour illustrer la performance globale du système dépendent du découpage *enrôlement - test*. Pour toutes ces raisons, il est nécessaire de calculer un intervalle de confiance à l'EER lors de la comparaison

des systèmes biométriques. Cet intervalle de confiance est surtout indispensable lors de la comparaison des systèmes biométriques ayant des taux d'erreurs similaires.

Bolle *et al.* [65] ont introduit une méthode non-paramétrique nommée *bootstrap* pour estimer l'intervalle de confiance associé aux taux d'erreurs FAR et FRR. Il s'agit d'une technique qui permet de faire de l'inférence statistique sur des nouveaux échantillons tirés à partir de l'échantillon initial. Ce ré-échantillonnage consiste en un tirage aléatoire (avec remise) avec remplacement de M exemples de la base de test. Pour l'estimation du FRR, nous pouvons prendre $M = N_i$ tandis que $M = N_l$ pour le FAR, N_i et N_l étant le nombre de scores des utilisateurs légitimes et d'imposteurs, respectivement. À chaque tirage k, nous avons une estimation de FAR (τ) et FRR (τ) pour une valeur du seuil τ, et ainsi une estimation de l'EER_k. Nous répétons ensuite cette procédure k fois pour calculer ainsi l'intervalle de confiance. Selon Allano [66], $k = 1000$ tirages sont suffisants.

D'après la loi de grands nombres, lorsque k tend vers l'infini, la variable à estimer (l'EER par exemple) tend vers une variable normale. L'intervalle de confiance (IC) peut être ainsi déterminé grâce aux percentiles de la distribution normale. L'intervalle de confiance à 95% est défini par :

$$IC = EER \pm 1.96 * \frac{\sigma}{\sqrt{k}} \tag{2.10}$$

où EER est le taux d'erreur global estimé sur l'échantillon initial, k est le nombre de tirages, et σ la variance des k taux d'erreurs calculés sur les différents tirages k. À noter que pour un intervalle de confiance à 90%, il suffit de remplacer la valeur de $1,96$ par $1,645$. L'intervalle de confiance représente ainsi une mesure de confiance sur le taux d'erreur estimé. Plus l'intervalle présenté dans l'équation 2.10 est petit, plus le taux d'erreur calculé est fiable.

2.2.1.6 Discussion

Traditionnellement, le point particulier EER est le plus couramment utilisé dans la littérature pour évaluer et comparer les systèmes biométriques. Dans la section précédente, nous avons vu la faiblesse d'utiliser seulement l'EER pour comparer les systèmes biométriques. Pour des systèmes biométriques ayant des taux d'erreur différents, l'utilisation de l'EER peut être suffisante pour affirmer qu'un système est meilleur qu'un autre. Tandis que dans le cas où les systèmes à comparer présentent des taux d'erreurs similaires (lors des compétitions), utiliser un métrique complémentaire devient indispensable. Nous avons vu qu'il existe dans la littérature d'autres

métriques complémentaire à l'EER (l'AUC, la courbe RC, la capacité et l'intervalle de confiance associé à l'EER) que nous pouvons utiliser afin de comparer les systèmes biométriques dans un cadre précis. L'AUC présentée dans [63], qui permet de quantifier la diversification de scores des utilisateurs légitimes et d'imposteurs, est un bon indicateur de performance complémentaire à l'EER. Il permet de bien représenter les performances globales de l'algorithme. La courbe robustesse RC présentée dans [61], est également une métrique efficace qui permet de quantifier la robustesse des systèmes face aux altérations. Cette métrique doit être prise en considération surtout dans les cas des systèmes basés sur l'analyse comportementale. Prenons le cas de la modalité dynamique de frappe au clavier, la figure 2.4 montre qu'une légère modification a un impact important sur les taux d'erreur EER, ce qui signifie que le système testé n'est pas très robuste face aux altérations. La capacité présentée dans [60] permet, en utilisant la base de données biométrique et la fonction de similarité, 1) de comparer les systèmes biométriques, et 2) de quantifier la performance de chaque utilisateur de la base, en illustrant le nombre d'utilisateurs proche de son modèle. Enfin, nous pouvons conclure que les métriques AUC, capacité et RC sont complémentaires à l'EER pour avoir une meilleure précision sur la performance du système testé.

2.2.2 Benchmarks

Afin d'évaluer les systèmes biométriques, nous avons besoin d'une base de données dédiée à cette évaluation. Cette base assure que les systèmes seront testés selon les mêmes conditions d'acquisition. Les bases de données biométriques peuvent servir également à régler les paramètres d'un système monomodal (paramétrage du seuil de décision) et multimodal (paramétrage des poids pour la fusion). Les bases de données biométriques collectées sont généralement divisées en deux types : bases de données réelles et synthétiques.

2.2.2.1 Bases de données réelles

Ces bases contiennent des données biométriques réelles acquises grâce à la participation de volontaires. Dans la littérature, il existe deux ensembles des bases : 1) bases monomodales et 2) multimodales. Un exemple de bases monomodales est *FACES94* [67], *AR* [68], *FERET* [69, 70], *FVC2002 DB*$_2$ [71], FRGC (Face Recognition Grand Challenge) [72], *USF Human ID Gait Baseline* [73], *ENSIB* [74], *GREYC-Keystroke* [29], *etc.* Un exemple des bases multimodales est *XM2VTSDB* [75], *BANCA* [76],

BIOSECURE [77], *etc.*

Dans cette section, nous présentons uniquement les bases de données utilisées dans le cadre de cette thèse que sont *FACES94*, *AR*, *FERET*, *FVC2002* DB_2 et *ENSIB*.

BD_1 *FACES94*

La base de données FACES94 a été collectée sur 152 personnes (20 images par personne). Les images de cette base ont été capturées avec des conditions différentes d'expression faciale. La figure 2.7 présente un exemple des images de cette base.

FIG. 2.7 – Exemple de visages de la base *FACES94* (source [67]).

BD_2 *AR*

La base de données AR a été créée par Aleix Martinez et Robert Benavente au *Computer Vision Center (CVC)*. Elle a été collectée sur 120 personnes (26 images par personne) avec des conditions différentes d'expression faciale, d'éclairage, et d'occultation (lunettes de soleil et écharpe). Un exemple de visages de cette base est présenté dans la figure 2.8.

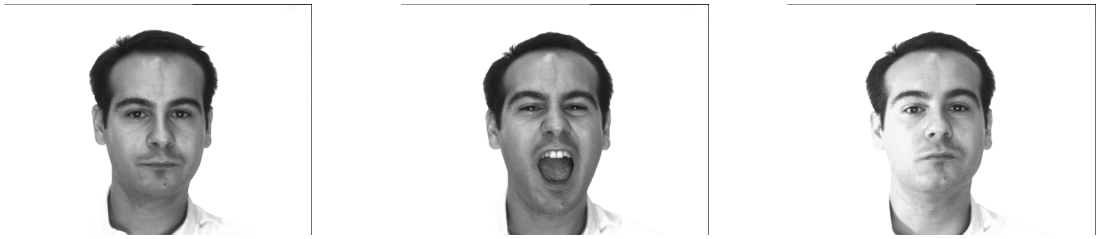


FIG. 2.8 – Exemple de visages de la base *AR* (source [68]).

BD_3 *FERET*

FERET est une base de données unimodale destinée à quantifier la performance des systèmes de reconnaissance faciale. Elle a été collectée sur 1199 personnes

en plusieurs sessions entre août 1993 et juillet 1996. Pour certains volontaires, le temps séparant la première et dernière image acquise est supérieur à deux ans. Les images de cette base ont été collectées avec des conditions différentes d'expression, de pose, d'éclairage et d'âge. La figure 2.9 représente un exemple de visages extraits de cette base de données.

FIG. 2.9 – Exemple de visages de la base *FERET* (source [69]).

BD_4 *FVC*2002 DB_2

La base DB_2 est une base de données d'empreintes digitales utilisée dans la compétition *Fingerprint Verification Competition* (FVC2002). Elle est composée de 100 personnes (8 images par personne). La figure 2.10 illustre un exemple des images de cette base.

FIG. 2.10 – Exemple d'empreintes digitales de la base *FVC2002* DB_2 (source [71]).

BD_5 *ENSIB*

ENSIB est une base de données de visages collectés en 2007. Elle est composée de 100 personnes (40 images par personne), contenant des images capturées avec des conditions différentes de pose (de gauche à droite). La figure 2.11 illustre un exemple des images de cette base.

FIG. 2.11 – Exemple de visages de la base ENSIB (source [74]).

2.2.2.2 Bases de données synthétiques

Ces bases contiennent des données synthétiques simulant des données biométriques réelles. Une base synthétique doit satisfaire deux propriétés. Premièrement, la performance issue d'une base synthétique doit être proche de celle obtenue avec une base de données réelle. Deuxièmement, une donnée dans la base synthétique ne doit pas représenter une donnée biométrique réelle d'un individu. La base *SFinGe* générée par le logiciel *SFinGe* [4], développé par le laboratoire italien *BioLab*[1], est un exemple de base synthétique. La figure 2.12 illustre quelques empreintes digitales générées par ce logiciel.

FIG. 2.12 – Exemple d'empreintes synthétiques générées par *SFinGe* (source [78]).

2.2.3 Compétitions

La comparaison des systèmes biométriques nécessite un protocole d'évaluation et une base de données bien définis. Le protocole d'évaluation assure que les systèmes biométriques sont testés sous les mêmes conditions. Prenons le cas d'un système de reconnaissance faciale, il est quasiment impossible de comparer l'EER d'un système utilisant une seule image pour l'enrôlement avec d'autres systèmes utilisant cinq images. Pour régler ce problème de disparité dans les protocoles d'évaluation des systèmes biométriques, plusieurs compétitions ont été organisées. Ces compétitions sont

1. http ://biolab.csr.unibo.it/

ouvertes à tous (industriels, laboratoires de recherche, et développeurs indépendants) et en général organisées par un organisme indépendant afin de pouvoir comparer dans un cadre précis les systèmes en question. Nous présentons dans cette section un aperçu sur les compétitions les plus citées dans la littérature et qui sont divisées en deux types : les compétitions monomodales et multimodales.

2.2.3.1 Compétitions monomodales

C_1 *Signature Verification Competition (SVC)*

SVC [79] est une compétition de vérification par dynamique de signature qui a été organisée en collaboration avec la conférence internationale ICBA (*International Conference on Biometric Authentication*) en 2004. Le taux d'erreur EER a été utilisé comme indicateur de performance ;

C_2 *Fingerprint Verification Competition (FVC)*

Les compétitions de reconnaissance d'empreinte digitale ont été organisées par plusieurs universités [2] en 2000, 2002, 2004 et 2006. Les participants ont testé leurs algorithmes en fournissant leurs fichiers exécutables correspondants aux phases d'*enrôlement* et de *vérification*. Quatre bases de données, dont trois sont réelles et une synthétique générée par le logiciel *SFinGe*, ont été employées durant la compétition *FVC2006*. Les différents indicateurs de performance utilisés sont : la distribution de scores des utilisateurs légitimes et d'imposteurs, la taille des modèles biométriques, le temps moyen d'enrôlement et de vérification, taux d'échec à l'enrôlement et les courbes ROC ;

C_3 *Face Recognition Vendor Test (FRVT)* et *Iris Challenge Evaluation (ICE)*

Les compétitions de reconnaissance faciale (FRVT) et d'iris (ICE) [80] sont des compétitions organisées par l'Institut National des Standards et de la Technologie (NIST). Ces compétitions utilisent les courbes ROC comme indicateur de performance ;

C_4 *Speaker Recognition Evaluation (SRE)*

Les compétitions de reconnaissance vocale SRE ont été organisées par le NIST. Il y a eu un grand nombre des compétitions SRE [3] depuis 1997 et pour la $12^{\text{ème}}$ fois en 2010.

2. `http://bias.csr.unibo.it/fvc2006/`
3. `http://www.itl.nist.gov/iad/mig/tests/sre/`

2.2.3.2 Compétitions multimodales

C_5 *BioSecure Multimodal Evaluation Campaign (BMEC)*

BMEC [4] est une compétition organisée par le réseau d'excellence BioSecure [5] en 2007. Cette compétition comprenait des parties pour l'évaluation des systèmes biométriques monomodaux et une partie pour les systèmes multimodaux (fusion au niveau de scores). La base multimodale BioSecure [77] a été utilisée lors de cette compétition. Les détails du protocole utilisé et les résultats sont présentés par Mayoue *et al.* [81] ;

C_6 *Multiple Biometric Grand Challenge (MBGC)*

MBGC [82] est une compétition multimodale organisée par le NIST en 2009. L'objectif de cette compétition consiste à améliorer les systèmes de reconnaissance basés sur le visage et l'iris selon différentes conditions d'acquisition. Elle consiste également à évaluer les algorithmes de fusion, au niveau des images et au niveau de scores de ces deux modalités.

2.2.4 Plateformes

Outre les compétitions, il existe plusieurs plateformes permettant aux chercheurs et aux développeurs de tester leurs algorithmes par rapport à l'existant dans l'état de l'art. Nous présentons dans cette section un aperçu sur les plateformes existantes que sont BioSecure Reference and Evaluation framework, le logiciel GREYC-Keystroke et FVC-onGoing.

P_1 *BioSecure Reference and Evaluation framework*

En plus de la compétition multimodale BMEC, le réseau d'excellence BioSecure a proposé un Framework [83] visant à comparer et évaluer les systèmes biométriques. Ce Framework comporte douze systèmes de référence relatifs aux modalités suivantes : visage, voix, iris, empreinte digitale, géométrie de la main et signature dynamique. Ces systèmes de références sont accessibles sur `http://svnext.it-sudparis.eu/svnview2-eph/ref_syst/`. Ces systèmes sont implémentés en quatre modules remplaçables (prétraitement, extraction de caractéristiques, génération du modèle biométrique et vérification) permettant aux développeurs et aux chercheurs de tester et d'évaluer une partie spécifique de leur système. En d'autres termes, un chercheur qui souhaite évaluer son algorithme de vérification peut le tester en remplaçant le

4. `http://biometrics.it-sudparis.eu/BMEC2007/`
5. `http://biosecure.it-sudparis.eu`

module correspondant dans le système de référence par cet algorithme. L'indicateur de performance utilisé est la courbe DET avec son EER correspondant ;

P_2 *GREYC-Keystroke*

Le logiciel *GREYC-Keystroke* [29], développé dans le laboratoire de recherche GREYC, est un outil qui permet de comparer les systèmes biométriques basés sur la dynamique de frappe au clavier. Il est également utilisé pour créer une base de données afin de comparer les différents algorithmes qui existent dans l'état de l'art sous les mêmes conditions d'acquisition. Les indicateurs de performance proposés sont : la distribution de scores des utilisateurs légitimes et d'imposteurs, les courbes ROC et le taux d'échec à l'acquisition (FTA) ;

P_3 *Fingerprint Verification Competition-onGoing (FVC-onGoing)*

FVC-onGoing est une plate-forme d'évaluation en ligne [6] dédiée aux algorithmes de reconnaissance des empreintes digitales. Elle est l'évolution des compétitions FVC, présentées dans la section précédente. La plate-forme propose plusieurs bases de données regroupées en deux parties : la première (*Fingerprint Verification*) quantifie les deux modules d'enrôlement et de vérification, tandis que la seconde (*Fingerprint Matching (ISO)*) quantifie seulement le module de vérification sur des Templates ISO [84] basés sur les minuties. Les indicateurs de performance proposés sont : le taux d'échec à l'enrôlement et l'acquisition (FTE et FTA), le taux de non-correspondance (FNMR) pour un taux de fausse correspondance fixé et vice versa, le temps moyen d'enrôlement et de vérification, la taille maximale requise pour stocker le modèle biométrique sur le support, la distribution de scores des utilisateurs légitimes et d'imposteurs, et la courbe ROC avec son EER correspondant.

2.2.5 Conclusion

Nous avons présenté dans la section 2.2 les outils pour évaluer les performances d'un système biométrique monomodal et multimodal. Les tableaux présentés dans l'annexe B récapitulent les principales propriétés de ces outils que sont les taux d'erreur, les bases de données, les compétitions et les plateformes, respectivement.

Les outils présentés sont efficaces pour évaluer et comparer les systèmes biométriques. Cependant, malgré la diversité de ces outils, cet aspect d'évaluation ne prend

6. `https://biolab.csr.unibo.it/FVCOnGoing`

pas en considération la qualité des données biométriques acquises, ni la manière dont l'utilisateur interagit avec le système biométrique. De plus, cet aspect d'évaluation ne mesure pas la robustesse du système contre la fraude. Nous présentons donc par la suite un état de l'art sur la qualité de données biométriques, suivie par les travaux liés à l'usage d'un système biométrique ainsi que sa robustesse contre la fraude.

2.3 Qualité des données biométriques

Les systèmes d'authentification basés sur la biométrie sont beaucoup moins précis que ceux basés sur *ce que l'on sait* ou *ce que l'on possède*. Contrairement aux systèmes traditionnels, la comparaison de deux vecteurs biométriques est exprimée en pourcentage de similarité qui n'atteint quasiment jamais 100%. Cette variation des résultats est plutôt liée aux artefacts d'acquisition qu'à l'échantillon biométrique de l'individu qui est généralement stable avec le temps. Prenons le cas des systèmes d'authentification par empreintes digitales : la qualité d'image de l'empreinte peut varier selon le degré de saleté de la peau du doigt, son niveau d'humidité, son aspect huileux ou son aspect dégradé (*ex.*, coupure). La pression que l'on exerce sur le capteur utilisé est aussi déterminante quant aux détails qui sont recueillis. Ainsi, pour qu'un système soit opérationnel et efficace contre les divers types de bruit d'acquisition, contrôler la qualité des données acquises devient indispensable.

L'évaluation de la qualité des données biométriques est un domaine de recherche récent en biométrie. Les travaux effectués sur la qualité sont très peu abordés par rapport aux autres recherches sur l'extraction de paramètres et la reconnaissance. Il a pris de plus en plus d'attention dans la communauté biométrique après les résultats de la compétition FVC [85]. En effet, les bases de données collectées pour les compétitions FVC en 2004 et 2006 ont été intentionnellement altérées par rapport à celles organisées en 2000 et 2002. Le tableau 2.1 montre clairement la dégradation de la performance de l'algorithme testé avec les images altérées. En effet, l'EER augmente en moyenne de $0,96$ en 2000 et 2002 (acquisition contrôlée) jusqu'à $2,115$ en 2004 et 2006 (acquisition intentionnellement altérée). Ceci montre clairement l'impact de la qualité des données acquises sur la performance globale d'un système biométrique.

Dans la suite, nous commençons par définir la qualité des données biométriques. Nous présentons ensuite les facteurs impactant sur la qualité des données dans la section 2.3.2, ainsi que les travaux de la littérature dans la section 2.3.3.

Base de données	**2000**	**2002**	**2004**	**2006**
DB_1	0.67%	0.1%	1.97%	5.56%
DB_2	0.61%	0.14%	1.58%	0.02%
DB_3	3.64%	0.37%	1.18%	1.53%
DB_4	1.99%	0.1%	0.61%	0.27%
Moyenne	0.96%		2.115%	

TAB. 2.1: Les valeurs de l'EER de l'algorithme le plus performant sur chaque base de données utilisée dans les quatre compétitions FVC (extrait de [86]).

2.3.1 Définition de la qualité des données biométriques

Selon l'Organisation Internationale de Normalisation ISO/IEC 29794-1 [87], la qualité des données biométriques peut être considérée selon trois points de vues différents comme le montre la figure 2.13 : 1) caractéristiques de la source (*character*), 2) fidélité (*fidelity*) et 3) utilité (*utility*). Les **caractéristiques de la source** se réfèrent à la qualité intrinsèque attribuable aux caractéristiques physiques de l'individu. La **fidélité** illustre le degré de similarité entre l'échantillon biométrique acquis et son modèle biométrique correspondant. L'**utilité** se réfère à l'impact de l'échantillon biométrique acquis sur la performance globale d'un système biométrique. En biométrie, la qualité des données biométriques est généralement reliée à l'utilité [6]. Ainsi, la qualité devrait être un facteur prédictif de la performance du système biométrique. La qualité des données biométriques peut être utilisée dans plusieurs applications :

- Acceptation pour l'enrôlement : pour avoir des modèles biométriques de bonne qualité surtout pour les applications utilisant une seule donnée biométrique. Un exemple d'une telle application est le passeport biométrique (une seule image de visage est utilisée pour l'enrôlement) ;

- Acceptation pour la vérification : parmi les données biométriques acquises, quelle donnée est à envoyer pour la vérification ? ou faut-il en acquérir plus ? ;

- Acceptation pour l'identification : pour vérifier si l'utilisateur fournit volontairement des mauvaises données biométriques pour éviter la reconnaissance ;

- Multimodalité : l'information sur la qualité peut être utilisée dans les approches multimodales [88, 89, 90]. Poh *et al.* [89] ont étudié la performance des méthodes de fusion multimodale lors du changement de la qualité des données biométriques (*i.e.*, changement qui peut être dû au remplacement du capteur d'acquisition). Les résultats de la campagne d'évaluation, sur 22 méthodes de fusion, montrent

que les meilleures méthodes sont celles qui exploitent la qualité lors du processus de fusion.

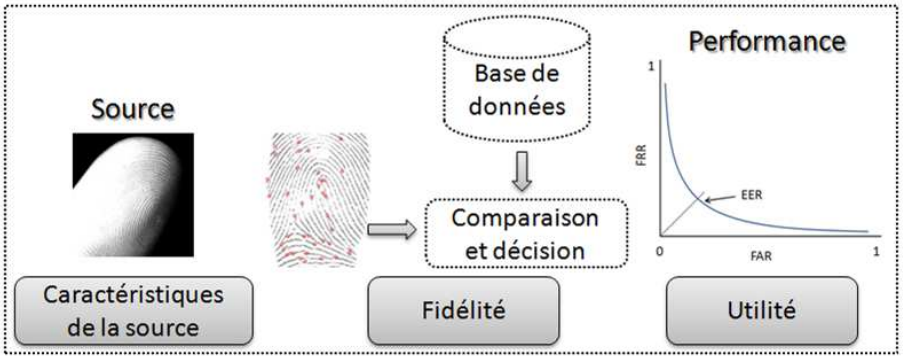


FIG. 2.13 – Définition de la qualité des données biométriques selon les trois axes : caractéristiques de la source, fidélité et utilité.

2.3.2 Les facteurs dégradant la qualité des données biométriques

Il existe plusieurs facteurs qui affectent la qualité des données biométriques. Ces facteurs peuvent être classés en trois familles que sont les facteurs physiologiques, comportementaux et environnementaux.

Les **facteurs physiologiques** sont totalement liés à l'individu et ainsi difficiles à contrôler. Le tableau 2.2 illustre l'impact de ces facteurs sur quelques modalités. Prenons le cas de l'âge : bien que l'empreinte digitale soit une donnée stable, la vérification est moins précise jusqu'à l'adolescence et au cours de la vieillesse.

Les **facteurs comportementaux** sont liés au comportement de l'individu lors de la phase d'enrôlement et de vérification. Ces facteurs sont également difficiles à contrôler. Le tableau 2.3 présente l'impact de ces facteurs sur les modalités utilisées. Prenons le cas de la nervosité en dynamique de frappe au clavier : d'une manière générale, un individu énervé tape de manière plus frénétique qu'à l'accoutumée. Les techniques de mise à jour des modèles biométriques [91], sont considérées comme un bon moyen pour tolérer ces variations.

Les **facteurs environnementaux** sont généralement liés aux artefacts d'acquisition. D'une manière générale, certains de ces facteurs peuvent être contrôlables contrairement aux facteurs physiologiques et comportementaux. L'impact de ces facteurs sur quelques modalités biométriques est donné au tableau 2.4. Prenons le

cas du lieu du capteur d'acquisition (intérieur/extérieur) : pour qu'un système vocal soit opérationnel, il faut s'assurer que le capteur d'acquisition soit installé dans un endroit calme pour éviter les parasites lors de l'acquisition.

Nous avons vu dans cette section qu'il existe trois familles de facteurs qui peuvent dégradés considérablement la performance globale d'un système biométrique. Certains de ces facteurs peuvent être contrôlés et d'autres sont non maîtrisables. Dans un système basé sur la dynamique de frappe au clavier, il sera plus difficile de contrôler le comportement d'un utilisateur qu'installer le capteur dans un endroit calme pour éviter les bruits d'acquisition. Pour toutes ces raisons, contrôler la qualité des données biométriques est un enjeu majeur en biométrie. Dans la prochaine section, nous présentons un état de l'art sur l'évaluation de la qualité d'une donnée biométrique.

Facteurs	**Empreinte digitale**	**Iris**	**Visage**	**Voix**	**Dynamique de frappe au clavier**	**Main**
Age	×	×	×	×	×	×
Genre			×	×		
Ethnicité		×	×			
Blessures	×	×	×	×	×	×

TAB. 2.2: Facteurs physiologiques ayant un impact sur la qualité des données biométriques (extrait de [86]).

Facteurs	**Empreinte digitale**	**Iris**	**Visage**	**Voix**	**Dynamique de frappe au clavier**	**Main**
Fatigue	×	×	×	×	×	×
Distraction	×	×	×	×	×	×
Coopérativité	×	×	×	×	×	×
Motivation	×	×	×	×	×	×
Nervosité	×	×	×	×	×	×
Distance		×	×	×	×	
Fermeture des yeux		×	×			
Pression sur le capteur	×				×	×
Expression faciale			×			
Coiffure, barbe et maquillage			×			
Vêtements			×			
Chapeau			×			
Bijoux	×		×			×
Lunettes ou lentilles de contact		×	×			

TAB. 2.3: Facteurs comportementaux ayant un impact sur la qualité des données biométriques (extrait de [86]).

Facteurs	**Empreinte digitale**	**Iris**	**Visage**	**Voix**	**Dynamique de frappe au clavier**	**Main**
Intérieur/extérieur	×	×	×	×	×	×
Fond			×			
Température	×					×
Humidité	×					×
Éclairage	×	×	×			
Réflexion de la lumière		×	×			
Bruits				×		

TAB. 2.4: Facteurs environnementaux ayant un impact sur la qualité des données biométriques (extrait de [86]).

2.3.3 État de l'art

Dans cette section, nous présentons un aperçu des méthodes existantes destinées à évaluer la qualité des données biométriques morphologiques. Nous n'abordons pas les méthodes de qualité comportementales dans le cadre de cette thèse.

Alonso-Fernandez *et al.* [92] ont présenté un aperçu des méthodes existantes visant à quantifier la qualité d'empreintes digitales. Les auteurs montrent l'impact des images de mauvaise qualité sur la performance globale des systèmes biométriques. D'autres méthodes pour mesurer la qualité d'empreintes digitales sont données dans [93, 94, 95]. Les méthodes présentées ont montré leur efficacité sur la prédiction de la qualité des images d'empreintes digitales. Cependant, ces méthodes dépendent de la modalité biométrique considérée, et ainsi ne peuvent pas être exploitables pour d'autres types de modalité (comme le visage). La métrique *NIST Fingerprint Image Quality* (NFIQ) proposée par Tabassi et Wilson [96] est un exemple de cette famille de méthodes. Cette métrique utilise un vecteur à 11 éléments basé sur la qualité des minuties extraites, et un processus d'apprentissage par réseaux de neurones pour prédire la classe de qualité (1 : excellente, . . . , 5 : pauvre) pour une image d'empreinte digitale.

Krichen *et al.* [97] ont proposé une métrique pour mesurer la qualité des images Iris basée sur l'utilisation d'un modèle de mélange gaussien (*Gaussian Mixture Model*, GMM). Les auteurs ont comparé l'efficacité de leur métrique avec d'autres de la littérature selon deux types d'altérations (occultation et flou) ayant un impact sur les systèmes de reconnaissance par Iris. D'autres méthodes liées à la qualité des images Iris sont présentées dans [98, 99]. Cependant, ces méthodes sont destinées pour mesurer la qualité d'Iris, et ainsi ne peuvent pas être utilisées pour d'autres types de modalités (comme les veines de la main).

He *et al.* [100] ont proposé un modèle hiérarchique pour calculer la qualité de l'échantillon biométrique à trois niveaux : base de données (q_1), classe (q_2) et image (q_3). La méthode proposée est basée sur les quantiles de la distribution de scores des utilisateurs légitimes et d'imposteurs. La méthode proposée est efficace puisqu'elle est basée sur la séparation de la distribution de scores des utilisateurs légitimes et d'imposteurs. Cependant, la méthode requière un minimum d'images pour chaque classe (*i.e.*, individu), ce qui limite son utilisation dans la pratique.

Zhang et Wang [101] ont présenté une méthode basée sur l'hypothèse d'asymétrie du visage. La méthode utilise le descripteur SIFT pour quantifier la qualité d'une image. La méthode propose trois métriques de qualité : q_1 mesure le rapport des points SIFT détectés sur les deux côtés du visage, q_2 et q_3 ajoutent les critères de localisation et des descripteurs SIFT sur les points d'intérêt détectés, respectivement. La méthode présentée a démontré sa robustesse face aux variations d'éclairage et de pose. D'autres méthodes basées sur l'asymétrie sont présentées dans [102, 103]. Cependant, ces méthodes ne peuvent pas être utilisées pour les autres types de modalité (comme l'empreinte digitale).

2.3.4 Discussion

La section 2.3.3 montre que la plupart des travaux existants sont dépendants de la modalité utilisée et du système de vérification. D'autres méthodes basées sur la distribution de scores des utilisateurs légitimes et d'imposteurs existent et requièrent un minimum d'images pour chaque classe, ce qui limite leurs utilisations d'une manière directe sur les images acquises. Nous proposons ainsi dans le chapitre 3 une nouvelle approche pour quantifier la qualité de données biométriques morphologiques. Cette approche est basée sur l'utilisation conjointe de deux types d'informations : 1) la qualité de l'image, et 2) la qualité des paramètres extraits en utilisant le descripteur SIFT [42]. L'approche proposée possède l'avantage plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé.

2.4 Usage des systèmes biométriques

Les systèmes biométriques nécessitent l'acquisition de données biométriques de personnes, ceci à l'aide d'un capteur. Cependant, l'interaction homme-machine n'est pas toujours intuitive surtout pour les utilisateurs inexpérimentés. Kukula et Proctor [9] suggèrent que le développement des systèmes biométriques doit prendre en compte

la manière dont les individus interagissent avec le capteur. Les auteurs arguent que l'absence de cette étude entraînera une dégradation de la performance des systèmes tels que : le taux d'échec à l'enrôlement (FTE), le taux d'échec à l'acquisition (FTA) et le taux de faux rejets (FRR). Jain *et al.* [8] ont illustré la complexité de la conception d'un système biométrique selon trois axes comme le montre la figure 2.14 : la performance en terme d'erreurs (FTE, FTA, EER, *etc.*), l'usage en terme d'acceptabilité, et la sécurité en terme de robustesse du système contre la fraude. De nos jours, certaines applications requièrent qu'un système biométrique soit fonctionnel dans les cas extrêmes d'un seul de ces trois axes. Cependant, pour qu'un système soit opérationnel et acceptable, il est nécessaire que le système en question prenne en compte simultanément ces axes. Pour toutes ces raisons, l'analyse de la perception des usagers lors de l'utilisation des systèmes biométriques constitue un enjeu majeur en biométrie.

L'évaluation de l'usage des systèmes biométriques permet de réduire la complexité d'utilisation de ces systèmes. Ceci permet d'une part, d'augmenter l'acceptabilité des usagers surtout dans les applications destinées au grand public. D'autre part, cela va permettre d'améliorer leur performance en terme d'erreurs (FTA, FTE, EER) [10]. Une telle évaluation permet également de quantifier l'opérationnalité d'un système dans un contexte d'utilisation spécifique et une cible utilisateurs. Le taux EER utilisé traditionnellement pour évaluer et comparer les systèmes biométriques est généralement calculé sur des données déjà collectées sans prendre en considération les altérations possibles lors de l'acquisition de données biométriques requises, ce qui réduit la précision de cette métrique. Pour cela, nous avons besoin de connaître leur perception sur le système en termes de confiance, d'acceptabilité, *etc.*

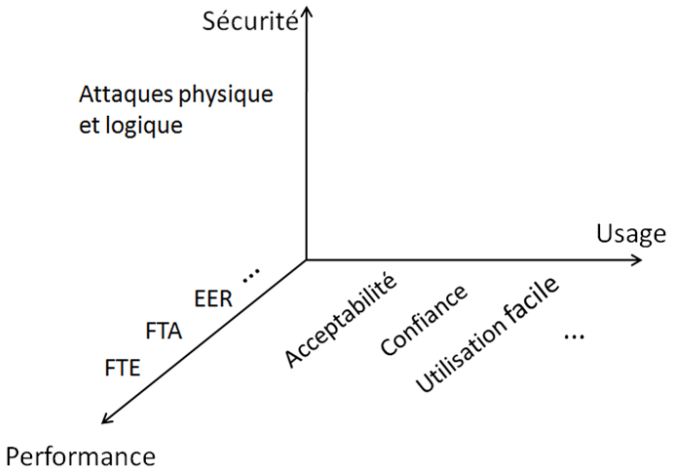


FIG. 2.14 – Conception d'un système biométrique : performance, usage et sécurité (extrait de [8]).

2.4.1 État de l'art

L'objectif de cette section est de présenter un aperçu de littérature portant sur l'usage des systèmes biométriques. Les travaux existants sont généralement basés sur une analyse statistique d'un questionnaire qui regroupe deux types d'informations. La première concerne les **questions socio-démographiques** consistant à collecter des informations sur l'individu tels que le genre, l'âge, la profession, *etc.* La seconde regroupe les **questions de perception** qui mesurent l'expérience des usagers sur la biométrie, leur perception et satisfaction lors de l'utilisation d'un système biométrique. Nous présentons dans la suite, les études existantes dans la littérature sur l'usage des systèmes biométriques :

- Deane *et al.* [104] ont mesuré l'acceptabilité des systèmes biométriques basés sur l'analyse morphologique (empreintes digitales, géométrie de la main et la rétine) et comportementales (signature dynamique, voix et dynamique de frappe au clavier). Cette étude, sur 76 volontaires, a montré que tous les systèmes biométriques ont été perçus comme moins acceptables que l'approche traditionnelle basée sur un secret. Et, contrairement aux attentes, les systèmes basés sur l'analyse comportementale ont été perçus comme étant moins acceptables que ceux basés sur l'analyse morphologique ;

- Le groupe de recherche *Opinion Research Corporation International (ORC)* [105] a effectué deux enquêtes téléphoniques, en septembre 2001 (1017 adultes) et en août 2002 (1046 adultes), auprès d'habitants des États-Unis. L'étude en 2001 a montré que 77% des personnes estiment que l'utilisation d'empreintes digitales protègent les individus contre la fraude. Pour l'atteinte à la vie privée, 87% en 2001 et 88% en 2002 ont exprimé des inquiétudes concernant la mauvaise utilisation de leurs données personnelles. L'enquête a montré également qu'il y a une forte acceptation pour que les autorités gouvernementales utilisent les empreintes digitales comme un moyen d'authentification : 88% pour les passeports, 84% pour l'accès aux bâtiments gouvernementaux et 82% dans les aéroports ;

- L'expérience du groupe de recherche *NCR Financial Solutions Division* [106, 107, 108] sur la biométrie a démontré qu'il y a des inquiétudes concernant le stockage et la mauvaise utilisation de données personnelles. Les auteurs trouvent que l'acceptation est liée au nombre d'utilisations de la biométrie en général. Les informations fournies par le dispositif biométrique peuvent aussi

améliorer l'acceptation ;

– Le groupe de recherche NIST a effectué deux études [109, 110] sur l'usage des systèmes de reconnaissance par empreintes digitales :

  • la première [109] a été effectuée sur 300 adultes (151 femmes et 149 hommes) recrutés à partir de 10.000 personnes. Parmi ces participants, 84% n'ont pas exprimé d'inquiétudes concernant le stockage de leurs empreintes digitales et 77% étaient d'accord pour fournir ces empreintes comme moyen d'identification pour les passeports ;

  • la seconde [110] a examiné l'impact de la hauteur (99 cm, 114, 3 cm et 124, 5 cm) et l'inclinaison du capteur (plat, 10 °, 20 ° et 30 °) sur la performance et l'usage du système. Cette étude a été menée sur 126 volontaires et a montré que l'inclinaison du capteur n'avait pas un impact sur la qualité de données biométriques acquises, ni sur le temps pris par cette tâche. Cependant, la hauteur du capteur a un impact sur la qualité de données acquises. Cette étude a montré également qu'avec l'augmentation de la hauteur, les participants préféraient les deux inclinaisons (20 ° et 30°) ;

– Pons et Polak [111] ont effectué une enquête sur 86 étudiants universitaires d'une grande université du sud des États-Unis. Le sondage comportait deux parties. La première visait à mesurer le degré de familiarité des participants avec la technologie biométrique. La seconde avait pour objectif de mesurer la perception des participants en termes d'avantages et d'utilisation de systèmes d'authentification biométrique. Les résultats de ce sondage ont signalé des inquiétudes concernant la vie privée. Ces résultats ont également montré que leur volonté de fournir leurs données biométriques était faible ;

– D'autres études [112, 113, 114, 115, 116] ont souligné plusieurs inquiétudes sur l'utilisation de la biométrie telles que :

  • l'utilisation d'un système biométrique nécessite le stockage de données biométriques sur un support pour qu'elles puissent être utilisées ultérieurement pendant la phase de vérification ou d'identification. En raison de ce stockage, les usagers ont exprimé des inquiétudes concernant l'atteinte à leur vie privée et la mauvaise utilisation de leurs données personnelles ;

  • des inquiétudes physiques ont été soulignées lors de l'interaction avec le capteur biométrique. Certains capteurs sont considérés comme intrusifs.

Prenons le cas de la reconnaissance par l'iris, certains usagers ont exprimé des inquiétudes concernant leur vision au cours du temps ;

- d'autres inquiétudes ont été également soulignées sur la propreté des capteurs, la facilité d'utilisation, *etc.*

2.4.2 Discussion

Les travaux présentés dans la section 2.4.1 soulignent plusieurs points importants. Premièrement, ces travaux montrent l'intérêt de la biométrie pour différents cas d'usage comme le cas d'opérations de contrôle aux frontières, avec notamment l'utilisation de passeports biométriques. Deuxièmement, certaines modalités sont considérées comme intrusives (comme le cas de l'iris, l'ADN, *etc.*). L'ADN est considéré parmi les méthodes biométriques les plus fiables. Cependant, cette méthode n'est pas utilisée pour du contrôle d'accès logique ou physique car peu acceptable par les usagers. Troisièmement, le stockage des données biométriques présente des risques en termes de respect des droits et des libertés fondamentales.

Les études de la littérature portant sur l'usage de la biométrie sont beaucoup moins nombreuses que celles liées à la performance. Ces études sont généralement basées sur des statistiques de réponses à un questionnaire, mais aucune analyse de données n'est réalisée pour comprendre les raisons associées. Comprendre les réponses permettrait de les prendre en compte dans la conception de nouveaux systèmes biométriques. De plus, certaines études sont dépendantes de la modalité utilisée [109, 110, 115], et ainsi ne peuvent pas être exploitables pour d'autres types de modalités. Afin d'apporter une solution à cette problématique, nous proposons dans le chapitre 4 une méthode générique (*i.e.*, indépendante de la modalité) pour évaluer l'usage des systèmes biométriques. La méthode proposée est basée sur un questionnaire et utilise le test de Kruskal-Wallis (KW), les réseaux bayésiens et les arbres de décision pour expliquer les réponses des usagers. Enfin, de notre point de vue, l'absence d'une méthodologie d'évaluation qui prend en compte la perception des utilisateurs limite l'utilisation des systèmes biométriques. Dans la prochaine section, nous nous intéressons à l'évaluation de la sécurité des systèmes biométriques.

2.5 Sécurité des systèmes biométriques

Les systèmes biométriques sont de plus en plus utilisés dans de nombreuses applications, pour améliorer la sécurité d'accès à des ressources physiques et logiques.

Malgré les avantages de ces systèmes par rapport aux systèmes d'authentification traditionnels, ils sont toujours vulnérables à des attaques spécifiques qui peuvent dégrader considérablement leur fonctionnalité. Dans la suite, nous présentons les travaux de la littérature qui montrent la vulnérabilité des systèmes biométriques, ainsi que les travaux existants traitant cette problématique.

2.5.1 Points de compromission d'un système biométrique

Ratha *et al.* [117] ont regroupé les attaques sur un système biométrique générique en 8 classes. La figure 2.15 (version simplifiée de la figure 1.4) illustre les emplacements possibles de ces attaques dans un système biométrique générique :

Classe 1
Données biométriques falsifiées : une reproduction de la donnée biométrique utilisée sera présentée au capteur biométrique (comme la présentation d'une copie d'une signature) ;

Classe 2
Transmission de données biométriques interceptées : une ancienne donnée biométrique enregistrée est rejouée dans le système sans passer par le capteur biométrique (comme la présentation d'une ancienne copie de l'image de l'empreinte) ;

Classe 3
Attaque sur le module d'extraction de paramètres : ce module pourrait être remplacé par un cheval de Troie de manière à produire des informations choisies par l'attaquant ;

Classe 4
Altération de paramètres extraits : après l'obtention de données par le module d'extraction de paramètres, ceux-ci sont altérés voire remplacés par d'autres données définies par l'attaquant ;

Classe 5
Le module de calcul de similarité est remplacé par un module malveillant : ce module pourrait être remplacé par un cheval de Troie afin de produire artificiellement de hauts ou bas scores ;

Classe 6
Altération de la base de données : la base de modèles biométriques est disponible localement, à distance ou distribuée sur plusieurs serveurs. Dans ce type d'attaque, l'attaquant modifie un ou plusieurs modèles afin d'autoriser un imposteur voire d'empêcher un utilisateur légitime d'y accéder ;

Classe 7

Attaque sur le canal entre la base de données et le module de calcul de similarité : dans ce type d'attaque, les modèles sont altérés sur le lien de transmission reliant la base de modèles et le module de calcul de similarité ;

Classe 8

Altération des décisions (accepté ou rejeté) : ce type d'attaque altère la décision booléenne (oui ou non) pris par le module de calcul de similarité. La dangerosité de cette attaque est élevée puisque même si le système est robuste en terme de performance, il a été rendu inutile par ce type d'attaque.

Les menaces relatives de ces attaques reposent généralement sur plusieurs facteurs que sont la modalité biométrique (il est plus difficile de reproduire la rétine que de forger une signature), le type du capteur (2D ou 3D, les capteurs 3D permettent de mieux détecter les tentatives de fraudes) et les paramètres de sécurité (illustrés par le FAR) du système.

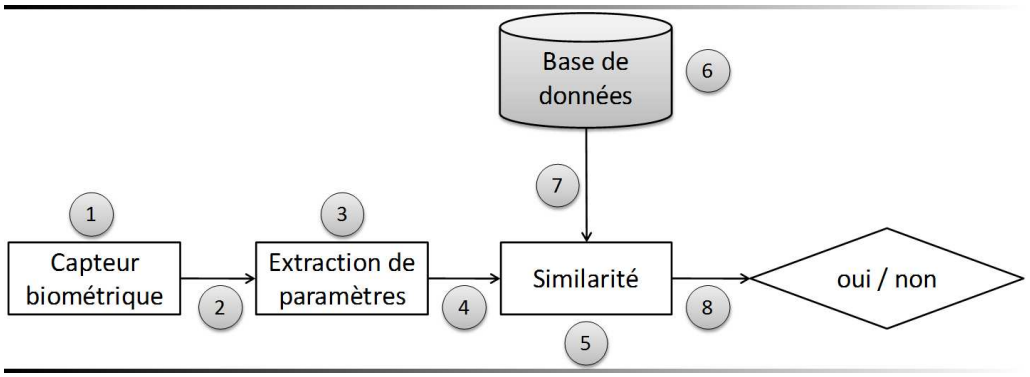


FIG. 2.15 – Emplacements des points de compromission d'un système biométrique (extrait de [117]).

2.5.2 Les différentes attaques

Schneier [118] présente deux inconvénients majeurs des systèmes biométriques que sont l'absence de secret et le problème d'irrévocabilité. Pour le premier inconvénient, tout le monde connaît nos traits biométriques (comme le visage). Tandis que pour la seconde, le trait biométrique ne peut pas être remplacé s'il est compromis.

Maltoni *et al.* [11] ont décrit les menaces possibles sur une **application générique** protégée par un système d'authentification biométrique. Une attaque par *déni de*

service (DoS), rend le système inaccessible de telle sorte que les utilisateurs légitimes ne puissent plus l'utiliser pour s'authentifier. Dans une attaque par *circonvention*, un attaquant accède à des ressources protégées de l'application (comme l'accès à des données personnelles d'un autre utilisateur). Dans une attaque par *répudiation*, un attaquant nie l'accès au système. Par exemple, un employé dans une banque qui a pu modifier illégalement certains enregistrement financiers, peut affirmer que le taux de fausses acceptations (FAR) du système a permis une autre personne d'accéder au compte de la victime. Dans une attaque par *contamination*, un attaquant obtient subrepticement des données biométriques des utilisateurs légitimes (comme le cas de reconnaissance faciale, récupération des images par des caméras cachées) pour les utiliser pour accéder à l'application. Dans une attaque par *collusion*, un utilisateur légitime avec de hauts privilèges (comme l'administrateur du système) est l'attaquant qui, d'une manière illégale, modifie le système. Dans une attaque par *coercition*, les attaquants forcent les utilisateurs légitimes à accéder au système (comme l'utilisation d'empreintes pour accéder à un guichet automatique bancaire sous la menace des armes).

2.5.3 État de l'art sur la sécurité

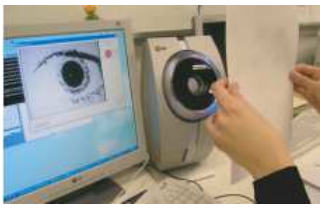
Dans cette section, nous présentons quelques études qui montrent la vulnérabilité des systèmes biométriques. Ensuite, nous présentons un aperçu des méthodes dans la littérature destinées à évaluer la sécurité des systèmes biométriques.

Les attaques de classe 1 [119, 120, 121] ont été identifiées comme des méthodes efficaces pour frauder les systèmes biométriques. Il s'agit principalement de présenter de fausses données biométriques aux capteurs comme le montre la figure 2.16. Un exemple de ce type d'attaque est présenté par Ruiz-Albacete *et al.* [120]. Leur étude montre, qu'en utilisant un système biométrique commercial, des taux élevés d'attaques réussies (de $49,32\%$ jusqu'à $82,41\%$).

Des attaques de classe 2 [122], 4 [123] et 6 [124] montrent également la vulnérabilité des systèmes biométriques. Adler [122] a proposé une attaque de classe 2 sur les systèmes de reconnaissance faciale. La méthode présentée est itérative, basée uniquement sur les scores retournés par l'algorithme de similarité entre une image d'entrée et l'image victime. Une illustration de ce type d'attaque est donnée à la figure 2.17.

Les travaux présentés par Ratha *et al.* [3], Maltoni *et al.* [11], Schneier [126] ainsi que les attaques présentées ci-dessus montrent clairement la vulnérabilité des

(a) Empreinte digitale (extrait de [125])



(b) Iris (extrait de [120])

FIG. 2.16 – Deux exemples d'attaque sur le capteur biométrique.

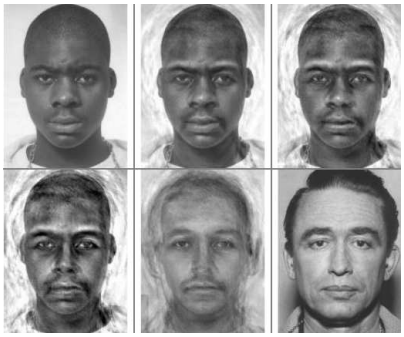


FIG. 2.17 – De gauche à droite, de haut en bas, progression des images synthétiques à différentes itérations (k=0, 200, 500, 3200) et l'image victime (extrait de [122]).

systèmes biométriques. Dès lors, l'évaluation de ces systèmes en terme de sécurité est un enjeu primordial dans ce domaine de recherche. Nous présentons dans la suite de cette section, les travaux de la littérature portant sur l'évaluation de la sécurité des systèmes biométriques.

L'arbre d'attaque introduit par Schneier [127], fournit une arborescence pour mener une analyse de sécurité des protocoles, des systèmes et des réseaux. Cependant, cette technique dépend de la modalité biométrique considérée, du système cible et de son contexte d'utilisation, ce qui la rend inexploitable dans un processus d'évaluation générique. Un exemple d'utilisation de cette approche pour analyser les attaques de classe 1 sur les systèmes de reconnaissance d'empreintes digitales est présenté dans [128].

Matyás et Ríha [129] proposent de classifier un système biométrique en quatre catégories (1 : très simple, 2 : simple, 3 : moyen, 4 : avancé). Selon la classification des auteurs, les systèmes avancés sont ceux les plus efficaces contre la fraude. La classification repose sur plusieurs critères selon que le système implémente des mécanismes de protection sur les liens de transmission, utilise un test pour détecter les fausses données biométriques, *etc.* La méthode proposée est simple à utiliser. Cependant, le

modèle de classification proposé ne pouvait pas être considéré comme discriminant pour évaluer et comparer les systèmes biométriques : selon cette classification, la plupart des systèmes existants sont classés dans la catégorie 1 et 2.

La norme de l'Organisation Internationale de Normalisation ISO/IEC FCD 19792 [48] porte sur les aspects de l'évaluation de la sécurité des systèmes biométriques. Cette norme n'avait pas pour but de définir une méthodologie pour l'évaluation de ces systèmes, mais définit les principales exigences à prendre en considération lors de l'évaluation de ces systèmes. Outre les menaces et les vulnérabilités présentées dans la section 2.5.2, la norme aborde des menaces liées à la performance du système et la qualité des données biométriques pendant la phase d'enrôlement. Un système ayant un taux de fausses acceptations (FAR) élevé, peut être facilement fraudé par la présentation de plusieurs tentatives d'un imposteur. Si les données biométriques de mauvaise qualité sont acceptées lors de la phase d'enrôlement, un attaquant peut ainsi plus facilement frauder le système en fournissant des données bruitées. D'autres facteurs ont été également signalés tels que le respect des droits et des libertés fondamentales. La norme allègue que la gestion de stockage et la politique d'accès aux modèles biométriques est un facteur primordial lors du processus d'évaluation de tels systèmes.

Dimitriadis et Polemi [130] présente une méthode d'évaluation quantitative pour la sécurité des systèmes biométriques. Ils présentent une liste de 12 vulnérabilités de ces systèmes et proposent quelques contre mesures associées. La méthode proposée calcule un facteur de risque associé à chaque vulnérabilité. Elle est facile à utiliser et efficace. Cependant, la méthode ne prend pas en considération les besoins de sécurité (confidentialité, intégrité, *etc.*), comme ils sont présents dans la méthode d'audit de sécurité EBIOS [131]. De plus, d'autres attaques et vulnérabilités doivent être prises en compte, surtout celles présentées dans la norme ISO/IEC FCD 19792 [48] qui ne sont pas prises en compte dans ces travaux, pour mieux évaluer et comparer les systèmes biométriques.

2.5.4 Discussion

La travaux de la littérature montrent que les systèmes biométriques sont vulnérables. La section 2.5.3 montre également que les travaux orientés sur l'évaluation de la sécurité des systèmes biométriques sont beaucoup moins nombreux que ceux liés à la définition de scénarios d'attaques de tels systèmes. De plus, la diversité de points de compromission et le nombre d'acteurs impliqués dans le processus biométrique, rend ces systèmes vulnérables. Pour toutes ces raisons, nous voyons que l'évaluation

des systèmes biométriques en terme de sécurité est une phase importante lors de la conception et l'évaluation de tels systèmes. Nous proposons ainsi dans le chapitre 5 : 1) une base commune d'attaques et de vulnérabilités des systèmes biométriques, et 2) une méthode générique (*i.e.*, indépendante de la modalité) pour évaluer quantitativement les systèmes biométriques en terme de sécurité. La méthode proposée est inspirée des Critères Communs [132] et de la méthode d'audit de sécurité EBIOS [131].

2.6 Conclusion

Dans ce chapitre, nous avons présenté les méthodes d'évaluation existantes des systèmes biométriques. Ces méthodes sont généralement réalisées sous quatre aspects que sont la performance en terme de taux d'erreur, la qualité des données acquises, l'usage en termes d'acceptabilité et de satisfaction et la sécurité en terme de robustesse contre la fraude.

Nous avons vu qu'il existe un grand nombre de métriques statistiques, de compétitions et de plateformes qui ont été mises en place pour évaluer la performance des systèmes biométriques. Cependant, le processus biométrique est réalisé par une interaction homme-machine. L'état de l'art montre que l'interaction de l'individu avec le capteur d'acquisition a un impact majeur sur la performance globale du système. Afin qu'un système soit acceptable et opérationnel, étudier ainsi cette interaction s'avère indispensable. Nous voyons que le manque de prise en compte du point de vue des usagers lors de la conception et du développement d'un système biométrique limite son utilisation.

Nous avons également vu que les systèmes biométriques sont vulnérables aux attaques spécifiques qui peuvent dégrader considérablement leur fonctionnalité et utilité. Ces systèmes présentent également des risques en termes de respect des droits et des libertés fondamentales : plusieurs inquiétudes ont été soulignées par les usagers concernant le stockage et la mauvaise utilisation de leurs données biométriques.

Ce chapitre montre également l'intérêt de la qualité pour éviter les artefacts d'acquisition et dans les approches multimodales. Malgré l'intérêt de la qualité, cet aspect d'évaluation est très peu abordé en biométrie. De plus, la plupart de ces travaux existants dépendent de la modalité utilisée et du système de vérification.

Nous pouvons voir que les travaux existants liés à la qualité, l'usage et la sécurité sont beaucoup moins abordés que ceux liés à la performance. Nous avons ainsi choisi d'axer cette thèse sur ces trois aspects complémentaires à la performance. L'objectif de cette thèse est la conception d'une méthodologie d'évaluation générique qui tient en compte simultanément les quatre aspects d'évaluation que sont la performance, la qualité, l'usage et la sécurité. Le chapitre suivant présente une méthode d'évaluation de la qualité de données biométriques morphologiques, tandis que les deux suivantes concernent l'usage et la sécurité des systèmes biométriques, respectivement.

Chapitre 3

Évaluation de la qualité des données biométriques morphologiques

Ce chapitre présente la méthode d'évaluation de la qualité de données biométriques que nous avons développée. La méthode proposée utilise deux types d'informations : la première est basée sur la qualité de l'image et l'autre sur la qualité de l'information contenue dans l'image en utilisant le descripteur Scale Invariant Feature Transform (SIFT). Cinq bases de données (quatre de visages et une d'empreintes digitales), et un système biométrique ont été utilisés pour montrer l'efficacité de la méthode proposée.

Sommaire

3.1 Introduction

Nous avons vu dans le chapitre 2 qu'il existe plusieurs facteurs qui dégradent la qualité des données biométriques acquises. Ces facteurs dépendent de la modalité utilisée et, dans la plupart de temps, ils sont non maîtrisables. Prenons par exemple le cas de la reconnaissance faciale, il est difficile de contrôler l'expression faciale de l'usager pendant la phase d'acquisition de la donnée biométrique. Ainsi,

l'évaluation de la qualité des données biométriques acquises est un facteur primordial dans le processus biométrique.

Dans le chapitre 2, nous avons également vu que les travaux existants sont dépendants de la modalité utilisée et du système de vérification. Nous présentons ainsi dans ce chapitre une nouvelle approche pour quantifier la qualité de données biométriques. Elle consiste à quantifier la qualité de l'image acquise afin d'améliorer la performance des systèmes d'authentification biométrique. L'avantage de cette méthode est qu'elle est plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé.

Dans la suite de ce chapitre, la section 3.2 présente la méthode proposée. Le protocole expérimental ainsi que les résultats expérimentaux sont donnés dans la section 3.3. Enfin, la section 3.4 conclut ce chapitre.

3.2 Méthode développée

La méthode proposée consiste à quantifier la qualité d'une donnée biométrique en utilisant deux types d'informations complémentaires (*cf.*, figure 3.1). Le principe retenu est le suivant : suite au calcul d'un critère de qualité d'image (section 3.2.1) et de plusieurs critères de qualité du descripteur (section 3.2.2), un processus de classification par apprentissage statistique est opéré à partir de l'ensemble des critères calculés (section 3.2.3).

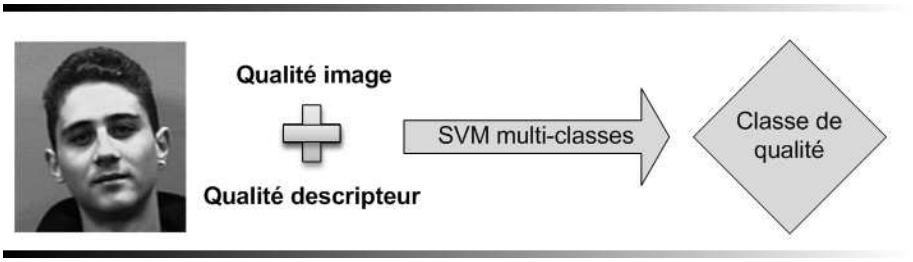


FIG. 3.1 – Principe de la méthode proposée.

3.2.1 Qualité image sans référence

L'évaluation de la qualité des images est utilisée pour valider un traitement appliqué sur des images numériques. Dans le cadre de la compression des images,

par exemple, une telle évaluation est utilisée pour quantifier la qualité de l'image reconstruite. Les métriques de qualité sont généralement classées en trois catégories : i) les métriques de qualité avec référence complète, notées FR (*Full Reference*)[133], qui comparent l'image à évaluer avec un modèle de référence de celle-ci ; ii) les métriques de qualité avec référence réduite, notées RR (*Reduced Reference*)[134], qui comparent une description de l'image à évaluer avec une description du modèle de référence ; et iii) les métriques de qualité sans référence, notées NR (*No Reference*)[135], qui quantifient la qualité de l'image à évaluer, à partir de connaissances a priori sur celles-ci (*i.e.*, sans utilisation de modèle ou de description de référence). Dans cette étude, étant donné que le signal de référence n'est pas disponible, nous avons cherché à utiliser une métrique de qualité sans référence (NR). La plupart des métriques NR existantes dépendent de l'artefact d'acquisition (effet de bloc [136], flou [137], *etc.*), ce qui limite leur utilisation en pratique. D'autres méthodes [138, 139] utilisent un algorithme d'apprentissage sur des paramètres extraits. L'efficacité de ces métriques dépend ainsi de la fiabilité et de la généralisation de ces paramètres. Dans le cadre de cette thèse, nous avons utilisé l'indice *BLind Image Integrity Notator using DCT Statistics* (BLIINDS) [140] qui ne dépend pas de l'artefact d'acquisition. Cet indice exploite la notion des statistiques de scènes naturelles. L'idée principale de cette approche repose sur l'hypothèse que les fonctions du système visuel humain ont évolué en fonction du temps et sont adaptées aux statistiques du monde dans lequel l'être humain évolue. L'indice BLIINDS est basé sur le calcul de quatre facteurs de dégradation dans le domaine de la DCT à différentes résolutions spatiales de l'image. Ces facteurs sont ensuite combinés afin de calculer la note finale de qualité. L'image est décomposée en bloc de taille 17×17. Les dégradations mesurées sont :

1. Distortion de contraste (v_1) : le contraste est une propriété intrinsèque d'une image qui désigne la différence entre les zones claires et foncées d'une image. Le contraste v_1 est calculé en utilisant les valeurs de contraste local de chaque bloc. Le contraste local du k-ième bloc est donné par :

$$c^k(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{x_{AC}^i}{x_{DC}} \tag{3.1}$$

avec N est la taille du bloc, x_{DC} représente le coefficient DC et l'ensemble $\{x_{AC}^i |\ i = 1 : N\}$ représente les coefficients AC. Le contraste de l'image v_1 est ainsi calculé par :

$$v_1 = \frac{1}{M} \sum_{i=1}^{M} c^i(x) \tag{3.2}$$

avec M est le nombre de blocs de l'image en question.

2. Distortion de structure (v_2) : les caractéristiques de structure sont obtenues en utilisant le kurtosis des coefficients (non DC) de fréquences DCT, calculés sur chaque bloc. Le kurtosis du $k^{\text{ème}}$ bloc est donné par :

$$\kappa^k(x_{AC}) = \frac{E(x_{AC} - \mu)^4}{\sigma} \tag{3.3}$$

avec μ est la moyenne des coefficients AC, et σ son écart-type. La mesure de distortion de structure v_2 est ainsi calculée par la moyenne des valeurs au dessous du $10^{\text{ème}}$ percentile.

3. Anisotropie d'orientation (v_3 *et* v_4) : les auteurs dans [141] montrent que la dégradation a un impact sur l'information directionnelle d'une scène. Par conséquent, l'anisotropie (qui dépend de l'information directionnelle d'une scène) est calculée en utilisant l'entropie de Rényi (qui est une généralisation de l'entropie de Shannon) sur les blocs DCT selon quatre orientations différentes $\theta = 0, 45, 90, 135$ en degrés. Les deux mesures v_3 et v_4 sont calculées comme suit : les coefficients DCT du $k^{\text{ème}}$ bloc autour de l'orientation θ sont notés par $P_\theta[k, j]$, avec j est l'indice du coefficient DCT. Chaque coefficient du bloc DCT est ensuite normalisé par :

$$\tilde{P}_\theta[k, j] = \frac{P_\theta[k, j]^2}{\sum_{j=1}^{N} P_\theta[k, j]^2} \tag{3.4}$$

avec N la taille du $k^{\text{ème}}$ bloc orienté et son entropie de Rényi R_θ^k est défini par :

$$R_\theta^k = \frac{1}{1-\beta} \log_2 \left(\sum_{j=1}^{N} \tilde{P}_\theta[k, j]^\beta \right) \tag{3.5}$$

où $\beta > 1$. Enfin, les deux mesures basées sur l'anisotropie sont définies par :

$$v_3 = \text{var}(E(R_\theta^k)) \text{ et } v_4 = \max(E(R_\theta^k)), \forall k, \forall \theta \tag{3.6}$$

Dans cette étude, on fixera β à 3. Étant donné que la perception visuelle de l'image dépend de la résolution de l'image, la distance entre le plan de l'image et l'observateur, et l'acuité des observateurs, une approche multi-échelles est appliquée afin de calculer un score global :

$$\text{BLIINDS} = \prod_{i=1}^{L} v_1^{\alpha_1^i} v_2^{\alpha_2^i} v_3^{\alpha_3^i} v_4^{\alpha_4^i} \tag{3.7}$$

avec $\sum_{j=1}^{4}\sum_{i=1}^{L}\alpha_j^i = 1$ et L représente le nombre de niveaux de décomposition utilisé. Les valeurs α_j^i ont été obtenues en calculant la corrélation de chacun des critères (v_i) avec les notes de qualité fournies par les observateurs humains [140]. Des exemples de cette métrique sont donnés à la figure 3.2.

13.58 11.15 9.35 8.50

FIG. 3.2 – Exemples de la métrique BLIINDS sur des images de la base de données *FACES94*. De gauche à droite, image de référence ensuite images altérées par un bruit gaussien.

3.2.2 Qualité du descripteur

La mesure de la qualité d'un descripteur est basée sur des mesures statistiques de points d'intérêt. Nous avons utilisé les points d'intérêt puisqu'ils décrivent de façon stable les régions de l'image où l'information est importante. Cette approche est généralement utilisée pour reconnaître des objets [142] et dans les algorithmes de reconnaissance biométrique [43]. Pour le calcul du vecteur descripteur au voisinage des points détectés, il existe de nombreuses méthodes tels que *Scale Invariant Feature Transform* (SIFT) [42], *Shape Contexts* [143], *Speed Up Robust Features* (SURF) [144]. Parmi ces algorithmes, l'algorithme SIFT proposé par Lowe [42] est retenu pour deux raisons principales. Premièrement, l'algorithme SIFT est efficace au changement d'échelle et à la rotation 2D. Deuxièmement, une étude comparative [145] de différents descripteurs montre que SIFT est le plus performant. L'algorithme SIFT a été également utilisé par Berretti *et al.* [146] dans le cas de reconnaissance faciale 3D.

L'algorithme SIFT, publié par Lowe en 1999 [42], est un algorithme de traitement d'images qui permet de détecter et de décrire les caractéristiques d'une image. Elle permet de transformer une image en un ensemble de caractéristiques, chacun étant invariant aux transformations suivantes : translation de l'image, changement d'échelle (*c.-à-d.*, redimensionnement), rotation et partiellement invariant aux changement d'éclairage. La détection des points d'intérêt présente dans l'algorithme SIFT se fait dans l'espace des échelles. Les emplacements des points d'intérêt sont définis comme

étant l'extremum du résultat de la différence de gaussiennes DoG appliqué dans l'échelle spatiale sur une série d'images lissées et rééchantillonnées. L'espace d'échelle d'une image est défini par une fonction $L(x, y, \sigma)$ produite grâce à la convolution entre la fonction gaussienne $G(x, y, \sigma)$ avec l'image d'entrée $I(x, y)$:

$$L(x, y, \sigma) \ = \ G(x, y, \sigma) * I(x, y) \tag{3.8}$$

avec $*$ le produit de convolution, et G la fonction gaussienne suivante :

$$G(x, y, \sigma) \ = \ \frac{1}{2\pi\sigma^2} \mathrm{e}^{\frac{x^2+y^2}{2\sigma^2}} \tag{3.9}$$

Ainsi, pour extraire des points d'intérêt stables dans l'espace échelle, Lowe a proposé d'utiliser l'extremum de la fonction de différences gaussiennes, qui peut être calculée à partir de la différence de deux échelles voisines séparées d'une constante multiplicative k. La fonction de différences gaussiennes DoG est donnée par :

$$D(x, y, \sigma) \ = \ L(x, y, k\sigma) - L(x, y, \sigma) \tag{3.10}$$

Cette première étape permet d'obtenir un grand nombre de points d'intérêt. Selon la méthode SIFT, une phase de filtrage de ces points est effectuée afin d'en extraire les plus intéressants. Afin qu'un point soit sélectionné, il doit satisfaire deux conditions : la première concerne le contraste du point d'intérêt tandis que la seconde concerne son rayon de courbure. Pour la première condition, il faut que la valeur de $D(x, y, \sigma)$ soit supérieure à un seuil prédéfini (fixée à $0,03$ par Lowe). Pour la seconde condition, on s'intéresse à la matrice hessienne de D, tout comme pour le critère de sélection de Harris et Stephen [147] :

$$M^L = \begin{bmatrix} \frac{\partial D^2}{\partial x} & \frac{\partial D}{\partial x}\frac{\partial D}{\partial y} \\ \frac{\partial D}{\partial x}\frac{\partial D}{\partial y} & \frac{\partial D^2}{\partial y} \end{bmatrix}$$

À partir de cette matrice, le critère de sélection de Lowe est le suivant :

$$R^L \ = \ \frac{Tr(M^H S)}{Det(M^L)} \ < \ \frac{(r+1)^2}{r} \tag{3.11}$$

avec

$$M^L = \mathrm{e}^{\frac{x^2+y^2}{2\sigma^2}} * \begin{bmatrix} \frac{\partial I^2}{\partial x} & \frac{\partial I}{\partial x}\frac{\partial I}{\partial y} \\ \frac{\partial I}{\partial x}\frac{\partial I}{\partial y} & \frac{\partial I^2}{\partial y} \end{bmatrix}$$

avec $Det(\cdot)$ le déterminant, $Tr(\cdot)$ la trace et r un paramètre du critère. Une valeur élevée du paramètre r permet de s'assurer que le point considéré est un point d'intérêt.

Un point d'intérêt est défini par 5 paramètres (x, y, σ, θ, v). Le couple (x, y) correspond à sa position dans l'image originale. Le couple (σ, θ) décrit son échelle et orientation. Et le vecteur v est son vecteur de descripteurs qui est calculé en utilisant son voisinage. Le voisinage est divisé par une grille 4×4. Ensuite, le gradient est calculé sur chacune des 16 localisations de la grille puis est quantifié selon un histogramme à 8 orientations. La concaténation de ces éléments nous permet d'obtenir un vecteur de descripteurs à 128 éléments. La figure 3.3 illustre la construction d'un descripteur SIFT d'un point d'intérêt. La figure 3.4 présente des exemples de détection de points d'intérêt par cet algorithme.

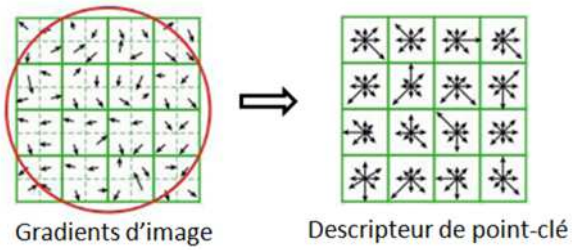


FIG. 3.3 – Construction d'un descripteur SIFT (source [42]).

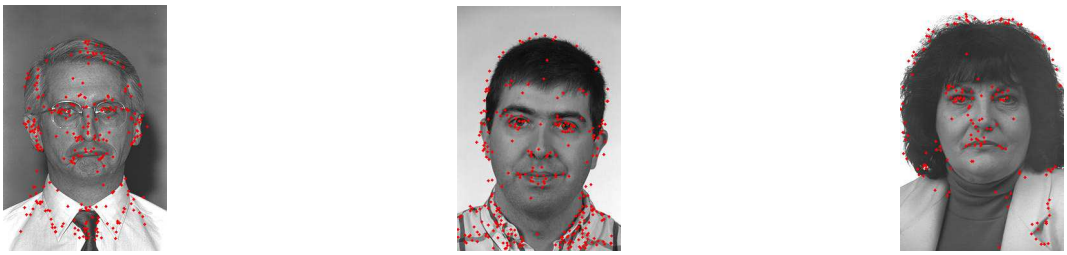


FIG. 3.4 – Exemples de détection de points d'intérêt.

En utilisant SIFT, l'image I est ainsi caractérisée par l'ensemble $Y(I) = \{k_i = (x_i, y_i, \sigma_i, \theta_i, v_i) \mid i = 1 : N(I)\}$ avec $N(I)$ le nombre de points d'intérêt détectés dans I ; (x_i, y_i) la position du point d'intérêt i dans I ; (σ_i, θ_i) l'échelle et l'orientation du point d'intérêt i ; et v_i le vecteur (à 128 éléments) de descripteurs du point d'intérêt i. À partir de ces caractéristiques, nous avons choisi d'utiliser quatre critères qui nous ont paru pertinents (*cf.*, section 3.3.2.1) pour prédire la qualité du descripteur : 1) le nombre de points d'intérêt détectés dans l'image I ; 2) le coefficient DC de la

matrice MAT, avec $N(I)$ lignes et 128 colonnes, contenant les vecteurs descripteurs des points d'intérêt détectés dans I ; 3) la moyenne et 4) l'écart-type du vecteur contenant l'échelle de chaque point d'intérêt détecté dans I.

Finalement, nous disposons de cinq critères (un dédié à la qualité de l'image et quatre sur la qualité du descripteur) pour établir le niveau de qualité de données biométriques. Au lieu de faire une opération arithmétique des valeurs ainsi obtenues, nous proposons d'utiliser un algorithme de classification de la qualité à 10 classes :

- Classe 1 correspond à une image de référence (c'est à dire non altérée) ;

- Classes 2 jusqu'à 10 correspondent à 3 types d'altérations et 3 niveaux pour chaque type d'altération, respectivement. Une description détaillée des altérations introduites est donnée à la section 3.3.1.1.

3.2.3 Machines à Vecteurs de Support (SVM)

La méthode de machine à vecteurs de support ou Séparateurs à Vastes Marges est une méthode de classification par apprentissage supervisé développée par Vapnik [148]. Elle est connue sous le terme anglais par *Support Vectors Machine* (SVM). Le but des SVM est de classifier un objet x à l'aide d'une marge maximale associée à des vecteurs de supports et d'une fonction noyau. Cette méthode est devenue populaire du fait de ces performances à traiter des données de grande dimension. La fonction noyau permet d'opérer un changement de repère dans un espace de plus grande dimension afin de retrouver un problème de séparation linéaire, lorsque les données ne sont pas linéairement séparables. Soit une base d'apprentissage $S_{apprentissage}$: $S_{apprentissage} = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ composée de m couples (vecteur d'attributs, classe) avec $x_i \in \mathbb{R}^n$ et $y_i \in \{-1, 1\}$. L'algorithme SVM projette les valeurs x_i dans un espace de travail $\mathcal{H}$ ($\phi : \mathbb{R}^n \to \mathcal{H}$). L'hyperplan optimal de séparation des deux classes dans l'espace $\mathcal{H}$ est ensuite recherché. Cet hyperplan (w, b) matérialise la frontière de séparation entre les deux classes. La classe y d'un nouvel exemple x est définie par :

$$y = \langle w, \Phi(\mathbf{x}) \rangle + b = \sum_{i=1}^{\ell} \alpha_i^* y_i K(\mathbf{x}_i, \mathbf{x}) + b \tag{3.12}$$

avec $\alpha_i^* \in R$ et $K(\cdot, \cdot)$ est la fonction noyau. Dans l'algorithme SVM, l'hyperplan est optimal s'il maximise la distance qui le sépare des exemples dont il est le plus

proche. Cette distance est appelée marge du classifieur. Les α_i^* qui maximisent le critère d'optimalité sont obtenus en résolvant :

$$\max_{\alpha_i} \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i K(\mathbf{x}_i, \mathbf{x}_j y_j) \tag{3.13}$$

sous les contraintes, $0 \leq \alpha_i \leq C$ et $\sum_{i=1}^{\ell} \alpha_i y_i = 0$, avec C est le coefficient de pénalisation. L'algorithme SVM de base a été développé pour les problèmes de classification à deux classes. Cependant, plusieurs approches peuvent être utilisées pour l'étendre aux problèmes multi-classes. Dans nos travaux, nous avons utilisé l'approche *un contre un* avec le critère de vote majoritaire pour la sélection de la classe finale. Nous avons utilisé un script python (*easy.py*) fourni par la librairie libsvm [149]. Une recherche exhaustive (*grid-search*) est effectuée pour la recherche des deux paramètres optimums C et γ, et le noyau utilisé est le noyau RBF défini par :

$$k(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2) \tag{3.14}$$

3.3 Validation

3.3.1 Protocole expérimental

Afin d'évaluer les performances de la méthode proposée, cinq bases de données (quatre de visages et une d'empreintes digitales) et un système de vérification ont été utilisés. Dans la section 3.3.1.1, nous présentons les bases de données utilisées, les artefacts d'acquisition introduits ainsi que le système de vérification utilisé. Le processus d'évaluation est ensuite donné à la section 3.3.1.2.

3.3.1.1 Outils utilisés

Les bases de référence utilisées sont celles présentées dans le chapitre 2 section 2.2.2.1 : *FACES94*, *ENSIB*, *FERET*, *AR* et *FVC2002 DB*$_2$. Le système de vérification utilisé est GREYC-Face présenté dans le chapitre 1 section 1.3.3.1.

Pour chacune des bases de référence utilisées, nous avons simulé plusieurs artefacts d'acquisition (mouvement, bruit gaussien et distance d'acquisition), et appliqué trois niveaux de dégradation pour chaque type d'altération :

– Altération par mouvement (flou) : les images altérées par flou sont obtenues par un filtre gaussien 2D en utilisant la méthode MATLAB *fspecial ('gaussian', hsize, σ)* ;

– Altération par bruit gaussien : ces images sont obtenues en utilisant la méthode MATLAB *imnoise (I, 'gaussian', μ, v)* ;

– Altération par distance d'acquisition (redimensionnement) : ces images sont générées en utilisant la méthode MATLAB *imresize (I, scale, 'nearest')*.

Le tableau 3.1 présente la valeur des paramètres requis des méthodes MATLAB. La figure 3.5 illustre un exemple de ces altérations sur une image de la base de données *FACES94*.

Type d'altération	**Méthode**	**Niveau 1**	**Niveau 2**	**Niveau 3**
Flou	fspecial	hsize = [7 7] et $\sigma = 1$	hsize = [7 7] et $\sigma = 2$	hsize = [7 7] et $\sigma = 6$
Bruit gaussien	imnoise	$\mu = 0.01$ et $v = 0.003$	$\mu = 0.01$ et $v = 0.01$	$\mu = 0.01$ et $v = 0.017$
Redimensionnement	imresize	scale = 0.8	scale = 0.6	scale = 0.4

TAB. 3.1: Paramètres des méthodes d'altérations MATLAB.

(a) Altération par flou

(b) Altération par bruit gaussien

(c) Altération par redimensionnement

FIG. 3.5 – Exemple d'altérations sur une image de la base de données *FACES94*. De gauche à droite, image de référence ensuite images altérées niveau 1, 2 et 3, respectivement.

3.3.1.2 Processus de validation

Selon Grother et Tabassi [6], les méthodes de qualité doivent être en mesure de prédire la performance des systèmes biométriques. Cela signifie qu'une méthode de qualité prend en entrée une donnée biométrique, et prédit sa catégorie de qualité lié au taux d'erreur associée à cette donnée. Afin de quantifier la performance de la méthode proposée, nous procédons comme suit :

- Apprentissage des SVM multi-classes : pour les bases de visages, nous avons généré quatre SVM multi-classes (*i.e.*, un SVM multi-classes par base), et un SVM multi-classes contenant des exemples de toutes les bases (SVM_{tout}). Pour la base d'empreintes digitales, nous avons généré un autre SVM multi-classes. Pour apprendre et tester les différents SVM multi-classes, nous avons découpé chaque base d'images en deux ensembles $S_{apprentissage}$ et S_{test} d'une manière équilibrée (*i.e.*, le même nombre d'exemples par classe existe dans les deux ensembles). Le choix du noyau utilisé et les paramètres requis sont présentés dans la section 3.2.3 ;

- Définition des catégories de qualité : la méthode SVM multi-classes proposée prédit une classe de qualité pour une image en entrée. Afin de quantifier la performance de cette méthode, nous devons tout d'abord définir les catégories de qualité pour le système de vérification utilisé. Selon le système de vérification utilisé, certaines altérations peuvent avoir un impact sur sa performance globale plus que d'autres. Par la suite, l'EER est utilisé pour illustrer la performance globale d'un système biométrique ;

- Corrélation des valeurs de l'EER avec les catégories de qualité : afin de quantifier l'efficacité de la méthode proposée pour prédire les performances du système testé, nous calculons l'EER de chaque catégorie de qualité. L'intérêt de la méthode proposée est ainsi quantifié par son efficacité pour prédire les performances du système testé. En d'autres termes, plus les données biométriques sont dégradées, plus la performance globale du système est dégradée (cela se traduit par une augmentation des valeurs de l'EER).

3.3.2 Résultats

Nous présentons dans la section 3.3.2.1 l'efficacité des cinq critères de qualité retenus en fonction des altérations. La section 3.3.2.2 montre l'intérêt de la méthode

proposée selon le processus de validation présenté dans la section 3.3.1.2. Une étude comparative entre la méthode proposée et la méthode NFIQ est présentée dans la section 3.3.2.3.

3.3.2.1 Comportement des attributs en fonction des altérations

Dans cette section, nous présentons le comportement des cinq critères de qualité utilisés dans la méthode proposée avec les altérations introduites dans la section 3.3.1.1. Il s'agit de quantifier l'efficacité de chacun de ces critères pour détecter les trois types d'altérations. Pour ce faire, nous utilisons le coefficient de corrélation linéaire de Pearson (*cf.*, annexe A) entre les critères de qualité utilisés et les trois types d'altérations. Nous définissons ainsi pour chaque type d'altération et pour chaque critère p du vecteur de qualité les deux variables comme suit :

- $X_p = \{X_{pk} |\ k = 1 : 4\}$ où X_{p1} est l'ensemble des valeurs du critère p de toutes les images de référence, (X_{p2}, X_{p3}, X_{p4}) sont les ensembles des valeurs de p de toutes les images altérées niveau 1, 2 et 3, respectivement ;

- Les niveaux d'altérations sont représentés par la variable Y (1 : pour les bases de référence, 2, 3 et 4 : pour les bases altérées niveau 1, 2 et 3, respectivement). Plus spécifiquement, $Y = \{y_k | y_k = 1,\ k = 1 : N;\ y_k = 2,\ k = N+1 : 2N;\ y_k = 3,\ k = 2N+1 : 3N;\ y_k = 4,\ k = 3N+1 : 4N\}$ où N correspond à la taille des quatre bases de visages de référence.

Le tableau 3.2 montre que les quatre critères de qualité du descripteur (Nombre de points d'intérêt, Coefficient DC, Moyenne et Ecart-type d'échelles) sont pertinents pour détecter les trois types d'altérations. Le critère de qualité image BLIINDS a montré son efficacité (avec la valeur absolue d'un coefficient de corrélation supérieur à $0,6$) pour détecter les altérations par flou et bruit gaussien. Pour l'altération par redimensionnement, le tableau montre que BLIINDS ne permet pas de la détecter. Ce résultat était attendu car BLIINDS est une métrique de qualité d'image sans-référence et multi-résolutions, et que le processus de redimensionnement n'introduit aucune distorsion de qualité en l'absence de référence.

3.3.2.2 Comportement de la méthode proposée

Les performances des six SVM multi-classes générés (cinq pour les bases de visages et un pour la base d'empreintes digitales) sont données dans le tableau 3.3. Nous avons mis le symbole «×» pour la base *FVC2002 DB*$_2$, car nous avons généré un seul

Critère	Flou	Bruit gaussien	Redimensionnement
Nombre de points d'intérêt	**-0.57**	0.39	-0.49
Coefficient DC	**-0.62**	**0.57**	**-0.53**
Moyenne échelles	**0.79**	**-0.56**	-0.4
Ecart-type échelles	0.35	-0.35	-0.47
BLIINDS	**0.63**	**-0.8**	-0.1

TAB. 3.2: Coefficients de corrélation de Pearson entre les critères de qualité utilisés et les altérations sur les toutes les bases de visages. Les valeurs en gras correspondent à des corrélations fortes (en valeur absolue).

SVM multi-classes pour cette base. Le tableau 3.3 montre l'intérêt de la méthode proposée pour détecter les trois types d'altérations réelles (flou, bruit gaussien et redimensionnement) des données, avec des taux de bonne classification satisfaisants (de $82,29\%$ jusqu'à $97,73\%$ sur la base d'apprentissage, et de $81,16\%$ jusqu'à $91,1\%$ sur la base de test).

Afin de définir les catégories de qualité, nous avons testé la robustesse du système contre les altérations introduites dans la section 3.3.1.1. La figure 3.6 montre l'impact des images dégradées sur la performance globale (illustrée par les valeurs de l'EER) du système testé. Les valeurs de l'EER sont calculées en utilisant la première image de référence pour l'enrôlement, et les autres pour le test (procédé d'enrôlement unique). Cette figure montre que toutes les altérations introduites ont un impact sur la performance du système biométrique étudié (*i.e.*, courbes croissantes en fonction des dégradations). Par conséquent, nous définissons dans le tableau 3.4, les catégories de qualité retenues pour le système biométrique utilisé. La figure 3.7 illustre des résultats d'évaluation par la méthode proposée sur des images de la base *FACES94*. La figure 3.8 présente les valeurs de l'EER de chaque catégorie de qualité en utilisant les quatre SVM multi-classes (un SVM multi-classes par base), et le SVM multi-classes généré à partir des exemples de toutes les bases de visage, respectivement. La méthode proposée a montré son efficacité à prédire les performances du système testé. En d'autres termes, plus les images sont dégradées, plus la performance globale du système est dégradée (cela se traduit par une augmentation des valeurs de l'EER). À partir de la figure 3.8, nous pouvons également déduire deux points :

- Pour les bases *FACES94*, *ENSIB* et *AR* il n'y avait pas une différence significative entre les deux valeurs de l'EER de la base de référence et la base prédite (catégorie I) par la méthode proposée ;

- Pour la base *FERET*, il y avait une différence de 5.62% (figure 3.8 à gauche) et

5.64% (figure 3.8 à droite). Cette variation est due à la complexité de cette base (la base de référence contient beaucoup d'images altérées par résolution). Malgré cela, la méthode a également montré son efficacité à prédire la performance du système utilisé sur cette base.

	SVM_{chaque}		SVM_{tout}	
	$S_{apprentissage}$	S_{test}	$S_{apprentissage}$	S_{test}
FACES94	91.01	86.69	85.68	85.28
ENSIB	97.73	89.82	94.92	91.1
FERET	82.33	81.2	82.29	81.16
AR	90.08	89.08	90.7	88.92
FVC2002 DB$_2$	×	×	91.7	83.68

TAB. 3.3: Précision (en %) des modèles SVM multi-classes sur les deux ensembles d'apprentissage ($S_{apprentissage}$) et de test (S_{test}).

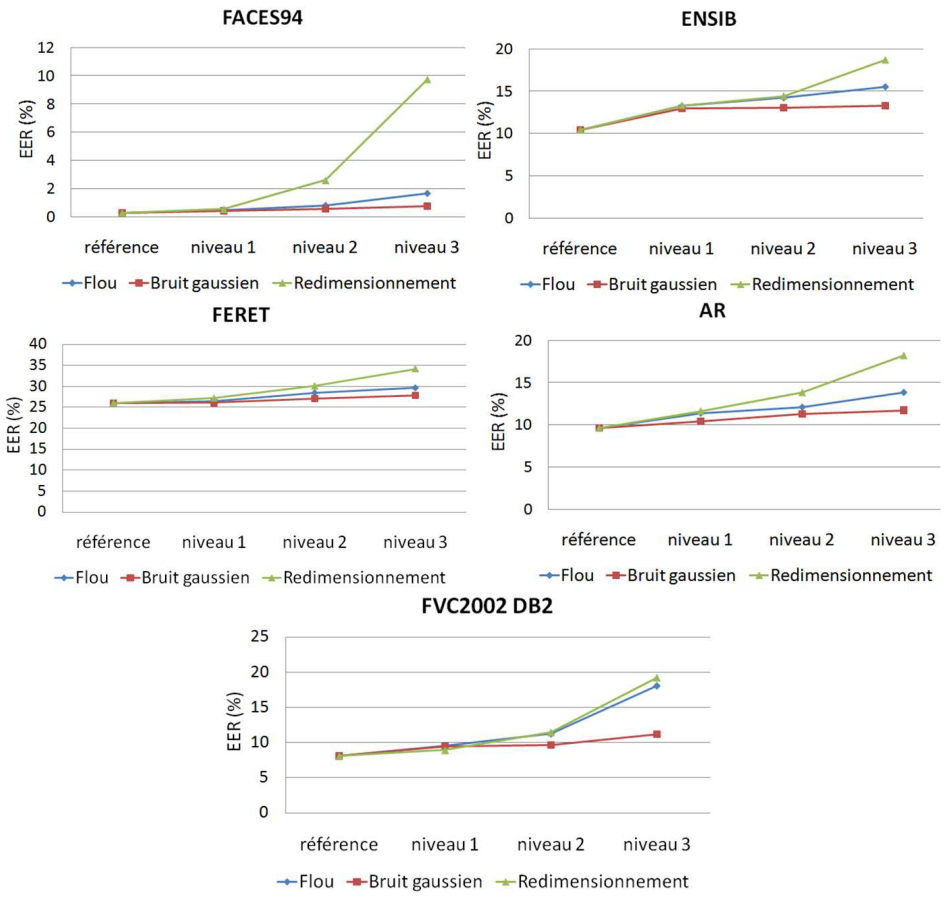


FIG. 3.6 – Impact des altérations sur la performance globale du système biométrique utilisé : valeurs de l'EER (en %) sur chaque base de données.

Catégorie de qualité	Label prédit par le SVM multi-classes	Description
I	1	Bonne
II	2, 5 et 8	Moyenne
III	3, 6 et 9	Mauvaise
IV	4, 7 et 10	Très mauvaise

TAB. 3.4: Catégories de qualité.

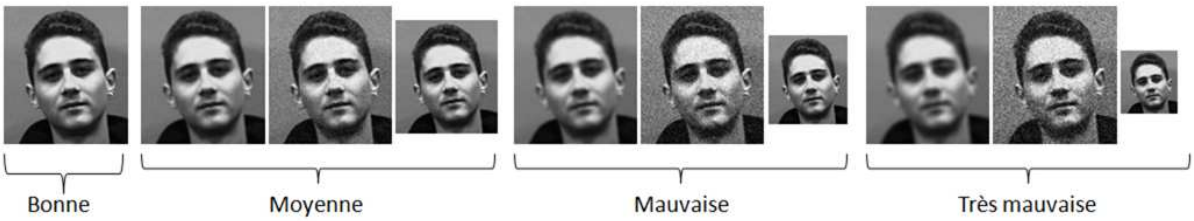


FIG. 3.7 – Exemple des résultats d'évaluation sur des images de la base *FACES94*.

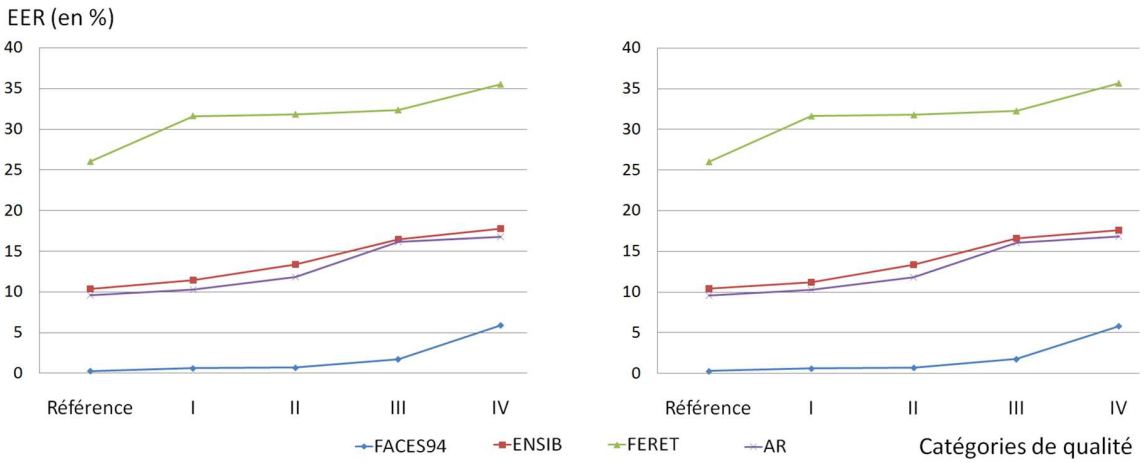


FIG. 3.8 – Les valeurs de l'EER des quatre bases de référence, et de chaque catégorie de qualité. Ces valeurs sont calculées en utilisant les quatre SVM multi-classes (à gauche) et le SVM multi-classes généré à partir des exemples de toutes les bases (à droite), respectivement.

3.3.2.3 Etude comparative entre la méthode proposée et NFIQ

Nous avons utilisé la base d'empreintes digitales *FVC2002 DB_2* (avec ses images altérées, le nombre total des images ainsi utilisées est égal à 8000) pour comparer la méthode proposée avec la métrique de qualité NFIQ [96] proposée par le NIST. Nous avons choisi NFIQ puisque cette dernière est la plus citée dans la littérature dans le cas d'empreintes digitales. Afin de comparer ces deux algorithmes de qualité, nous avons suivi la démarche suggérée par Grother et Tabassi [6]. Nous avons utilisé le test de Kolmogorov-Smirnov (KS) [150] pour mesurer le chevauchement des deux

distributions de scores des utilisateurs légitimes (scores intra) et d'imposteurs (scores inter). Ce test statistique retourne une valeur définie entre 0 et 1 : une valeur proche de 0 signifie que les deux distributions sont homogènes (*i.e.*, dépendantes), tandis qu'une valeur proche de 1 signifie que les deux distributions sont indépendantes. Ainsi, plus les images sont de bonne qualité, plus une valeur statistique KS importante (proche de 1) est attendue.

Le tableau 3.5 décrit les valeurs statistiques du test KS. La méthode proposée a montré son efficacité pour mieux séparer les deux distributions des scores intra et inter que la métrique NFIQ. Pour la catégorie IV (*i.e.*, images de très mauvaise qualité), la méthode NFIQ est légèrement plus efficace (statistique KS égale à $0,64$) que la méthode proposée (statistique KS égale à $0,626$) pour séparer la distribution des scores intra et inter. Tandis que, pour les autres trois catégories de qualité (I, II et III), la méthode proposée (statistiques KS allant de $0,797$ jusqu'à $0,869$) est nettement meilleure que la méthode NFIQ (statistiques KS allant de $0,632$ jusqu'à $0,82$).

Méthode	**Catégorie I**	**Catégorie II**	**Catégorie III**	**Catégorie IV**
Méthode proposée	0.869	0.828	0.797	0.626
NFIQ	0.82	0.698	0.632	0.64

TAB. 3.5: Comparaison entre la méthode proposée et NFIQ. Test de Kolmogorov-Smirnov (KS) pour un intervalle de confiance égal à 95%.

3.4 Conclusion

Nous avons présenté dans ce chapitre, une méthode pour prédire la qualité des données biométriques morphologiques (représentées par une image). La méthode proposée utilise deux types d'informations complémentaires : 1) la qualité de l'image, et 2) la qualité des paramètres extraits en utilisant le descripteur SIFT. L'approche proposée est plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé. Nous avons montré son intérêt pour détecter trois types d'altérations réelles (flou, bruit gaussien et redimensionnement) des données, qui ont un impact majeur sur la performance globale des systèmes biométriques. Nous avons également montré que la méthode proposée est plus efficace que la méthode de qualité NFIQ sur la base d'empreintes digitales *FVC2002 DB*$_2$. Le chapitre suivant présente une méthode générique pour évaluer l'usage des systèmes biométriques.

Chapitre 4

Évaluation de l'usage d'un système biométrique

Ce chapitre présente l'évaluation des systèmes biométriques en termes d'acceptabilité et de satisfaction. La contribution de ce chapitre est double : nous proposons 1) un questionnaire pour recueillir les caractéristiques socio-démographiques, l'expérience et l'attitude des usagers qui pourraient avoir un impact sur leur acceptabilité et leur satisfaction et 2) une méthode pour évaluer l'usage d'un système biométrique qui s'articule autour de trois étapes : la collection de données, la préparation des données collectées et l'analyse de données. La méthode proposée est générique, ce qui permet de l'utiliser pour n'importe quel type de modalité biométrique, et ne se contente pas d'une analyse descriptive mais également explicative.

Sommaire

4.1 Introduction

Analyser la perception des usagers lors de l'utilisation des systèmes biométriques a de nombreux intérêts. Nous avons vu dans le chapitre 2 que la manière dont

l'utilisateur interagit avec le capteur biométrique a un impact significatif sur la performance globale des systèmes. Nous avons également vu que certaines modalités sont considérées comme intrusives (comme l'iris), et que certaines modalités sont plus acceptées que d'autres pour des raisons sociales et/ou culturelles. Pour toutes ces raisons, étudier l'usage de la biométrie est un facteur important à prendre en compte lors de la conception et du développement des systèmes biométriques.

Dans le chapitre 2, nous avons également vu que les travaux existants sont dépendants de la modalité biométrique considérée. De plus, ces travaux sont basés sur des statistiques à des réponses à un questionnaire, mais aucune analyse de donnée n'est réalisée pour comprendre les raisons associées. Nous présentons ainsi dans ce chapitre une nouvelle approche (indépendante de la modalité) qui ne se contente pas à une analyse quantitative des réponses, mais également explicative.

Dans la suite de ce chapitre, la section 4.2 présente le principe de la méthode proposée. Ensuite, dans la section 4.3, nous illustrons la méthode proposée sur deux systèmes d'authentification biométrique. Enfin, la section 4.4 conclut ce chapitre.

4.2 Méthode développée

La méthode que nous avons développée est composée de trois étapes, comme nous pouvons le voir sur la figure 4.1 :

1. Collection de données : elle consiste à saisir des données en utilisant un questionnaire de satisfaction. Le questionnaire permet d'évaluer l'acceptabilité et la satisfaction des usagers lors de l'utilisation des systèmes biométriques ;
2. Préparation de données : préalablement à toute tâche d'extraction de connaissances, une phase de pré-traitement des données est nécessaire, et souvent prépondérante quant à la qualité des résultats ;
3. Analyse de données : une fois les réponses des usagers au questionnaire obtenues, deux types d'analyse sont réalisés. La première consiste à extraire la dépendance (s'il en existe) entre les caractéristiques socio-démographiques (comme le genre) et les questions de satisfaction. Afin qu'un système soit acceptable dans un cadre d'utilisation spécifique et une cible utilisateurs, il est nécessaire de vérifier si le genre, par exemple, affecte l'utilisation du système en question. Dans la deuxième analyse, nous nous intéressons à déterminer les dépendances entre les questions de satisfaction afin d'expliquer les réponses des usagers.

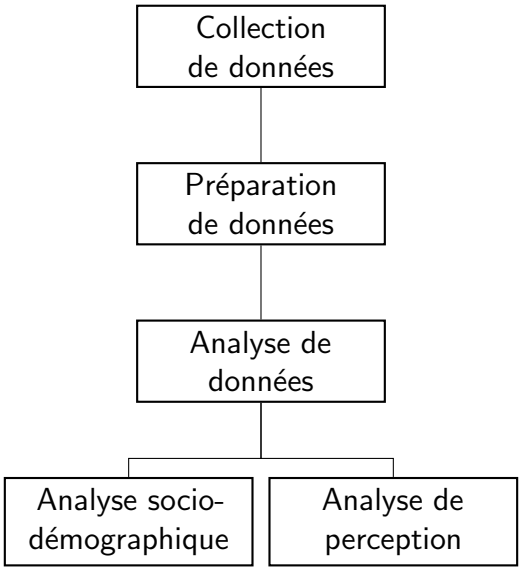


FIG. 4.1 – Principe de la méthode proposée.

4.2.1 Collection de données

La première étape de la méthode proposée consiste à créer un questionnaire de satisfaction. Dans le chapitre 2, nous avons vu qu'il existe plusieurs facteurs à extraire pour atteindre cet objectif, tels que :

– *Les facteurs socio-démographiques* tels que le genre, *etc.* ;

– *La confiance* dénote comment la performance du système biométrique est perçue par les utilisateurs. Elle dépend principalement du retour d'expérience des utilisateurs ;

– *La facilité d'utilisation* du système biométrique dépend de la qualité du capteur biométrique et de l'ergonomie de l'interface. Elle dépend aussi du temps requis pour l'identification. Si le système biométrique prend plusieurs minutes entre l'acquisition de la donnée biométrique et l'identification de l'utilisateur, les utilisateurs considèrent que le système biométrique n'est pas facile d'utilisation ;

– *Protection de la vie privée :* le processus d'authentification biométrique nécessite l'acquisition d'information biométrique pour générer un modèle propre à chaque individu durant la phase d'enrôlement. Généralement, ce modèle est stocké dans une base de données ou un support (carte à puce, clef USB, *etc.*) pour être utilisé durant la phase de vérification et d'identification. Cependant,

ce stockage constitue un risque potentiel concernant la mauvaise utilisation de données personnelles, ce qui est considéré comme une atteinte de la vie privée et violation des libertés personnelles des utilisateurs ;

– *Impact d'intrusion physique :* l'acquisition de données biométriques nécessite l'interaction de l'utilisateur avec le capteur biométrique. Dépendant de la modalité utilisée, l'acquisition de données est effectuée sans ou avec contact avec le capteur biométrique. D'après nos connaissances, aucun document n'a souligné de dommages physiques pour les utilisateurs de ces systèmes. Cependant, plusieurs inquiétudes ont été soulevées le long de cette interaction. Prenons le cas de reconnaissance par la rétine, cette technologie assure une bonne fiabilité et une haute barrière contre la fraude. Malgré l'efficacité de cette technologie, sa prolifération a été limitée car elle est considérée comme intrusive par les usagers (comme le risque de dégradation de la cornée).

Le questionnaire ainsi créé consiste à extraire ces facteurs. Plus particulièrement, il a été conçu pour recueillir les caractéristiques socio-démographiques, l'expérience et l'attitude des usagers qui pourraient avoir un impact sur leur niveau d'acceptabilité et leur degré de satisfaction vis-à-vis du système biométrique. Le questionnaire prend en considération les travaux de littérature présentés dans le chapitre 2. Il contient également d'autres questions que nous ont paru intéressantes lors de la création des deux bases (GREYC-Keystroke et ENSIB) de données biométriques, l'étude d'acceptabilité [151] du logiciel GREYC-Keystroke, et l'avis de deux experts en psychologie (Patrice Georget et Cécile Sénémeaud) de l'UCBN [1]. Les opinions des experts étaient importantes surtout pour la formulation et l'ordre des questions. En effet, la question doit être aussi neutre que possible afin d'éviter les réponses biaisées. Nous avons utilisé l'échelle Likert paire [152] pour les réponses du questionnaire. L'échelle paire est utilisée pour éviter d'avoir beaucoup des réponses neutres (ni positive et ni négative). Le questionnaire contient 18 questions qui sont divisées en deux catégories :

– **Perception générale des systèmes biométriques** qui contient 7 questions qui visent à comprendre l'expérience des utilisateurs sur la technologie biométrique ;

– **Perception du système biométrique testé** qui contient 11 questions qui visent à mesurer l'acceptabilité et la satisfaction des utilisateurs sur le système

1. Université de Caen Basse-Normandie

testé.

En plus de ces questions, nous avons posé d'autres questions pour avoir quelques informations sur l'individu tels que : le genre, l'âge et la profession. Ces caractéristiques sont nécessaires pour vérifier s'ils ont un impact sur la perception du système biométrique testé. Nous avons également posé la question Q_{16} pour déterminer où l'utilisation du système testé serait la plus appropriée pour l'utilisateur (*i.e.*, accès physique, logique ou les deux).

Enquête sur l'usage d'un système biométrique

Partie I. Caractéristiques socio-démographiques

Date de naissance (année)	...
Genre	☐ masculin ☐ féminin
Dans quel continent habitez-vous ?	☐ asie ☐ afrique ☐ europe ☐ amérique du nord ☐ amérique du sud ☐ autre
Niveau d'éducation	☐ universitaire ou grande école ☐ lycée ☐ autre
Statut professionnel	☐ étudiant ☐ salarié ☐ retraité ☐ autre

Partie II. Perception générale des systèmes biométriques

Q_1. Avez-vous déjà entendu parler de systèmes biométriques **(avant notre étude)** ?	☐ oui ☐ non
Q_2. Avez-vous déjà utilisé un système biométrique **(avant notre étude)** ?	☐ oui ☐ non
Q_3. Avez-vous déjà été personnellement victime d'une fraude d'identité ?	☐ oui ☐ non
Q_4. Comment évaluez-vous vos connaissances sur la technologie biométrique ?	☐ pas du tout importantes ☐ plutôt pas importantes ☐ plutôt importantes ☐ tout à fait importantes ☐ je ne sais pas
Q_5. Comment évaluez-vous vos connaissances concernant le vol d'identité ?	☐ pas du tout importantes ☐ plutôt pas importantes ☐ plutôt importantes ☐ tout à fait importantes ☐ je ne sais pas
Q_6. Selon vous, est-ce que les solutions basées sur un secret (*ex.*, mot de passe) sont une solution pertinente contre la fraude (par ex. commerce électronique) ?	☐ pas du tout d'accord ☐ plutôt pas d'accord ☐ plutôt d'accord ☐ tout à fait d'accord ☐ je ne sais pas
Q_7. Selon vous, est-ce que la biométrie peut être une solution pertinente contre la fraude (par ex. commerce électronique) ?	☐ pas du tout d'accord ☐ plutôt pas d'accord ☐ plutôt d'accord ☐ tout à fait d'accord ☐ je ne sais pas

Partie III. Perception du système biométrique testé

Q_8. Avez-vous déjà utilisé cette modalité **(avant notre étude)** ?	☐ oui ☐ non
Q_9. Ressentez-vous une gêne à utiliser ce système biométrique particulier ?	☐ pas du tout gênant ☐ plutôt pas gênant ☐ plutôt gênant ☐ tout à fait gênant ☐ je ne sais pas
Q_{10}. Trouvez-vous que l'utilisation de cette technologie est une atteinte à votre vie privée ?	☐ pas du tout intrusive ☐ plutôt pas intrusive ☐ plutôt intrusive ☐ tout à fait intrusive ☐ je ne sais pas
Q_{11}. Trouvez-vous l'utilisation de ce système biométrique facile et agréable ?	☐ pas du tout facile ☐ plutôt pas facile ☐ plutôt facile ☐ tout à fait facile ☐ je ne sais pas
Q_{12}. La vérification est-elle rapide ?	☐ pas du tout rapide ☐ plutôt pas rapide ☐ plutôt rapide ☐ tout à fait rapide ☐ je ne sais pas
Q_{13}. La réponse pour la vérification est-elle correcte ?	☐ jamais ☐ rarement ☐ parfois ☐ toujours ☐ je ne sais pas
Q_{14}. Selon vous, est-ce que le système utilisé peut-il être facilement contourné ?	☐ pas du tout d'accord ☐ plutôt pas d'accord ☐ plutôt d'accord ☐ tout à fait d'accord ☐ je ne sais pas
Q_{15}. Seriez-vous prêt à utiliser ce système biométrique dans le futur ?	☐ pas du tout d'accord ☐ plutôt pas d'accord ☐ plutôt d'accord ☐ tout à fait d'accord ☐ je ne sais pas
Q_{16}. Si vous êtes prêt à utiliser le système dans le futur, seriez-vous prêt l'utiliser pour gérer le contrôle d'accès logique (accès à un ordinateur) ou physique (accès à un lieu) ?	☐ physique ☐ logique
Q_{17}. Avez-vous confiance dans ce système ?	☐ non ☐ pas vraiment ☐ plutôt ☐ oui ☐ je ne sais pas
Q_{18}. Quelle est votre appréciation générale de ce système ?	☐ pas du tout satisfait ☐ plutôt pas satisfait ☐ plutôt satisfait ☐ tout à fait satisfait ☐ je ne sais pas

4.2.2 Préparation de données

Cette étape consiste à collecter les informations pertinentes, en supprimant de l'étude les vecteurs réponses ayant un nombre prédéfini de questions sans réponse. Cette phase améliore la précision et la fiabilité des informations extraites et déduites.

4.2.3 Analyse socio-démographique

Cette analyse consiste à extraire les dépendances significatives (s'il en existe) entre les caractéristiques socio-démographiques (comme le genre) et les questions de satisfaction (comme l'appréciation générale). Comme nous avons vu dans le chapitre 2 section 2.4.1, il existe des facteurs culturels (voire religieux) qui peuvent avoir un

impact sur la perception des usagers lors de l'utilisation des systèmes biométriques. Prenons le cas de l'Arabie Saoudite, le genre a un impact sur l'utilisation des systèmes de reconnaissance faciale : les femmes ne peuvent pas utiliser ces systèmes pour des raisons religieuses. Il sera également utile de vérifier si l'âge ou la profession a un impact sur la perception des usagers : dans un système biométrique idéal, son utilisation doit être acceptable par tous les usagers de la population. Dès lors, pour qu'un système soit opérationnel et utilisable, étudier ces dépendances s'avère indispensable dans certains contextes d'utilisation et pour une cible utilisateurs. Pour ce faire, nous utilisons le test de Kruskal-Wallis (KW) [153]. Le choix du test KW est retenu puisqu'il nous permet d'identifier s'il existe une différence significative en terme de moyenne entre deux distributions de scores (*i.e.*, scores provenant de l'échelle Likert utilisée). En utilisant ce test, nous pouvons valider statistiquement nos conclusions statistiques à un intervalle de confiance égale à 95%. Une description du test KW, ainsi que son critère de décision est donné à l'Annexe A.

4.2.4 Analyse de perception

4.2.4.1 Les classifieurs

Pour certains domaines d'application, il est essentiel d'avoir des outils qui permettent de produire des procédures de classification compréhensibles par l'utilisateur. Les réseaux bayésiens [154, 155] et les arbres de décision [156] répondent à cette contrainte car ils représentent graphiquement un ensemble de règles facilement interprétables. Nous avons ainsi utilisé ces deux classifieurs, qui permettent d'expliquer les réponses des usagers, pour mesurer leur satisfaction.

4.2.4.1.1 Réseaux Bayésiens

Les réseaux bayésiens [154, 155] ont été développés pour tenter de résoudre des problèmes de prédiction et d'abduction, courants en intelligence artificielle. Un réseau bayésien (RB) est un modèle probabiliste graphique permettant d'acquérir, de capitaliser et d'exploiter des connaissances. C'est un graphe causal, orienté, acyclique, dont les noeuds sont des variables d'intérêts du domaine, les arcs des relations de dépendance entre ces variables. L'ensemble des noeuds et des arcs forme ce que l'on appelle la structure du réseau bayésien.

On définit un réseau bayésien de la facon suivante :
$RB = (G, \theta)$ est un réseau bayésien si $G = (X, E)$ est un graphe acyclique orienté dont les sommets représentent un ensemble de variables aléatoires $X = \{X_1, \ldots, X_n\}$,

et si $\theta_i = [P(X_i/X_{Pa(X_i)})]$ est la matrice des probabilités conditionnelles du noeud i connaissant l'état de ses parents $Pa(X_i)$ dans G.

Dans certaines applications, les connaissances de l'expert sur la structure que peut avoir le réseau bayésien ne sont que partielles. Fauré [157] présente une liste de ces connaissances a priori que sont : la déclaration d'un noeud racine, la déclaration d'un noeud feuille, l'existence (ou l'absence) d'un arc entre deux noeuds précis, l'indépendance de deux noeuds conditionnellement à certains autres, la déclaration d'un ordre (partiel ou complet) sur les variables, la déclaration d'un noeud cible (dans le cas de classification).

La construction d'un réseau bayésien est constituée à la fois d'un graphe causal et d'un ensemble de probabilités conditionnelles. L'apprentissage d'un tel réseau est généralement divisé en deux étapes : apprentissage de la structure et apprentissage des paramètres.

Étape 1. Apprentissage de la structure

L'algorithme d'apprentissage de la structure utilisé dans notre étude est celui présenté dans [158] de la libraire Weka [2]. Il consiste à rechercher les relations causales qui existent entre les variables en utilisant les tests d'indépendance conditionnelle.

Nous allons tout d'abord introduire quelques notations. Etant donnée $(x_i)_{1\leq i\leq n}$ n variables aléatoires à étudier avec la base des observations D de ses variables, nous définissons ce qui suit :

- Soit r_i le cardinal de la variable x_i.
- Nous désignons par q_i le cardinal de l'ensemble des parents de la variable x_i dans la structure B_S. q_i est ainsi calculé comme étant le produit des cardinalités des parents de x_i ($q_i = \prod_{x_j \in pa(x_i)} r_j$). À noter que, $pa(x_i) = \emptyset$ implique $q_i = 1$.
- Nous désignons par N_{ij} $(1 \leq i \leq n,\ 1 \leq j \leq q_i)$ le nombre d'enregistrements dans D pour lesquelles $pa(x_i)$ prend sa $j^{\text{ème}}$ valeur.
- Nous désignons par N_{ijk} $(1 \leq i \leq n,\ 1 \leq j \leq q_i,\ 1 \leq k \leq r_i)$ le nombre d'enregistrements dans D pour lesquelles $pa(x_i)$ prend sa $j^{\text{ème}}$ valeur et x_i prend sa $k^{\text{ème}}$ valeur. N_{ij} est ainsi égal à $\sum_{k=1}^{r_i} N_{ijk}$.
- Nous désignons par N le nombre d'enregistrements de la base des observations D.

2. http://www.cs.waikato.ac.nz/ml/weka/

L'algorithme d'apprentissage du réseau bayésien retenu est divisé en deux parties :

– Apprentissage du squelette : à partir d'un graphe complet non orienté, nous essayons de trouver des relations d'indépendance conditionnelle $\langle x, y|Z\rangle$ dans la base des observations D. Pour chaque couple de variables x, y, nous considérons des ensembles Z de cardinalités 0, puis 1 jusqu'à un maximum défini par l'utilisateur. En outre, l'ensemble Z est un sous-ensemble de variables qui sont des voisins de x et y. Si une indépendance est identifiée, l'arrête entre x et y est supprimée à partir du squelette. Pour tester si les variables x et y sont conditionnellement indépendantes étant donné un ensemble de variables Z, la structure des arcs $\{z \rightarrow y \ / \ \forall_{z \in Z}\}$ est comparée avec $\{x \rightarrow y\} \cup \{z \rightarrow y \ / \ \forall_{z \in Z}\}$ en utilisant l'un des scores suivants :

  - Entropie
$$H(B_S, D) = -N \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} \frac{N_{ijk}}{N} \log \frac{N_{ijk}}{N_{ij}} \tag{4.1}$$
  - Critère d'information d'Akaike (*Akaike Information Criterion*, abrégé par AIC)
$$Q_{AIC}(B_S, D) = H(B_S, D) + K \tag{4.2}$$
où $K = \sum_{i=1}^{n} (r_i - 1).q_i$
  - Minimum de la longueur de description (*Minimum Description Length*, abrégé par MDL)
$$Q_{MDL}(B_S, D) = H(B_S, D) + \frac{K}{2} \log N \tag{4.3}$$
  - Bayésien
$$Q_{Bayes}(B_S, D) = \prod_{i=0}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(r_i - 1 + N_{ij})!} \prod_{k=1}^{r_i} N_{ijk}! \tag{4.4}$$

– Apprentissage du graphe orienté acyclique (DAG) : la première étape d'orientation des arrêtes consiste à chercher toutes les configurations de la forme $x - -z - -y$ où x et y ne sont pas reliés dans le squelette. Nous désignons par Z l'ensemble des variables qui a justifié la suppression du lien entre x et y. Si $z \notin Z$, alors nous associons la direction $x \rightarrow z \leftarrow y$.
Enfin, un ensemble de règles graphiques est appliqué à diriger les arrêtes restantes.

- règle 1 : $i \rightarrow j - -k$ & $i - / - k \Rightarrow j \rightarrow k$
- règle 2 : $i \rightarrow j \rightarrow k$ & $i - -k \Rightarrow i \rightarrow k$
- règle 3 : $i \rightarrow j \leftarrow k$ & la structure $S_1 \Rightarrow m \rightarrow j$
- règle 4 : $i \rightarrow j$ & la structure $S_2 \Rightarrow i \rightarrow m$ & $k \rightarrow m$
- règle 5 : s'il reste toujours des arrêtes non orientées, nous les orientons au hasard.

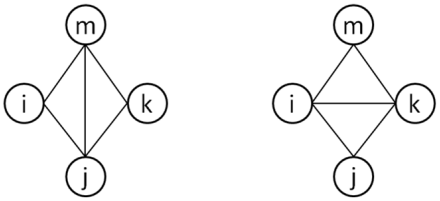


FIG. 4.2 – Les deux structures S1 (à gauche) et S2 (à droite).

Étape 2. Apprentissage des paramètres

Une fois la structure du réseau fixée, il faudra estimer les probabilités conditionnelles associées. Pour ce faire, il existe deux méthodes d'apprentissage de paramètres :

– À partir de données complètes : la probabilité d'un événement est estimée par la fréquence d'apparition de l'événement dans la base de données. Cette approche, appelée maximum de vraisemblance (MV), est décrite par :

$$\hat{P}(X_i = x_k \mid parent(X_i) = c_j) = \hat{\theta}_{i,j,k} = \frac{N_{i,j,k}}{\sum_k N_{i,j,k}} \tag{4.5}$$

où $N_{i,j,k}$ est le nombre d'événements dans la base de données pour lesquels la variable X_i est dans l'état x_k et ses parents sont dans la configuration c_j.

– À partir de données incomplètes : dans certaines applications, les bases de données sont souvent incomplètes. Certaines variables ne sont observées que partiellement ou même jamais. La méthode la plus utilisée lorsque les données sont incomplètes repose sur l'algorithme itératif Expectation-Maximisation (EM) présentée dans [159].

4.2.4.1.2 Arbres de décision

Un arbre de décision, introduit par Breiman *et al.* [156], est un outil d'aide à la décision et à l'exploration de données. C'est la représentation graphique d'une procédure de classification. En effet, à toute description complète est associée une seule feuille

de l'arbre de décision. Cette association est définie en commençant à la racine de l'arbre et en descendant dans l'arbre selon les réponses aux tests qui étiquettent les noeuds internes. La classe associée est alors la classe par défaut associée à la feuille qui correspond à la description. La procédure de classification obtenue a une traduction immédiate en terme de règles de décision. Les systèmes de règles obtenus sont particuliers car l'ordre dans lequel on examine les attributs est fixé, et les règles de décision sont mutuellement exclusives.

L'apprentissage des arbres de décision consiste généralement à diviser récursivement et le plus efficacement possible les exemples de l'ensemble des observations. Cette tâche est effectuée par des tests définis à l'aide des attributs jusqu'à ce que l'on obtienne des sous-ensembles d'exemples ne contenant (presque) que des exemples appartenant tous à une même classe. Dans toutes les méthodes d'apprentissage par arbre de décision dans la littérature, nous trouvons les trois primitives suivantes : 1) décider si un noeud est terminal, 2) sélectionner un test à associer à un noeud et 3) affecter une classe à une feuille. Le schéma général des ces méthodes est illustré dans l'**Algorithme** 1. Ces méthodes diffèrent par les choix effectués pour le test (comme l'utilisation du gain ou de la fonction entropie) et le critère d'arrêt (*i.e.*, comment décider si un noeud est terminal ?).

Algorithme 1 Algorithme d'apprentissage générique par arbre de décision

Entrées: la base des observations D
Sorties: l'arbre de décision
début
Initialiser à l'arbre vide
La racine est le noeud courant
répéter
Décider si le noeud courant est terminal
si le noeud est terminal **alors**
Affecter une classe
sinon
Sélectionner un test et créer le sous-arbre
finsi
Passer au noeud suivant non exploré s'il en existe
jusqu'à obtenir un arbre de décision
fin

Avant de présenter les algorithmes d'apprentissage par arbres de décision, nous allons introduire quelques notations. Le langage de représentation est constitué d'un certain nombre d'attributs. Ces attributs peuvent être binaires, qualitatifs (*i.e.*, à

valeurs dans un ensemble fini) ou continus (*i.e.*, à valeurs réelles). Nous considérons que les noeuds d'un arbre de décision sont repérés par des positions qui sont des mots sur $\{1, \ldots, p\}^*$, où p est l'arité maximale des noeuds. Étant donnés une base d'observations D, une variable classe $C = \{1, \cdots, c\}$ et un arbre de décision t, nous avons : chaque position p de t correspond à un sous-ensemble de D qui est l'ensemble des observations qui satisfont les tests de la racine jusqu'à cette position. Par conséquent, nous définissons, pour toute position p de l'arbre t, les quantités suivantes :

- $N(p)$ est le cardinal de l'ensemble des observations associé à p ;
- $N(k/p)$ est le cardinal de l'ensemble des observations associé à p qui sont de classe k ;
- $P(k/p) = \frac{N(k/p)}{N(p)}$ est la proportion d'éléments de classe k à la position p ;
- Les fonctions permettant de mesurer le degré du mélange des classes pour tout échantillon pour toute position de l'arbre en construction sont définies comme suit :

$$Entropie(p) = -\sum_{k=1}^{c} P(k/p) \times \log(P(k/p)) \tag{4.6}$$

$$Gini(p) = 1 - \sum_{k=1}^{c} P(k/p)^2 \tag{4.7}$$

- La fonction gain pour choisir le test qui doit étiqueter le noeud courant de l'arbre est définit comme suit :

$$Gain(p, test) = f(p) - \sum_{i=1}^{n} P_i \times f(p_i) \tag{4.8}$$

où p désigne une position, f est la méthode adoptée (*Entropie* ou *Gini*), *test* un test d'arité n, P_i est la proportion d'éléments de D à la position p qui vont en position p_i (*i.e.*, qui satisfont la $i^{\text{ème}}$ branche du test *test*).

Dans le cadre de cette thèse, nous avons testé deux algorithmes d'apprentissage par arbres de décision : *Classification And Regression Tree* (CART [156]) et l'algorithme *C4.5* développé par [160]. Dans la suite, nous présentons une brève description de ces deux algorithmes.

4.2.4.1.2.1 L'algorithme CART

La méthode CART, introduite par Breiman *et al.* [156], permet d'inférer des arbres de décision binaires. La phase d'expansion de la méthode CART est définie comme suit. Pour une base des observations D :

1. Décider si un noeud est terminal : un noeud p est terminal si $Gini(p) \leq i_0$ ou $N(p) \leq n_0$, où i_0 et n_0 sont des paramètres à fixer.

2. Sélectionner un test à associer à un noeud : soit p une position, la méthode sélectionne le test *test* qui maximise le gain défini dans l'équation 4.9 en utilisant la fonction *Gini* pour mesurer le degré de mélange :

$$Gain(p, test) = Gini(p) - (P_{gauche} \times Gini(p_1) + P_{droite} \times Gini(p_2)) \quad (4.9)$$

   où P_{gauche} (respectivement P_{droite}) représente la proportion d'éléments de l'ensemble des exemples associés à p qui vont sur le noeud en position p_1 (respectivement p_2).

3. Affecter une classe à une feuille : l'affectation est effectuée selon la classe majoritaire.

4.2.4.1.2.2 L'algorithme C4.5

La méthode $C4.5$ [160] est une extension de la méthode *Iterative Dichotomiser 3* ($ID3$) développée par Quinlan [161] en 1986. La phase d'expansion de la méthode $C4.5$ est défini comme suit. Pour une base d'observations D, il faut :

1. Décider si un noeud est terminal : un noeud p est terminal si tous les éléments associés à ce noeud sont dans une même classe ou si aucun test n'a pu être sélectionné. En d'autres termes, à chaque étape, dans l'ensemble des tests disponibles, ne peuvent être envisagés que les tests pour lesquels il existe au moins deux branches ayant au moins deux éléments dans la base des observations D. Si aucun *test* ne satisfait cette condition alors le noeud est considéré comme terminal.

2. Sélectionner un test à associer à un noeud : soit p une position, la méthode sélectionne le test *test* qui maximise le gain définie dans l'équation 4.8 en utilisant la fonction *entropie* pour mesurer le degré du mélange. La fonction *Gain*, ainsi définie, privilégie les attributs ayant un grand nombre de valeurs. Elle est donc pondérée par une fonction qui pénalise les tests qui répartissent les éléments en un trop grand nombre de sous-classes. Elle est définie par :

$$GainRatio(p, test) = \frac{Gain(p, test)}{SplitInfo(p, test)} \quad (4.10)$$

avec

$$SplitInfo(p, test) = -\sum_{i=1}^{n} P\prime(i/p) \times \log(P\prime(i/p)) \tag{4.11}$$

où n est l'arité du test $test$ et $P\prime(i/p)$ est la proportion des éléments présents à la position p satisfont la $i^{\text{ème}}$ branche du test $test$.

3. Affecter une classe à une feuille : l'affectation est effectuée selon la classe majoritaire.

À noter que, la performance des algorithmes d'apprentissage par réseaux bayésiens et arbres de décision dépendent du jeu de données en question. Ainsi, aucun algorithme d'apprentissage n'est meilleur que les autres existants. Pour cela, nous présentons dans la prochaine section les critères d'évaluation de tels classifieurs.

4.2.4.2 Evaluation de classifieurs

Les classifieurs, réseaux baysésiens et arbres de décision, sont des outils qui sont couramment utilisés dans l'analyse des décisions, afin d'aider à identifier une stratégie susceptible d'atteindre un objectif. Ils possèdent l'avantage d'être compréhensible par tout utilisateur et d'avoir une traduction immédiate en termes de règles de décision. Cependant, les méthodes d'apprentissage des classifieurs sont basées sur des mesures statistiques et des heuristiques, et ainsi ne sont pas fiables à 100%. Il sera ainsi nécessaire d'utiliser des métriques pour choisir le «meilleur» classifieur. Le terme «meilleur» répond souvent, mais **pas toujours**, au taux d'erreur en généralisation. Ricco [162] a détaillé dans sa thèse l'évaluation et la comparaison empirique de classifieurs. Les critères d'évaluation ainsi retenus sont :

– La précision : elle montre la capacité intrinsèque du classifieur à reconnaître la variable à prédire dans la population. Elle est quantifiée par le pourcentage d'instances correctement classifiées ;

– L'aire sous la courbe ROC (AUC) : L'AUC est devenue une meilleure alternative que la précision pour évaluer des classifieurs [163]. L'AUC d'un classifieur est équivalente à la probabilité qu'un classifieur donne un meilleur rang à un élément positif par rapport à un élément négatif, tous deux choisis aléatoirement dans la base. Une description détaillée de l'utilité d'AUC pour évaluer les classifieurs est donnée par Ling *et al.* [163]. L'AUC est définie par :

$$\widehat{AUC} = \frac{S_0 - n_0(n_0+1)/2}{n_0 n_1} \tag{4.12}$$

où

- n_0 : correspond au nombre d'éléments étiquetés positifs dans la base ;
- n_1 : correspond au nombre d'éléments étiquetés négatifs dans la base ;
- $S_0 = \sum r_i$, où r_i est le rang du $i^{\text{ème}}$ élément étiqueté positif.

– La compréhensibilité : elle permet de quantifier l'exploitabilité du modèle produit. Plus un modèle est complexe, moins on aura de facilité à le comprendre. Prenons le cas d'un réseau bayésien, le nombre important des parents pour un noeud donné affecte l'identification des fortes dépendances avec eux.

Comme nous l'avons dit, la performance des algorithmes dépend du jeu de données en question. Ainsi, aucun algorithme d'apprentissage n'est meilleur que les autres existants. L'idéal est d'utiliser plusieurs algorithmes d'apprentissage, et ensuite retenir le meilleur modèle selon les critères présentés par Ricco [162] en termes de la précision, d'aire sous la courbe ROC (AUC) et de compréhensibilité. En utilisant ces critères de comparaison, nous pouvons ainsi dire, qu'un algorithme est meilleur que les autres sur le jeu de données en question.

4.3 Résultats expérimentaux

4.3.1 Le protocole

Deux systèmes biométriques et un échantillon de 100 volontaires (étudiants, enseignants-chercheurs, secrétaires et employés) ont participé à l'étude permettant de montrer l'intérêt de la méthode proposée. L'âge et le genre des volontaires sont présentés dans la figure 4.3. Les deux systèmes testés sont GREYC-Keystroke (G-Keystroke) et GREYC-Face (G-Face) présentés dans le chapitre 1 section 1.3. La performance des deux systèmes testés est donnée à la figure 4.4. Le choix de la comparaison de l'usage d'une modalité basée sur des caractéristiques morphologiques (G-Face) et comportementales (G-Keystroke) est effectué pour diverses raisons. Premièrement, nous souhaitons quantifier l'usage de ces deux systèmes développés dans le laboratoire de recherche GREYC afin de tester leur opérationnalité (en terme d'acceptabilité). Deuxièmement, il est utile de déterminer quel type de modalité (morphologique ou comportementale) est la mieux perçue par les usagers, et pour quel type d'accès (accès

à des ressources physiques ou logiques). Troisièmement, les résultats de l'évaluation de ces deux systèmes différents (en termes d'enrôlement et d'utilisation) devraient être aussi différents que possible. Pour toutes ces raisons, nous avons choisi d'illustrer cette méthodologie d'évaluation sur les deux systèmes biométriques G-Face et G-Keystroke.

Avant de répondre au questionnaire, les volontaires ont testé les deux systèmes biométriques comme suit : tout d'abord, une brève explication a été donnée pour chaque utilisateur afin de présenter les deux systèmes étudiés, le but et la finalité de cette étude. Ensuite, il y a eu la phase d'utilisation des deux systèmes consistant en un enrôlement, puis en plusieurs tentatives de vérification en jouant le rôle d'un utilisateur légitime et d'un imposteur.

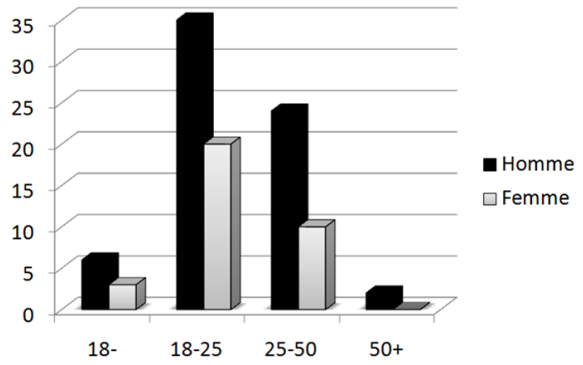


FIG. 4.3 – Âge et genre des volontaires.

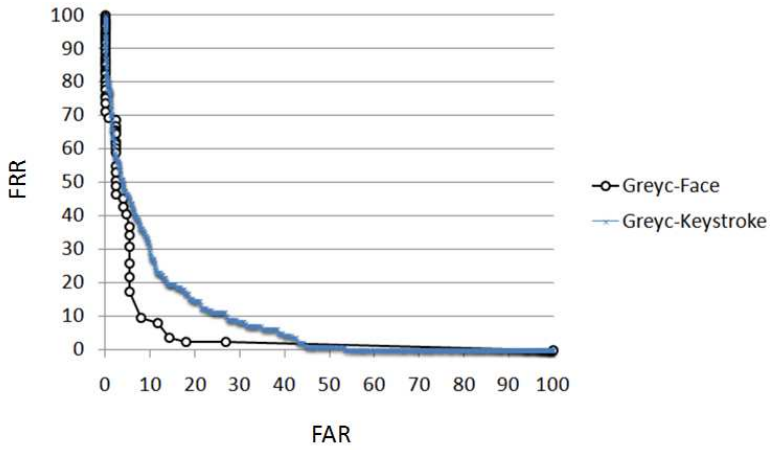


FIG. 4.4 – Les courbes ROC des deux systèmes étudiés : GREYC-Face (EER=8,76%) et GREYC-Keystroke (EER=17,51%).

4.3.2 Préparation de données

La première étape de la méthode proposée consiste à supprimer les vecteurs réponses ayant un nombre prédéfini X de questions sans réponses. Pour $X \geq 3$, 2 vecteurs réponses, 1 pour chaque système, ont été retirés de notre étude. Nous analysons ainsi les réponses de 99 vecteurs de chaque système.

4.3.3 Analyse socio-démographique

Nous étudions dans cette section les relations de dépendance, pour chaque système, entre les caractéristiques socio-démographiques et les réponses aux questions de satisfaction. Le tableau 4.1 illustre les p-valeurs du test KW pour un intervalle de confiance de 95%. Les valeurs en gras indiquent des relations significatives en utilisant le critère défini dans l'Annexe A (équation A.6). L'analyse de données de la présente recherche indique les résultats suivants :

- L'âge a un impact sur leur perception concernant la robustesse de G-Kesytroke contre les attaques : les volontaires ayant un âge ≥ 28 ans considèrent que le système est moins robuste par rapport aux autres ;

- Pour G-Kesytroke, le niveau d'éducation est significativement lié à leurs opinions sur divers facteurs que sont la gêne lors de son utilisation, l'inquiétude concernant la vie privée, le temps de vérification et sa performance :
  - les volontaires ayant un diplôme universitaire sont moins gênés lors de son utilisation que les autres sujets ;
  - les volontaires non diplômés ont exprimé plus d'inquiétudes concernant l'atteinte à la vie privée que les autres ;
  - les volontaires ayant un diplôme universitaire trouvent que la phase de vérification est plus rapide contrairement aux lycéens ;
  - la performance du système est mieux perçue par les volontaires ayant un diplôme universitaire que les lycéens.

- Le niveau d'éducation est significativement lié à l'utilisation de G-Face : les lycéens trouvent que son utilisation est plus compliquée contrairement au reste des sujets.

questions de perception	Genre	Age	Education	Profession
Connaissances sur la technologie biométrique	0.089	0.652	0.134	0.911
Connaissances concernant le vol d'identité	0.247	0.831	**0.028**	0.578
Solutions basées sur un secret contre la fraude	0.482	0.804	0.213	0.778
Solutions basées sur la biométrie contre la fraude	0.71	0.324	0.74	0.546
Gêne	0.419/0.385	0.509/0.473	0.096/**0.007**	0.696/0.428
Vie privée	0.206/0.556	0.65/0.649	0.196/**0.044**	0.756/0.36
Utilisation facile	0.145/0.438	0.063/0.522	**0.012**/0.392	0.46/0.406
Vérification rapide	0.135/0.144	0.226/0.294	0.195/**0.039**	0.095/0.202
Performance	0.982/0.075	0.779/0.717	0.78/**0.034**	0.369/0.058
Le système peut être facilement contourné	0.276/0.226	0.492/**0.02**	0.558/0.204	0.405/0.22
Utilisation future	0.079/0.48	0.604/0.883	0.721/0.821	0.33/0.255
Confiance	0.414/0.689	0.469/0.218	0.709/0.947	0.372/0.068
Appréciation générale	0.078/0.129	0.984/0.873	0.3/0.768	0.459/0.917

TAB. 4.1: Facteurs socio-démographiques et questions de satisfaction, analyse du test Kruskal-Wallis (KW) : les lignes avec 2 p-valeurs correspondent à des questions pour un système spécifique (G-Face / G-Keystroke).

4.3.4 Analyse descriptive

Nous présentons dans cette section, l'expérience des volontaires sur la technologie biométrique ainsi qu'une étude comparative sur les deux systèmes étudiés. Le test de KW est également utilisé pour extraire les différences statistiques des réponses sur les deux systèmes. Le tableau 4.2 présente les résultats pour un intervalle de confiance de 95%. Les valeurs en gras indiquent les différences significatives en utilisant le critère défini dans l'Annexe A (équation A.6). À partir des résultats statistiques et du tableau 4.2, nous pouvons déduire les points suivants :

- Les volontaires sont familiers avec la technologie biométrique : $74,75\%$ ont déjà entendu parlé de systèmes biométriques, $40,4\%$ ont déjà utilisé un système biométrique, et $38,38\%$ ont de bonnes connaissances sur cette technologie ;

- En utilisant le test de KW (p-valeur $< 0,01$), les volontaires considèrent que la biométrie $(90,9\%)$ est une solution plus pertinente contre la fraude que les solutions basées sur un secret $(33,33\%)$;

- Les résultats montrent qu'il n'y a pas de différence significative sur la comparaison des facteurs suivants : la facilité d'utilisation des deux systèmes, le temps de vérification, leur volonté d'utilisation de tels systèmes dans le futur et la confiance dans de tels produits. $24,24\%$ des volontaires trouvent que l'utilisation de G-Face n'est pas facile ($14,14\%$ pour G-Keystroke). $10,1\%$ trouvent que le temps de vérification de G-Face est lent ($14,14\%$ pour G-Keystroke). $22,22\%$ hésitent voire refusent l'utilisation de G-Face dans le futur ($17,17\%$ pour G-Keystroke). $32,32\%$ des volontaires n'ont pas confiance dans G-Face

(20, 2% pour G-Keystroke) ;

– D'autre part, il y a des différences significatives sur la gêne lors de l'utilisation des deux systèmes, l'inquiétude concernant la protection de la vie privée, perception sur la performance des deux systèmes, perception sur la robustesse des deux systèmes contre les attaques et l'appréciation générale. Les volontaires étaient plus gênés lors de l'utilisation de G-Face (25, 25%) que de G-Kestroke (16, 16%). Les volontaires ont exprimé plus d'inquiétudes concernant la vie privée lors de l'utilisation de G-Face (47, 47%) que de G-Keystroke (13, 13%). Les volontaires considèrent que la performance de G-Keystroke est meilleure que celle de G-Face. Les volontaires trouvent que G-Keystroke (53, 53%) est plus robuste contre les attaques que G-Face (34, 34%). Pour l'appréciation générale, ils étaient plus satisfaits lors de l'utilisation de G-Keystroke (88, 89%) que de G-Face (75, 75%) ;

– Enfin, 23, 23% des sujets préfèrent utiliser G-Face (62, 62% pour G-Keystroke) pour gérer l'accès logique, 41, 41% d'entre eux préfèrent utiliser G-Face (14, 14% pour G-Kesytroke) pour l'accès physique. Ces résultats montrent que G-Kesytroke est plutôt adapté pour l'accès logique, tandis que G-Face pour l'accès physique (ce qui était attendu).

Questions de perception	**G-Face**	**G-Kesytroke**	**p-valeur**
Gêne	1.916	1.67	**0.034**
Vie privée	2.427	1.626	**<< 0.05**
Utilisation facile	3.133	3.242	0.355
Vérification rapide	3.329	3.357	0.611
Performance	3.176	3.51	**0.007**
Le système peut être facilement contourné	2.659	2.3	**0.005**
Utilisation future	2.989	3.112	0.238
Confiance	2.84	3	0.176
Appréciation générale	2.823	3.175	**0.0021**

TAB. 4.2: Etude comparative des réponses entre les deux systèmes étudiés : G-Face et G-Keystroke.

4.3.5 Discussion

Les résultats statistiques et l'étude comparative des deux systèmes étudiés ont montré des taux surprenants. Nous avons trouvé un taux élevé (47, 47%) d'inquiétude sur le respect de la vie privée lors de l'utilisation de G-Face. Les résultats montrent également une forte inquiétude concernant la perception qu'ont les sujets sur la performance des deux systèmes et l'appréciation générale qu'ils en ont. Les volontaires

considèrent que la performance de G-Keystroke (avec un $EER = 17,51\%$) est meilleure que telle de G-Face (avec un $EER = 8,76\%$), et ils étaient plus satisfaits lors de l'utilisation de G-Keystroke ($88,89\%$) que de G-Face ($75,75\%$). Ces résultats statistiques montrent que le système le moins performant (*i.e.*, en utilisant la valeur de l'EER) était le mieux perçu. Il est ainsi important d'expliquer les réponses données par les utilisateurs afin d'expliquer ce comportement inattendu. Ce genre d'analyse a pour objectif également de mieux comprendre le comportement d'usagers, leurs besoins et perceptions, c'est ce que nous présentons dans la section suivante.

4.3.6 Analyse de perception

Dans cette section, nous analysons la dépendance des questions de satisfaction (section 4.2.1, parties II et III) en utilisant les réseaux bayésiens et les arbres de décision. Cette analyse va nous permettre de mieux comprendre le comportement des usagers afin d'expliquer les résultats surprenants présentés dans la section 4.3.4 (par exemple, la différence entre la performance perçue et réelle des deux systèmes testés). Pour ce faire, nous procédons comme suit :

1. Partition des questions

Afin de générer les modèles de décision, nous avons divisé les questions de satisfaction en deux ensembles (S_{Cause} et S_{Effet}) selon la relation de cause à effet. Cette division va nous permettre d'identifier les causes (comme la facilité d'utilisation) qui ont un impact sur leur satisfaction (comme l'appréciation générale) lors de l'utilisation du système testé. D'autre part, puisque la variable classe doit être nominale (*i.e.*, nature de construction des classifieurs), nous avons défini ces deux ensembles comme suit :

- $S_{Cause} = \{Q_i / i = 4, 5, 6, 7, 11, 12, 14\}$. Les questions appartenant à S_{Cause} sont considérées comme des attributs numériques en transformant les choix nominaux par des entiers. Prenons par exemple la question Q_4, le choix nominal «pas du tout importantes» est remplacé par 1, «plutôt pas importantes» par 2, «plutôt importantes» par 3 et «tout à fait importantes» par 4.
- $S_{Effect} = \{Q_i / i = 9, 10, 13, 15, 17, 18\}$. Puisque les questions appartenant à S_{Effet} sont des variables classes (connues en anglais par *target questions*), elles sont ainsi considérées comme des attributs nominaux.

2. Remplacement des valeurs manquantes

Afin de générer les modèles de décision, les valeurs manquantes (*i.e.*, les questions sans réponses) sont gérées pour les deux types d'attributs. Pour les attributs nominaux, les valeurs manquantes sont remplacées par la valeur nominale la plus fréquente et pour les attributs numériques, la valeur moyenne.

3. Sélection du modèle de décision

Suite aux résultats présentés dans la section 4.3.4, nous nous souhaitons à apporter une explication sur le comportement des usagers relatif au respect de la vie privée (Q_{10}) lors de l'utilisation de G-Face, la performance perçue (Q_{13}) et l'appréciation générale (Q_{18}) lors de l'utilisation des deux systèmes G-Face et G-Keystroke. Pour chacune de ces questions, nous avons ainsi généré les modèles de décision par réseau bayésien et arbre de décision. Le modèle retenu est le meilleur (en utilisant les critères de performance présentés dans la section 4.2.4.2) sur le jeu de données de 100 usagers.

En utilisant les modèles de décision générés, nous pouvons déduire les points suivants :

- En utilisant le modèle de décision présenté dans le tableau 4.3, l'inquiétude des usagers concernant la vie privée lors de l'utilisation du système G-Face est reliée à leur perception de sa robustesse contre la fraude. Parmi les usagers qui ont exprimé des inquiétudes sur la vie privée, la plupart d'entre eux ($65,96\%$) trouvent que le système peut être facilement contourné. Puisque plus de la moitié des usagers ($51,52\%$) ne considère pas que le système est robuste contre la fraude, ceci explique le taux élevé obtenu ($47,47\%$) concernant la protection de la vie privée. En utilisant la clause C_1 (tableau 4.3), nous pouvons également déduire que la facilité d'utilisation du système G-Face est une autre raison qui a influé la perception des usagers sur le respect de la vie privée.

- En utilisant le tableau 4.4, la performance perçue du système G-Face peut être expliquée par l'opinion des usagers sur sa robustesse contre la fraude et leurs connaissances concernant le vol d'identité. Tandis que pour G-Keystroke (tableau 4.5), la performance est reliée à sa robustesse contre la fraude et sa facilité d'utilisation. Nous pouvons déduire que la robustesse du système contre la fraude était un facteur principal influençant la performance perçue des deux systèmes testés. Puisque les usagers trouvent que G-Keystroke est plus robuste que G-Face (p-valeur=$0,005$) contre la fraude, et que la plupart

d'entre eux (85, 86%) considèrent que G-Kesytroke est facile à utiliser, ceci explique pourquoi la performance de G-Keystroke est mieux perçue que celle de G-Face.

– La robustesse du système contre la fraude et sa facilité d'utilisation sont deux facteurs majeurs influençant l'appréciation générale des usagers lors de l'utilisation des deux systèmes (*cf.*, tableaux 4.6 et 4.7). Puisque la plupart des volontaires trouvent que l'utilisation de G-Face est facile, nous pouvons déduire que sa robustesse contre la fraude est un facteur majeur influençant négativement leur appréciation générale sur le système. Tandis que pour G-Keystroke, sa facilité d'utilisation et sa robustesse perçue contre la fraude influencent positivement leur appréciation générale sur le système. Nous pouvons ainsi déduire que la robustesse des deux systèmes contre la fraude est un facteur majeur influençant leur appréciation générale lors de l'utilisation des deux systèmes testés.

Si (Le système peut être facilement contourné <=2) alors
... ...
Si (Le système peut être facilement contourné >2) alors
... ...
Si (Le système peut être facilement contourné <=3) alors
Si (Solutions basées sur la biométrie contre la fraude <=2) alors
... ...
Si (Solutions basées sur la biométrie contre la fraude >2) alors
Si (Vérification rapide <=3) alors
Si (Utilisation facile <=3) alors plutôt intrusive (20.0/3.0) (C_1)
Si (Utilisation facile >3) alors
... ...
Si (Solutions basées sur un secret contre la fraude >2) alors plutôt intrusive (5.0/2.0)
Si (Vérification rapide >3) alors
Si (Connaissances concernant le vol d'identité <=2) alors
... ...
Si (Utilisation facile >3) alors plutôt intrusive (6.0/3.0)
Si (Connaissances concernant le vol d'identité >2) alors plutôt intrusive (6.0/2.0)
Si (Le système peut être facilement contourné >3) alors
Si (Utilisation facile <=3) alors tout à fait intrusive (5.0/1.0)
... ...
Précision : 76.77%
AUC :
pas du tout intrusive : 0.94, plutôt pas intrusive : 0.938, plutôt intrusive : 0.916 et tout à fait intrusive : 0.919

TAB. 4.3: Extrait de l'arbre de décision expliquant les réponses des usagers sur l'impact à la vie privée lors de l'utilisation de G-Face. Les règles en gras indiquent les explications utiles.

Si (Utilisation facile <=3) alors Si (Connaissances concernant le vol d'identité <= 2) alors Si (Connaissances sur la technologie biométrique <= 1) alors Si (Le système peut être facilement contourné <= 2) alors toujours (6.0/2.0) **Si (Le système peut être facilement contourné > 2) alors rarement (2.0)** Si (Connaissances sur la technologie biométrique > 1) alors ... ... **Si (Connaissances concernant le vol d'identité <= 1) alors rarement (4.0/1.0)** ... ... Si (Connaissances concernant le vol d'identité > 2) alors ... ... Si (Utilisation facile > 3) alors Si (Le système peut être facilement contourné <= 2) alors toujours (19.0/2.0) **Si (Le système peut être facilement contourné > 2) alors** ... ... **Si (Solutions basées sur un secret contre la fraude > 2 alors rarement (3.0/1.0)** ... ...
Précision : 73.74% AUC : jamais : 0.962, rarement : 0.869, parfois : 0.877 et toujours : 0.901

TAB. 4.4: Extrait de l'arbre de décision expliquant la performance perçue par les usagers du système G-Face. Les règles en gras indiquent les explications utiles.

Si (Connaissances concernant le vol d'identité <=2) alors Si (Solutions basées sur la biométrie contre la fraude <= 2) alors rarement (2.0) Si (Solutions basées sur la biométrie contre la fraude >2) alors Si (Utilisation facile <=3) alors ... ... **Si (Utilisation facile >3) alors toujours (25.0/5.0)** Si (Connaissances concernant le vol d'identité >2) alors **Si (Le système peut être facilement contourné <=2) alors** Si (Solutions basées sur un secret contre la fraude <=1) alors ... ... **Si (Solutions basées sur un secret contre la fraude >1) alors toujours (24.0/3.0)** Si (Le système peut être facilement contourné >2) alors ... ...
Précision : 81.82% AUC : rarement : 0.894, parfois : 0.887 et toujours : 0.867

TAB. 4.5: Extrait de l'arbre de décision expliquant la performance perçue par les usagers du système G-Keystroke. Les règles en gras indiquent les explications utiles.

Si (Le système peut être facilement contourné <= 2) alors ... ... **Si (Le système peut être facilement contourné > 2) alors** Si (Utilisation facile <= 3) alors **Si (Solutions basées sur la biométrie contre la fraude <= 2 alors plutôt pas satisfait (2.0)** Si (Solutions basées sur la biométrie contre la fraude > 2) alors ... ... **Si (Utilisation facile <= 2) alors** **Si (Déjà entendu parlé =no) alors pas du tout satisfait (3.0)** ... ... Si (Utilisation facile > 2) alors ... ... **Si (Solutions basées sur un secret contre la fraude <= 1) alors plutôt pas satisfait (2.0)** **Si (Solutions basées sur un secret contre la fraude > 1) alors pas du tout satisfait (6.0/1.0)** ... ... Si (Utilisation facile > 3) alors plutôt satisfait (22.0/7.0)
Précision : 77.78% AUC : pas du tout satisfait : 0.958, plutôt pas satisfait : 0.872, plutôt satisfait : 0.761, et tout à fait satisfait : 0.738

TAB. 4.6: Extrait de l'arbre de décision expliquant l'appréciation générale des usagers lors de l'utilisation du système G-Face. Les règles en gras indiquent les explications utiles.

Si (Le système peut être facilement contourné <=1) alors **Si (Connaissances sur la technologie biométrique <=2) alors tout à fait satisfait (9.0/2.0)** **Si (Connaissances sur la technologie biométrique >2) alors plutôt satisfait (4.0)** Si (Le système peut être facilement contourné >1) alors ... ... **Si (Solutions basées sur un secret contre la fraude <=2) alors** **Si (Utilisation facile <=1) alors tout à fait satisfait (2.0/1.0)** **Si (Utilisation facile >1) alors plutôt satisfait (48.0/12.0)** Si (Solutions basées sur un secret contre la fraude >2) alors Si (Vérification rapide <=2) alors plutôt pas satisfait (4.0/2.0) Si (Vérification rapide >2) alors **Si (Utilisation facile <= 3) alors plutôt satisfait (10.0/1.0)** ... ...
Précision : 76.77% AUC : pas du tout satisfait : 0.964, plutôt pas satisfait : 0.953, plutôt satisfait : 0.817, et tout à fait satisfait : 0.814

TAB. 4.7: Extrait de l'arbre de décision expliquant l'appréciation générale des usagers lors de l'utilisation du système G-Keystroke. Les règles en gras indiquent les explications utiles.

4.3.7 Discussion

À partir des résultats présentés dans les deux sections 4.3.4 et 4.3.6, nous pouvons déduire plusieurs points. Premièrement, les usagers considèrent que les méthodes basées sur la biométrie sont plus pertinentes que celles basées sur un secret. Ceci montre que la biométrie est une solution acceptable qui va être de plus en plus utilisée dans les applications quotidienne. Deuxièmement, contrairement aux attentes, le système G-Keystroke (avec un EER égal à $17,51\%$) était plus acceptable que G-Face (avec un EER égal à $8,76\%$). Ceci montre que l'utilisation de la performance est nécessaire mais **pas suffisante** pour quantifier l'intérêt d'un nouveau système biométrique dans un cadre précis. Troisièmement, la facilité d'utilisation du système G-Face ainsi que sa robustesse contre la fraude sont deux facteurs majeurs sur la satisfaction des usagers lors de son utilisation, ainsi que son impact sur la vie privée ($47,47\%$). Pendant les expérimentations, la plupart des volontaires ont signalé que G-Face peut être facilement fraudé en présentant un photo devant le capteur d'acquisition. Nous trouvons ainsi qu'il faut intégrer à G-Face un détecteur de fakes [164] pour améliorer l'usage de ce système. Les résultats montrent également que la plupart des usagers trouvent que le temps de traitement des deux systèmes lors de la phase de vérification est raisonnable. La rapidité de traitement du système est un facteur important surtout dans le cas des systèmes d'identification.

4.4 Conclusion

Nous avons présenté dans ce chapitre une méthodologie générique permettant d'étudier l'acceptabilité et la satisfaction des usagers lors de l'utilisation de systèmes biométriques. La méthodologie proposée est basée sur une enquête et elle utilise 1) le test de KW pour extraire les dépendances significatives entre les caractéristiques socio-démographiques et les questions de satisfaction, et 2) les modèles de décision, les réseaux bayésiens et les arbres de décision, afin d'expliquer les réponses des usagers. Afin de montrer l'intérêt de la méthodologie proposée, nous l'avons appliqué sur deux systèmes d'authentification biométrique développés dans le laboratoire de recherche GREYC (GREYC-Face et GREYC-Keystroke). L'étude a été menée sur 100 volontaires. L'analyse de données montre que la perception des usagers lors de l'utilisation des systèmes biométriques doit être prise en considération lors du développement et d'évaluation des systèmes biométriques. Même si la performance d'un système est meilleure qu'un autre, cela n'implique pas forcément qu'il sera plus opérationnel ou acceptable. Enfin, un système performant en termes d'erreurs et d'acceptabilité mais pas robuste contre les divers types d'attaques, pose également

un problème. Nous présentons ainsi dans le chapitre suivant une méthode générique pour évaluer la sécurité des systèmes biométriques.

Chapitre 5

Évaluation de la sécurité d'un système biométrique

*Ce chapitre aborde l'évaluation des systèmes biométriques en terme de sécurité. La contribution de ce chapitre est double : nous proposons 1) une base commune d'attaques et de vulnérabilités des systèmes biométriques, et 2) une méthode générique (*i.e.*, indépendante de la modalité) pour évaluer quantitativement les systèmes biométriques. La méthode proposée est inspirée de l'approche des Critères Communs (ISO/CEI 15408) et de la méthode d'audit de sécurité EBIOS.*

Sommaire

5.1 Introduction

Malgré les avantages des systèmes biométriques par rapport aux systèmes traditionnels, ils sont vulnérables à des attaques spécifiques qui peuvent dégrader considérablement leur fonctionnalité et l'intérêt de déployer de tels systèmes. Ainsi, l'évaluation de la sécurité des systèmes biométriques est devenue indispensable pour

garantir l'opérationnalité de ces systèmes.

Dans le chapitre 2, nous avons vu que les travaux liés à l'évaluation des systèmes biométriques en terme de sécurité sont beaucoup moins abordés que ceux liés à la définition de scénarios d'attaques de tels systèmes. De plus, nous voyons que cet aspect d'évaluation est devenu indispensable surtout après la sortie de la norme ISO/IEC FCD 19792 [48] qui présentent les principales exigences à prendre en considération lors de l'évaluation des systèmes biométriques. Ainsi, notre contribution dans ce chapitre est double. Premièrement, nous proposons une base commune d'attaques et de vulnérabilités des systèmes biométriques. Cette base pourra ainsi être utilisée par les évaluateurs pour analyser (quantitativement ou qualitativement) le niveau de sécurité d'un système biométrique. Deuxièmement, nous proposons une méthode quantitative pour évaluer la sécurité des systèmes biométriques. La méthode proposée est indépendante de la modalité, et inspirée des Critères Communs (ISO/CEI 15408) [132] et de la méthode d'audit de sécurité EBIOS (Expression des Besoins et Identification des Objectifs de Sécurité) [131]. Dans la suite de ce chapitre, la section 5.2 décrit la méthode proposée. La base commune d'attaques et de vulnérabilités des systèmes biométriques est donnée dans la section 5.3. Les résultats expérimentaux sur deux systèmes d'authentification biométrique sont présentés dans la section 5.4. Enfin, la section 5.5 conclut ce chapitre.

5.2 Méthode développée

Selon l'Organisation Internationale de Normalisation ISO/IEC FCD 19792 [48], l'évaluation de la sécurité des systèmes biométriques est généralement composée de deux types d'évaluation complémentaires : **type 1)** évaluation du système biométrique (capteurs et algorithmes) et **type 2)** évaluation de l'environnement (est-ce que le système est utilisé en intérieur ou extérieur ?) et des conditions opérationnelles (politique d'administration telles que les mesures effectuées par les administrateurs du système pour n'enrôler que des utilisateurs légitimes, *etc.*). La méthode développée consiste à définir une méthode d'évaluation quantitative générique de **type 1** en utilisant une base commune d'attaques et de vulnérabilités, et la notion de facteurs de risque. La base utilisée s'appuie sur un modèle étendu (*cf.*, figure 5.1) de celle proposée par Ratha *et al.* [117]. Le principe de la méthode proposée repose sur quatre étapes : 1) étude du contexte, 2) expression des besoins de sécurité, 3) appréciation des risques et 4) calcul d'un indice de sécurité.

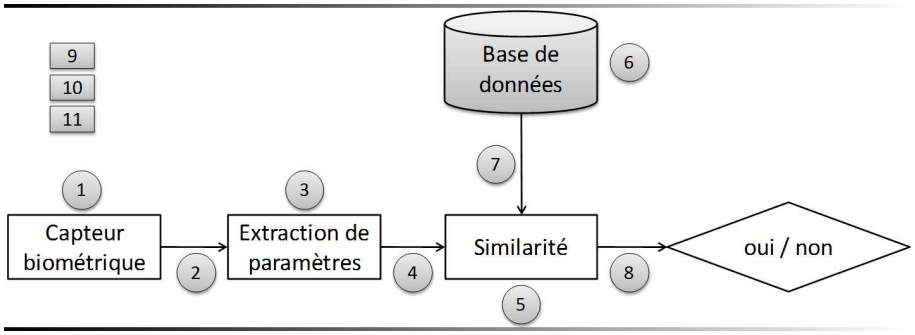


FIG. 5.1 – Modèle d'évaluation sécuritaire d'un système biométrique : les ◯ correspondent aux emplacements des attaques présentées par Ratha *et al.* [117], et les □ correspondent aux trois vulnérabilités globales ajoutées.

5.2.1 Étude du contexte

La première étape de la démarche consiste à identifier précisément l'utilité, l'architecture et le fonctionnement du système cible. Il s'agit également de décrire précisément les différents composants et éléments essentiels du système cible (appelé également les biens), qui dépendent généralement de la modalité biométrique considérée. On en retrouve de plusieurs natures, et certains sont généralement communs à tout système d'authentification biométrique. En utilisant l'architecture d'un système biométrique générique présentée dans la figure 5.1, nous avons choisi de protéger trois types de biens (information, fonction et matériel) d'un système biométrique générique. Les biens retenus sont décrits dans le tableau 5.1.

Référence	**Type**	**Description**
I_DONNEE_BIO	Information	Donnée biométrique de l'utilisateur qui souhaite s'authentifier
I_MODELE	Information	Modèle biométrique de l'utilisateur
I_DECISION	Information	Décision du système d'authentification (oui ou non)
F_TRAITEMENT	Fonction	Fonction de traitement des données biométriques issues du capteur biométrique
F_COMPARAISON	Fonction	Fonction de comparaison entre la donnée biométrique et le modèle biométrique
M_CAPTEUR	Matériel	Capteur biométrique
M_PROCESSUS	Matériel	Matériel sur lesquels les processus F_TRAITEMENT et F_COMPARAISON sont exécutés
M_BDD	Matériel	Support du stockage des modèles biométriques
M_CANAUX	Matériel	Canaux de transmission qui relient les différents composants du système

TAB. 5.1: Les biens d'un système d'authentification biométrique générique.

5.2.2 Expression des besoins de sécurité

Une fois le système précisément analysé et décrit, l'étape suivante consiste à répertorier les exigences en sécurité auxquelles il doit répondre. Ces exigences pourront ensuite être rapprochées des risques que présente le système pour pouvoir définir des contre-mesures de sécurité à mettre en place. Un système d'authentification biométrique présente dans la majorité des cas des besoins de sécurité génériques, parmi lesquels figurent la confidentialité, l'intégrité, la disponibilité et l'authenticité.

- La confidentialité (C) :
  Selon l'Organisation internationale de normalisation (ISO), la confidentialité est définie comme *le fait de s'assurer que l'information n'est seulement accessible qu'à ceux dont l'accès est autorisé.* La confidentialité garantit que les biens considérés ne sont accessibles qu'aux utilisateurs autorisés. Ce critère couvre également la notion de protection de la vie privée, qui est considérée comme un enjeu majeur en biométrie ;

- L'intégrité (I) :
  Ce critère garantit l'exactitude des biens considérés, et que les données biométriques n'ont pas été modifiées pendant le stockage ou la transmission ;

- La disponibilité (D) :
  Ce critère garantit l'accessibilité des biens considérés au moment voulu par les utilisateurs autorisés. Ce critère garantit également l'authentification des utilisateurs légitimes du système en question ;

- L'authenticité (A) :
  L'authentification est l'utilité principale d'un système biométrique pour que les utilisateurs légitimes puissent s'authentifier. Plus particulièrement, ce critère garantit que l'utilisateur qui présente la donnée biométrique est bien celui qu'il prétend être.

5.2.3 Appréciation des risques

Analyser les risques d'un système d'information (SI) est considéré comme un facteur indispensable à sa conception. L'analyse des risques repose sur deux phases [165] : la première consiste à identifier les causes des risques, tandis que la seconde détermine le niveau des risques identifiés (*i.e.*, dangerosité). Il existe plusieurs approches pour

analyser les risques d'un SI [166, 167, 168, 169], et elles sont généralement divisées en deux familles : quantitative et qualitative.

Dans le cadre de cette thèse, nous avons choisi d'utiliser une approche quantitative pour diverses raisons. Premièrement, c'est une approche qui permet d'estimer quantitativement les conséquences des risques identifiés, ce qui facilite nettement la comparaison des systèmes biométriques. Deuxièmement, l'approche quantitative est plus exploitable que l'approche qualitative pendant la phase de réduction des risques, puisque ces derniers sont quantitativement ordonnés selon leur degré de dangerosité. Une description détaillée des approches quantitatives et qualitatives est abordée par Rot [170], ainsi que leurs limites d'utilisation. La méthode ainsi proposée utilise la notion de facteurs de risque et la base commune d'attaques et de vulnérabilités présentée dans la section 5.3. Un facteur de risque, pour chaque attaque identifiée, est considéré comme un indicateur de sa dangerosité. Nous avons utilisé les facteurs de risque pour décrire quantitativement l'importance des attaques identifiées et des vulnérabilités globales retenues du système cible. Nous présentons ci-après le calcul des facteurs de risques des attaques identifiées, tandis que dans la section 5.2.3.2 sont décrits ceux liés aux vulnérabilités globales.

5.2.3.1 Facteurs de risque des attaques identifiées

Afin de calculer le facteur de risque pour chaque attaque identifiée, nous utilisons une approche quantitative inspirée de l'Analyse Multi-Critères (MCA) [171]. Plus spécifiquement, nous utilisons deux critères pour calculer un facteur de risque (f_r) pour chaque attaque identifiée :

$$f_r = c_1 \cdot c_2 \tag{5.1}$$

- La gravité (c_1) représente l'impact de l'attaque en terme de criticité. Ce facteur est défini sur une échelle entre 0 et 10 (le plus haut score 10 correspond à une attaque très critique). Il est arbitrairement fixé selon les besoins de sécurité retenus (confidentialité, intégrité, disponibilité et authenticité) sur les biens identifiés ;

- La facilité (c_2) représente la difficulté pour implémenter une attaque réussie. Ce facteur est défini sur une échelle entre 0 et 10 (le score 0 correspond à une attaque impossible, tandis que le score 10 correspond à une attaque très facile). Il est arbitrairement fixé selon les trois informations suivantes :

i) la vulnérabilité du système (failles liées à l'architecture de déploiement), ii) le coût en terme d'équipements spécifiques nécessaires pour implémenter l'attaque en question et iii) le niveau d'expertise requis (l'attaquant devrait maîtriser des connaissances techniques pour développer un programme malveillant).

5.2.3.2 Facteurs de risque des vulnérabilités globales

Pour les trois vulnérabilités globales retenues d'un système biométrique (*cf.*, section 5.3.2), nous avons développé un ensemble de règles pour estimer le facteur de risque associé comme décrit dans le tableau 5.2. Pour la performance du système, nous avons multiplié le taux d'erreur par deux puisque un système biométrique possédant un taux d'erreur (selon le système cible, EER ou HTER) supérieur ou égal à 50% n'est pas possible (pour un système d'authentification). Pour de tels systèmes, nous mettons ainsi le facteur de risque associé à 100. Pour la qualité, nous avons défini quatre règles selon le niveau de mise en oeuvre d'un contrôle de qualité du système pendant la phase d'enrôlement. Pour la base des modèles biométriques, nous avons également défini un ensemble de règles selon le degré de mise en oeuvre de mécanismes de protection (chiffrement et révocabilité des modèles biométriques) du système sur cet élément essentiel. Les quantifications du facteur de risque dans ce tableau sont certes arbitraires mais nous comptons solliciter la communauté en biométrie pour obtenir un consensus des experts.

Point	**Description**	**Conditions**	**Facteur de risque (f_r)**
9	Performance du système	Panel suffisant d'utilisateurs	2 * Taux d'erreur (limité à 100)
10	La qualité du modèle biométrique pendant l'enrôlement	- Plusieurs captures avec contrôle de qualité - Une seule capture avec contrôle de qualité - Plusieurs captures sans contrôle de qualité - Une seule capture sans contrôle de qualité	0 40 60 100
11	Mécanismes de protection de la base des modèles biométriques	- Base sécurisée et stockage local - Base sécurisée et stockage central - Base non sécurisée et stockage local - Base non sécurisée et stockage central	0 40 60 100

TAB. 5.2: Vulnérabilités globales des systèmes biométriques : les facteurs de risque.

5.2.4 Indice de sécurité

Estimer le niveau de sécurité global d'un système biométrique (plus généralement d'un SI), est une tâche difficile puisque le nombre d'acteurs impliqué dans le processus d'authentification est important. Dès lors, il est toujours avantageux d'illustrer la sécurité globale du système cible par un indice $(0 - 100)$ pour faciliter l'évaluation et la comparaison des tels systèmes [172]. Pour ce faire, nous avons utilisé la notion d'aire sous la courbe sur les facteurs de risque retenus pour calculer l'indice de sécurité globale du système cible. L'indice de sécurité est ainsi défini par :

$$\text{Indice} \;=\; \alpha \;\left(1 - \frac{AUC(f(x))}{AUC(g(x))}\right) = \alpha \;\left(1 - \frac{\int_1^n f(x)\,\mathrm{d}x}{\int_1^n g(x)\,\mathrm{d}x}\right) \tag{5.2}$$

où $\alpha = 100$, $n = r + s$, avec r le nombre d'emplacements des points de compromission (dans notre cas, $r = 8$) et s le nombre des vulnérabilités globales retenues (dans notre cas, $s = 3$). La fonction $f(x)$ représente la courbe obtenue à partir des facteurs de risque retenus sur les 11 points (le facteur de risque maximal est retenu à chaque point de compromission). Tandis que la fonction $g(x)$ représente la courbe obtenue à partir de l'ensemble des facteurs de risque le plus élevé qu'on peut avoir à chaque point de compromission (dans notre cas, ils sont égaux à 100). Plus l'indice est proche de 100, meilleure est la robustesse du système cible contre la fraude.

5.3 Base commune d'attaques et de vulnérabilités

La base d'attaques et de vulnérabilités prend en considération les attaques et vulnérabilités présentées dans [126, 11, 173, 174, 128], ainsi que les vulnérabilités présentées par l'Organisation Internationale de Normalisation ISO/IEC FCD 19792 [48]. La base contient également d'autre types d'attaques et de vulnérabilités qui nous ont paru pertinentes lors de la création des deux bases biométriques (*GREYC-Keystroke* et *ENSIB*), et l'étude portée sur l'usage des systèmes biométriques [175]. Nous présentons ci-après la base commune d'attaques identifiées, tandis que les vulnérabilités globales sont données dans la section 5.3.2.

5.3.1 Attaques des systèmes biométriques

La base commune d'attaques identifiées aux huit emplacements de compromission (*cf.*, figure 5.1) est donnée sous la forme suivante : «description» de l'attaque ainsi que son «atteinte» sur les besoins de sécurité retenus dans la section 5.2.2. Cette illustration va nous permettre de quantifier la gravité (c_1) de chaque attaque identifiée

lors de l'évaluation de la sécurité du système cible.

Point 1. Capteur

A_{11} – *Description* L'attaquant présente au capteur une fausse donnée biométrique (*e.g.*, doigts prothétiques) pour usurper l'identité d'un utilisateur légitime.
– *Atteintes* Authenticité sur I_DECISION.

A_{12} – *Description* L'attaquant exploite la similarité des données biométriques (le cas de jumeaux identiques et les systèmes biométriques utilisant des modalités spécifiques tels que le visage et l'ADN).
– *Atteintes* Authenticité sur I_DECISION.

A_{13} – *Description* Les utilisateurs légitimes fournissent volontairement leur donnée biométrique à l'attaquant (photo d'iris de bonne qualité).
– *Atteintes* Authenticité sur I_DECISION.

A_{14} – *Description* L'attaquant fournit sa propre donnée biométrique (tentative zéro effort) pour usurper l'identité d'un utilisateur légitime. En général, les attaquants choisissent des victimes ayant un faible modèle biométrique (image d'un visage de mauvaise qualité).
– *Atteintes* Authenticité sur I_DECISION.

A_{15} – *Description* L'attaquant récupère et exploite une donnée biométrique résiduelle (image d'une empreinte) sur le capteur afin d'usurper l'identité du dernier utilisateur authentifié.
– *Atteintes* Confidentialité sur I_DONNEE_BIO ; Authenticité sur I_DECISION.

A_{16} – *Description* L'attaquant dégrade le capteur biométrique pour le mettre hors d'état de fonctionnement.
– *Atteintes* Disponibilité sur M_CAPTEUR.

Points 2 et 4. Canaux de transmission

A_{241} – *Description* L'attaquant intercepte et rejoue une donnée biométrique à partir d'un canal de transmission.
– *Atteintes* Confidentialité sur I_DONNEE_BIO ; Authenticité sur I_DECISION.

A_{242} – *Description* L'attaquant détruit le canal de transmission afin de rendre le système indisponible pour les utilisateurs légitimes.
– *Atteintes* Disponibilité sur M_CANAUX.

A_{243} – *Description* L'attaquant altère l'information transportée sur le canal pour empêcher les utilisateurs légitimes de s'authentifier.
– *Atteintes* Intégrité sur I_DONNEE_BIO ; Intégrité sur M_CANAUX.

A_{244} – *Description* L'attaquant tente en permanence de s'authentifier par le système, en injectant des données au module de traitement (image d'une empreinte) ou au module de comparaison (minuties) [123].
– *Atteintes* Authenticité sur I_DECISION.

A_{245} – *Description* L'attaquant injecte en permanence des données pour rendre le système inaccessible aux utilisateurs légitimes.
– *Atteintes* Disponibilité sur M_CANAUX.

Points 3 et 5. Les composants logiciels

A_{351} – *Description* Les composants logiciels du système peuvent être remplacés par un programme du type cheval de Troie qui fonctionne selon les spécifications de ses concepteurs.
– *Atteintes* Confidentialité sur I_DONNEE_BIO ; Confidentialité sur I_MODELE ; Disponibilité sur F_TRAITEMENT ; Disponibilité sur F_COMPARAISON.

Point 6. Base des modèles biométriques

A_{61} – *Description* L'attaquant accède en mode lecture à la base des modèles biométriques.
– *Atteintes* Confidentialité sur I_MODELE ; Authenticité sur I_DECISION.

A_{62} – *Description* L'attaquant modifie (ajout, remplacement ou suppression) les modèles biométriques de la base.
– *Atteintes* Disponibilité sur I_MODELE ; Intégrité sur I_MODELE.

Point 7. Canal de transmission

A_{71} – *Description* L'attaquant intercepte un modèle biométrique à partir du canal de transmission afin d'être rejoué.

– *Atteintes* Confidentialité sur I_MODELE ; Authenticité sur I_DECISION.

A_{72} – *Description* L'attaquant altère les modèles biométriques transportés sur le canal pour empêcher les utilisateurs légitimes de s'authentifier.
– *Atteintes* Intégrité sur I_MODELE ; Intégrité sur M_CANAUX.

A_{73} – *Description* L'attaquant détruit le canal de transmission afin de rendre le système indisponible pour ses utilisateurs légitimes.
– *Atteintes* Disponibilité sur M_CANAUX.

Point 8. Canal de transmission

A_{81} – *Description* L'attaquant altère le résultat transporté (oui/non) pour empêcher l'accès d'un utilisateur légitime, ou ouvrir l'accès d'un imposteur.
– *Atteintes* Intégrité sur I_DECISION ; Authenticité sur I_DECISION.

A_{82} – *Description* L'attaquant détruit le canal de transmission afin de rendre le système indisponible pour ses utilisateurs légitimes.
– *Atteintes* Disponibilité sur M_CANAUX.

5.3.2 Vulnérabilités globales des systèmes biométriques

Point 9. Limites de performance

En comparaison aux systèmes d'authentification traditionnels qui offrent une réponse binaire (oui ou non), les systèmes biométriques sont moins précis et sont soumis à des erreurs telles que les taux de fausses acceptations (FAR) et de faux rejets (FRR). Cette variation illustrée par les taux d'erreurs peut affecter les systèmes biométriques en terme de sécurité. Doddington *et al.* [176] divisent les utilisateurs légitimes en quatre catégories que sont les moutons, les agneaux, les chèvres et les loups. Les moutons sont ceux qui peuvent être facilement reconnus (ils contribuent à une faible valeur du FRR). Les agneaux sont ceux qui peuvent être facilement imités (ils contribuent à un FAR élevé). Les chèvres sont ceux qui peuvent être difficilement reconnus (ils contribuent à un FRR élevé). Les loups sont ceux qui ont la capacité d'usurper facilement d'autres utilisateurs légitimes (ils contribuent

à un FAR élevé). Ainsi, un système biométrique peu efficient en terme de performance, peut être vulnérable face aux agneaux et loups. Dès lors, il est indispensable de prendre en considération la performance du système dans le processus d'évaluation.

Point 10. Limites de qualité pendant la phase d'enrôlement

La qualité des données acquises pendant la phase d'enrôlement est un facteur important à prendre en compte lors du développement des systèmes biométriques. L'absence d'un test de qualité augmente la possibilité d'avoir de faibles modèles biométriques. Ces modèles augmentent nettement la probabilité de réussite des attaques par zéro effort, hill-climbing et force brute [123].

Point 11. Mécanismes de protection des modèles biométriques

L'utilisation de la biométrie présente des vulnérabilités en termes de respect des droits et des libertés fondamentales. Le fait de conserver des modèles biométriques dans une base de données centrale constitue une invasion de la vie privée. Ces données sont donc des données sensibles, qui ne sont pas encore protégées de façon spécifique par une norme internationale (même si la norme ISO/IEC 27000 [177] adresse la protection des données personnelles). Parmi les solutions envisagées, on peut rendre les bases de données anonymes, et plus généralement intégrer la notion de respect de la vie privée dès la conception du système biométrique. Une autre solution plus efficace est d'utiliser le concept de biométrie révocable [178]. Il s'agit de transformer les données biométriques brutes, à l'aide d'une fonction choisie, de telle sorte que les données transformées soient sûres, révocables et respectent la vie privée des utilisateurs (intraçabilité par exemple).

5.4 Résultats expérimentaux

5.4.1 Étude du contexte et besoins de sécurité

Nous avons utilisé deux systèmes d'authentification biométrique pour montrer l'intérêt de la méthode proposée. Le premier système est le logiciel GREYC-Keystroke présenté dans le chapitre 1 section 1.3.2.1. Le deuxième, est le système Fingerprint lock présenté dans le chapitre 1 section 1.3.3.2. Nous avons choisi de comparer ces deux systèmes différents (en terme de conception) puisque les résultats de leur évaluation et leur comparaison devraient pouvoir être aussi différents que possible.

Les biens ainsi que les besoins de sécurité retenus sont ceux présentés dans le tableau 5.1 et la section 5.2.2, respectivement.

5.4.2 Appréciation des risques

Afin d'analyser les risques sur les deux systèmes cibles présentés dans la section précédente, nous avons utilisé la base commune d'attaques et de vulnérabilités des systèmes biométriques (*cf.*, section 5.3), et la notion de facteurs de risque (*cf.*, section 5.2.3). Les deux tableaux 5.3 et 5.4, présentent l'analyse des deux systèmes cibles GREYC-Keystroke et Fingerprint lock, respectivement. Nous avons mis le symbole «-» dans les trois dernières lignes de chacun de ces deux tableaux, puisque les facteurs de risque des vulnérabilités globales retenues sont évalués à l'aide des règles comme le montre le tableau 5.2.

Pour les attaques possibles sur le capteur (point 1), nous avons identifié trois attaques possibles (A_{14}, A_{15}, A_{16}) sur le système GREYC-Keystroke, et quatre ($A_{11}, A_{13}, A_{15}, A_{16}$) pour Fingerprint lock. Pour le système cible GREYC-Keystroke, l'attaque A_{13} n'était pas possible puisque un utilisateur légitime ne peut pas donner à un imposteur sa façon de taper sur un clavier. Pour Fingerprint lock, l'attaque A_{12} n'était pas possible puisque l'empreinte digitale est unique pour chaque personne (même pour le cas de jumeaux identiques [179]). Prenons le cas de l'attaque de type écoute et rejoue (A_{15}) présente dans les deux systèmes cibles. Pour le facteur «gravité (c_1)» de l'attaque A_{15}, nous avons mis les valeurs 8 et 10 pour le système GREYC-Keystroke et Fingerprint lock, respectivement. Nous avons mis la valeur 8, puisque l'attaquant ne récupère que des événements de temps, qui n'a pas d'impact tangible sur la vie privée. Tandis que dans le cas du système Fingerprint lock, nous l'avons mis à 10 puisque l'accès à la donnée biométrique brute constitue une intrusion dans la vie privée. Pour le facteur «facilité (c_2)» de l'attaque A_{15}, nous avons mis les valeurs 3 et 6 pour le système GREYC-Keystroke et Fingerprint lock, respectivement. Nous avons plus pénalisé le système Fingerprint lock que le système GREYC-Keystroke, puisque c'est beaucoup plus facile de récupérer une image d'empreinte digitale résiduelle sur le capteur que de récupérer des événements de temps. Henniger *et al.* [128] et Marcela Espinoza [125] montrent clairement la facilité de récupérer une image résiduelle d'une empreinte digitale pour usurper l'identité d'un utilisateur légitime.

La figure 5.2 illustre une étude comparative des facteurs de risque (le facteur de risque maximal est retenu à chaque point de compromission) entre les deux systèmes cibles. En utilisant cette figure, nous pouvons déduire plusieurs résultats tels que :

Fingerprint lock est plus vulnérable au point 1 que GREYC-Keystroke, les deux systèmes ne sont pas vulnérables au point 4, et le système GREYC-Keystroke est plus vulnérable que le deuxième système aux points de compromission $2, 3, 5, 6, 7, 8$ et 9. D'autre part, en utilisant les facteurs de risque ainsi calculés et l'équation 5.2, les indices du sécurité du système GREYC-Keystroke et Fingerprint lock sont ainsi égaux à $56,7\%$ et 86%, respectivement. Ces résultats montrent que le système GREYC-Kesytroke est globalement plus vulnérable aux attaques que le système Fingerprint lock. Ce résultat pouvait être attendu dans la mesure ou aucune mesure particulière n'a été prise pour sécuriser GREYC-Keystroke dont sa vocation est de fournir un logiciel de démonstration d'une modalité biométrique peu répandue.

Enfin, puisque le système Fingerprint lock est une boîte noire (*i.e.*, pas de documentation), nous n'avons pas pu identifier d'attaques sur quelques points de compromission. Malgré ça, un attaquant pourrait être en mesure de les trouver, grâce à du reverse engineering (matériel et logiciel). Cependant, l'utilisation du système commercial dans cette étude est retenue afin de montrer l'intérêt de la méthode proposée pour évaluer et comparer les systèmes d'authentification biométrique. D'une manière plus générale, lors de l'évaluation d'un système biométrique, les concepteurs doivent fournir tous les détails (architecture, caractéristiques, *etc.*) du système cible aux évaluateurs.

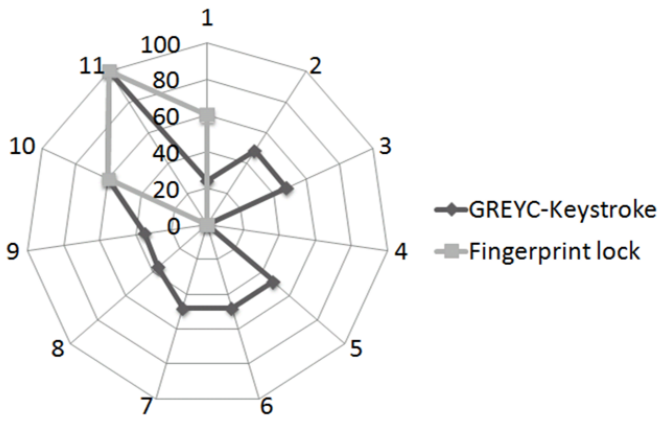


FIG. 5.2 – Une illustration comparative des deux systèmes cibles selon notre modèle d'analyse sécuritaire : huit points de compromission et trois vulnérabilités globales d'un système biométrique générique.

Point	Attaque	C	I	D	A	Gravité (c_1)	Facilité (c_2)	Risque (f_r)
1	A_{14}				×	6	2	12
	A_{16}			×		2	10	20
	A_{15}	×			×	8	3	24
2	A_{245}			×		2	6	12
	A_{243}		×			2	6	12
	A_{242}			×		2	10	20
	A_{244}				×	6	4	24
	A_{241}	×			×	8	6	48
3	A_{351}	×		×		8	6	48
5	A_{351}	×		×		8	6	48
6	A_{62}		×	×		8	4	32
	A_{61}	×			×	8	6	48
7	A_{72}		×			2	6	12
	A_{73}			×		2	10	20
	A_{71}	×			×	8	6	48
8	A_{82}			×		2	10	20
	A_{81}		×		×	6	6	36
9	Performance du système				×	-	-	35.02
10	Plusieurs captures sans contrôle de qualité				×	-	-	60
11	Base non sécurisée et stockage centrale	×	×	×	×	-	-	100

TAB. 5.3: Analyse sécuritaire du système GREYC-Keystroke (C : Confidentialité, I : Intégrité, D : Disponibilité, A : Authenticité).

Point	Attaque	C	I	D	A	Gravité (c_1)	Facilité (c_2)	Risque (f_r)
1	A_{16}			×		2	10	20
	A_{11}				×	6	8	48
	A_{13}				×	6	8	48
	A_{15}	×			×	10	6	60
9	Performance du système				×	-	-	0.1
10	Plusieurs captures sans contrôle de qualité				×	-	-	60
11	Base non sécurisée et stockage centrale	×	×	×	×	-	-	100

TAB. 5.4: Analyse sécuritaire du système Fingerprint lock (C : Confidentialité, I : Intégrité, D : Disponibilité, A : Authenticité).

5.5 Conclusion

L'évaluation des systèmes biométriques est considérée comme un enjeu majeur en biométrie. Malgré les travaux d'évaluation existants, peu d'études se sont focalisées sur l'évaluation de ces systèmes en terme de sécurité. La principale contribution de ce chapitre est la présentation 1) d'une base commune d'attaques et de vulnérabilités des systèmes biométriques, 2) d'une méthode générique pour évaluer quantitativement la sécurité des systèmes biométriques. La méthode proposée est inspirée des Critères

Communs (ISO/CEI 15408) et de la méthode d'audit de sécurité EBIOS. Nous avons montré son intérêt pour évaluer et comparer les systèmes d'authentification biométrique.

Conclusions et perspectives

Bilan

Au cours de cette thèse, nous nous sommes intéressés à l'évaluation des systèmes biométriques. Nous avons cherché à concevoir une méthodologie d'évaluation d'un système biométrique générique. Après une présentation générale de cette thèse, une présentation sur la biométrie, ses caractéristiques, ses limitations, ainsi que les enjeux liés à ses évaluations est effectuée dans le chapitre 1. Une présentation des travaux de la littérature liés aux aspects d'évaluation des systèmes biométriques est fournie dans le chapitre 2. Ces deux premiers chapitres définissent le contexte de cette thèse, ainsi que les outils préalables à notre étude. Nous avons ensuite présenté les trois contributions de cette thèse liées à la qualité de données biométriques morphologiques (chapitre 3), l'usage (chapitre 4) et la sécurité (chapitre 5) des systèmes biométriques.

La première contribution, présentée dans le chapitre 3, concerne la proposition d'une méthode d'évaluation de la qualité de données biométriques morphologiques. La méthode présentée utilise deux types d'informations complémentaires : 1) la qualité de l'image et 2) la qualité des paramètres extraits en utilisant le descripteur SIFT. La méthode proposée est plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé. Nous l'avons validé sur quatre bases publiques de visages (*FACES94*, *ENSIB*, *FERET* et *AR*) et une d'empreintes digitales ($FVC2002\ DB_2$). Les résultats expérimentaux montrent l'intérêt de la méthode pour détecter trois types d'altérations réelles des données (flou, bruit gaussien et redimensionnement) qui ont un impact majeur sur la performance globale des systèmes biométriques. Les résultats expérimentaux montrent également que la méthode proposée est plus efficace que la méthode NFIQ proposée par le NIST pour prédire les performances du système biométrique testé. Ces travaux sont publiés dans [180, 181, 182, 183].

La deuxième contribution, présentée dans le chapitre 4, est une méthodologie d'évaluation liée à l'usage des systèmes biométriques. Il s'agit de mesurer l'acceptabilité et la satisfaction des usagers lors de l'utilisation de tels systèmes. L'intérêt de la méthodologie présentée est double. Premièrement, nous avons utilisé un questionnaire pour recueillir les caractéristiques socio-démographiques, l'expérience et l'attitude des usagers qui pourraient avoir un impact sur leur satisfaction. Le questionnaire a été créé en collaboration avec deux experts en psychologie (Patrice Georget et Cécile Sénémeaud) de l'UCBN. Deuxièmement, nous avons proposé une méthode pour évaluer l'usage d'un système biométrique qui s'articule autour de trois étapes : la collection de données, la préparation des données collectées et l'analyse de données. La méthodologie présentée est indépendante de la modalité biométrique considérée. Deux systèmes biométriques (GREYC-Kestroke et GREYC-Face) ont été utilisés avec un échantillon de 100 volontaires pour montrer l'intérêt de la méthode développée. Les expériences ont permis de montrer que, contrairement aux attentes, le système GREYC-Keystroke (avec un EER égal à $17,51\%$) est mieux perçu par les usagers que le système GREYC-Face (avec un EER égal à $8,76\%$). Ceci montre que l'usage final des systèmes biométriques est un facteur primordial à prendre en compte lors de la conception de tels systèmes. La méthodologie présentée permet non seulement une analyse statistique des réponses des volontaires, mais aussi une analyse explicative pour mieux comprendre leur perception. Cette méthodologie d'évaluation est présentée dans [175, 184].

La troisième contribution, présentée dans le chapitre 5, est liée à l'évaluation des systèmes biométriques en terme de sécurité. Il s'agit de mesurer la sécurité d'un système (algorithmes et capteurs) face à divers types d'attaques. La contribution de la méthode proposée est double. Premièrement, nous avons présenté une base commune d'attaques et de vulnérabilités des systèmes biométriques. La base présentée est indépendante de la modalité biométrique considérée, et pourra être ainsi utilisée par les évaluateurs indépendamment de la méthode d'évaluation retenue. Deuxièmement, une méthode générique pour évaluer quantitativement les systèmes biométriques est proposée. Cette méthode s'articule autour de quatre étapes : 1) étude du contexte, 2) expression des besoins de sécurité, 3) appréciation des risques et 4) calcul d'un indice de sécurité. La méthode proposée est inspirée des Critères Communs (ISO/CEI 15408) et de la méthode d'audit de sécurité EBIOS. La méthode présentée est indépendante de la modalité biométrique considérée. Nous avons montré son intérêt sur deux systèmes biométriques (GREYC-Face et un système commercial

d'empreintes digitales). Cette méthode d'évaluation est présentée dans [185, 186].

Perspectives

Les perspectives de cette thèse sont nombreuses. Nous les articulons autour de trois axes.

1. Plate-forme d'évaluation biométrique en ligne : il s'agit de rassembler les trois aspects d'évaluation (performance, usage et sécurité), ainsi que le module dédié à la qualité de données biométriques dans une plate-forme librement accessible en ligne à destination des chercheurs et industriels en biométrie. L'intérêt d'une telle plate-forme est double. Premièrement, elle permettra aux chercheurs d'évaluer leurs systèmes d'une manière simple pour les améliorer. Deuxièmement, cette plate-forme a pour objectif principal d'étendre la base commune d'attaques et de vulnérabilités présentée en utilisant les commentaires et les suggestions des chercheurs. Il faut noter que l'analyse sécuritaire de la plate-forme en ligne sera potentiellement impliquée dans le comité technique en biométrie (TC4) de IAPR [1] et en normalisation (AFNOR [2]). Quelques captures d'écran de la plate-forme en cours de développement liées à l'aspect d'analyse sécuritaire d'un système biométrique sont présentées aux figures 5.3, 5.4 et 5.5.

2. Évaluation de la qualité de données biométriques : dans nos travaux, nous avons proposé une méthode d'évaluation permettant de détecter trois types d'altérations réelles (flou, bruit gaussien et redimensionnement) de données biométriques morphologiques, qui ont un impact majeur sur la performance globale des systèmes biométriques. Nous comptons ajouter un sixième critère pour détecter l'altération par luminance, qui a un impact significatif sur la plupart des systèmes de reconnaissance faciale existants. Nous comptons également mesurer l'efficacité de cette méthode sur d'autres types de modalités, notamment les veines de la main, les plis du doigt et l'iris. De plus, la métrique NFIQ comparée avec notre méthode est basée sur la qualité des minuties extraites. Ainsi, une étude comparative entre la méthode proposée et NFIQ en utilisant d'autres systèmes d'authentification basés sur les minuties extraites (tels que le système de vérification BOZORTH3 [187] développé par le NIST) s'avère utile. Outre la qualité de données morphologiques, nous comptons travailler sur la

1. International Association for Pattern Recognition
2. Association Française de Normalisation

qualité de données comportementales (*ex.*, dynamique de frappe au clavier). Cette information est utile puisque l'acquisition de données comportementales est sensible à divers types d'artefact liés au comportement des usagers (individu coopératif) et au lieu d'acquisition (environnement contrôlé). À noter que les méthodes de qualité comportementales sont basées sur des calculs de statistiques sur les caractéristiques extraites (événements de temps, pression, *etc.*).

3. Évaluation de la sécurité des systèmes biométriques : dans nos travaux, nous avons proposé une méthode d'évaluation de **type 1** (algorithmes et capteurs). Nous comptons étendre cette méthode pour qu'elle prenne en considération le **type 2** d'évaluation complémentaire (évaluation de l'environnement et des conditions opérationnelles). Les graphes de privilèges proposés par Dacier et Deswarte [188] nous paraît utile pour une telle évaluation. La méthode d'évaluation de **type 1** proposée présente l'avantage d'être indépendante de la modalité biométrique considérée, et facile à utiliser puisque l'approche retenue est quantitative. Cependant, le calcul des facteurs de risque est arbitrairement fixé. De plus, la base d'attaques doit être complétée avec plusieurs attaques publiées par les chercheurs. Afin de résoudre ces deux limites, nous comptons faire une évaluation subjective de la méthode proposée. Pour cela, nous avons prévu de faire évaluer par des experts les résultats d'analyse de sécurité (notamment, les valeurs subjectives de facteurs de risque). La comparaison des résultats d'analyse sécuritaire provenant des experts avec ceux obtenus par notre méthode d'évaluation nous permettra de l'améliorer. La plate-forme d'évaluation en ligne permet également de dépasser ces limites, puisque l'un de ses objectifs est d'étendre la base d'attaques présentée en utilisant les commentaires et les suggestions des chercheurs. La figure 5.5 présente une capture d'écran sur l'ajout d'une attaque par un contributeur sur le capteur biométrique. Nous comptons également implémenter plusieurs types d'attaques sur les capteurs afin de mesurer physiquement la robustesse des capteurs biométriques contre la fraude. Un exemple de type d'attaque est la présentation d'un faux doigt ou la présentation d'une image de visage de bonne qualité pour usurper l'identité d'un utilisateur légitime. Une base de contre-mesures pour la réduction des risques identifiés sera également utile.

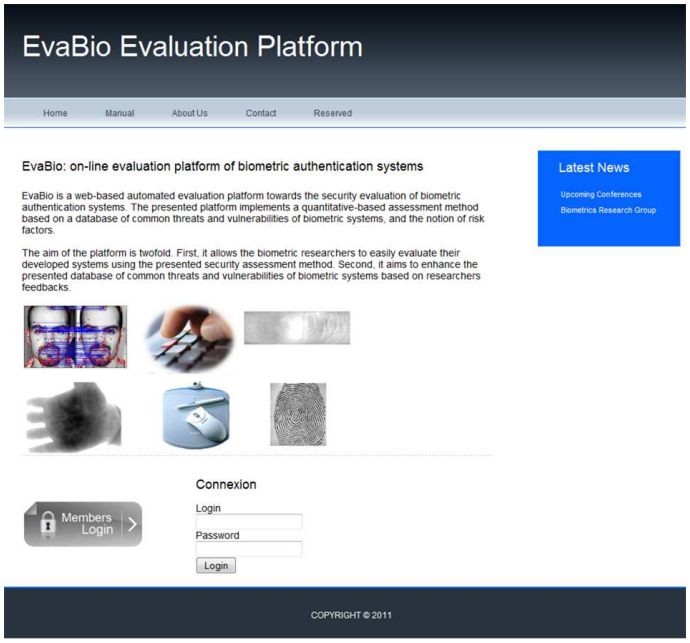


FIG. 5.3 – Plate-forme d'analyse sécuritaire d'un système biométrique.

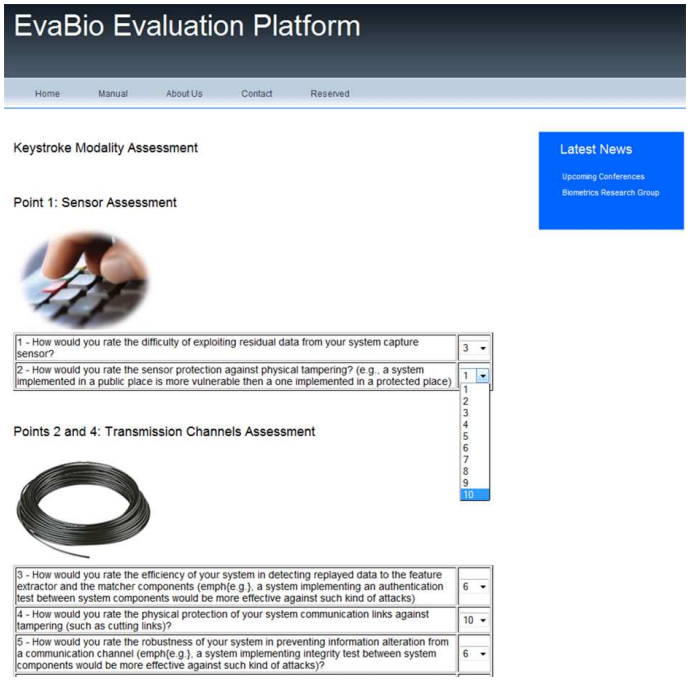


FIG. 5.4 – Analyse sécuritaire d'un système biométrique basé sur la dynamique de frappe au clavier.

FIG. 5.5 – Ajout d'une attaque par un contributeur sur le capteur d'un système biométrique basé sur la dynamique de frappe au clavier.

Publications de l'auteur

Chapitre de livre international

[1] Giot R., **El Abed M.**, Rosenberger C., «Keystroke dynamics», Intech Book on Biometrics, ISBN 978-953-307-618-8, p. 1-25, 2011.

Revues internationales à comité de lecture

[2] **El Abed M.**, Giot R., Hemery B., Rosenberger C., «Evaluation of Biometric Systems : A Study of Users' Acceptance and Satisfaction», Inderscience International Journal of Biometrics, p. 1-26, 2011.

[3] Mahier J., Hemery B., **El Abed M.**, El-Allam M. T., Bouhaddaoui M. Y., Rosenberger C., «Computation EvaBio : A Tool for erformance Evaluation in Biometrics», International Journal of Automated Identification Technology (IJAIT), p. 1-24, 2011.

[4] Giot R., **El Abed M.**, Hemery B., Rosenberger C., «Unconstrained Keystroke Dynamics Authentication with Shared Secret», Elsevier Journal of Computers & Security, p. 1-39, 2011.

Revues nationales avec comité de lecture

[5] **El Abed M.**, Hemery B., Charrier C., and Rosenberger C., «Evaluation de la qualité de données biométriques». Revue des Nouvelles Technologies de l'Information (RNTI), numéro spécial «Qualité des Données et des Connaissances /

Evaluation des méthodes d'Extraction de Connaissances dans les Données», p. 1-18, 2011.

[6] Pauchet A., **El Abed M.**, Merabti T., Prieur E., Lecroq T., Darmoni S.J., «Identification de répétitions dans les navigations au sein d'un catalogue de santé», numéro spécial RIA : Intelligence Artificielle et Web Intelligence, p. 113-132, 2009.

Conférences internationales avec comité de lecture et avec actes

[7] **El Abed M.**, Lacharme P., Rosenberger C., «Security EvaBio : An Analysis Tool for the Security Evaluation of Biometric Authentication Systems», the 5th IAPR/IEEE International Conference on Biometrics (ICB), New Delhi India, p. 1-8, 2012.

[8] **El Abed M.**, Giot R., Hemery B., Charrier C., Rosenberger C., «A SVM-Based Model for the Evaluation of Biometric Sample Quality», SSCI 2011 CIBIM - 2011 IEEE Workshop on Computational Intelligence in Biometrics and Identity Management, p. 115-122, 2011.

[9] Belguechi R., Cherrier E., **El Abed M.**, Rosenberger C., «Evaluation of Cancelable Biometric Systems : Application to Finger-Knuckle-Prints», IEEE International Conference on Hand-Based Biometrics, p. 1-6, 2011.

[10] Mahier J., **El Abed M.**, Hemery B., Rosenberger C., «Towards a Distributed Benchmarking Tool for Biometrics», IEEE International Conference on High Performance Computing & Simulation Conference (HPCS), p. 1-7, 2011.

[11] **El Abed M.**, Giot R., Hemery B., Schwartzmann J.J., Rosenberger C., «Towards the Security Evaluation of Biometric Authentication Systems», IEEE International Conference on Security Science and Technology (ICSST), Chongqing China, p. 167-173, 2011.

[12] **El Abed M.**, Giot R., Hemery B., Rosenberger C., «A Study of Users' Acceptance and Satisfaction of Biometric Systems», IEEE International Carnahan Conference on Security Technology (ICCST), San Jose USA, p.170-178, 2010.

[13] Giot R., **El Abed M**, Rosenberger C., «Fast Learning for Multibiometrics Systems Using Genetic Algorithms», IEEE International Conference on High Performance Computing & Simulation (HPCS), p. 266-273, 2010.

[14] Giot R., **El Abed M.**, Rosenberger C., «Keystroke Dynamics Authentication For Collaborative Systems», IEEE International Symposium on Collaborative Technologies and Systems (CTS), p. 172-179, 2009.

[15] Giot R., **El Abed M.**, Rosenberger C., «GREYC Keystroke : a Benchmark for Keystroke Dynamics Biometric Systems», IEEE International Conference on Biometrics : Theory, Applications and Systems (BTAS), p. 1-6, 2009.

[16] Giot R., **El Abed M.**, Rosenberger C., «Keystroke dynamics with Low Constraints SVM Based Passphrase Enrollment», IEEE International Conference on Biometrics : Theory, Applications and Systems (BTAS), p. 425-430, 2009.

Conférence internationale sans comité de lecture

[17] **El Abed M.**, Giot R., Rosenberger C., «Evaluation of Biometric Systems», poster in IEEE International Conference on High Performance Computing & Simulation Conference (HPCS), 2009.

Conférences nationales avec comité de lecture et avec actes

[18] **El Abed M.**, Hemery B., Charrier C., and Rosenberger C., «Un modèle SVM pour l'évaluation de la qualité des données biométriques». XXIII Colloque GRETSI, Bordeaux, p. 1-4, 2011. À paraître.

[19] **El Abed M.**, Giot R., Charrier C., Rosenberger C., «Evaluation of Biometric Systems : A SVM Based Quality Index», The third Norsk Information security conference (NISK), Norway, p. 1-12, 2010.

[20] Giot R., **El Abed M.**, Rosenberger C., «Authentification faiblement contrainte par dynamique de frappe au clavier», 17^{e} congrès francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA), p. 1-8, 2010.

Bibliographie

[1] S. G. Stubblebine and P. C. van Oorschot. Addressing online dictionary attacks with login histories and humans-in-the-loop. In *Financial Cryptography*, pages 39–53, 2004. [cité p. 1]

[2] A. K. Jain and A. Ross. Multibiometric systems. *Communications of the ACM*, 47 :34–40, 2004. [cité p. 2]

[3] N. K. Ratha, J. H. Connell, and R. M. Bolle. An analysis of minutiae matching strength. In *Audio- and Video-Based Biometric Person Authentication*, pages 223–228, 2001. [cité p. 3, 4, 22, 56]

[4] R. Cappelli, D. Maio, and D. Maltoni. Synthetic fingerprint-database generation. In *International Conference on Pattern Recognition (ICPR)*, pages 744–747, 2002. [cité p. 3, 40]

[5] ISO/IEC 19795-1. Information technology – biometric performance testing and reporting – part 1 : Principles and framework, 2006. [cité p. 3, 14, 20, 21, 22, 23, 26, 27, 33, 154]

[6] P. Grother and E. Tabassi. Performance of biometric quality measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 :531–543, 2007. [cité p. 4, 45, 71, 75]

[7] N. Poh, J.V. Kittler, and T. Bourlai. Quality-based score normalization with device qualitative information for multimodal biometric fusion. *IEEE Transactions on Systems, Man, and Cybernetics*, 40 :539–554, 2010. [cité p. 4, 23]

[8] A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, and A. Ross. Biometrics : A grand challenge. *International Conference on Pattern Recognition (ICPR)*, 2 :935–942, 2004. [cité p. 4, 50, 155]

[9] E. P. Kukula and R. W. Proctor. Human-biometric sensor interaction : Impact of training on biometric system and user performance. In *Proceedings of the Symposium*

on Human Interface 2009 on Human Interface and the Management of Information. Information and Interaction. Part II, volume 5618, pages 168–177, 2009. [cité p. 4, 49]

[10] E. P. Kukula, C. R. Blomeke, S. K. Modi, and S. J. Elliott. Effect of human-biometric sensor interaction on fingerprint matching performance, image quality and minutiae count. *International Journal of Computer Applications in Technology*, 34(4) :270–277, 2009. [cité p. 4, 50]

[11] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer-Verlag, 2003. [cité p. 4, 55, 56, 109, 149]

[12] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition : Security and privacy concerns. *IEEE Security & Privacy*, 1 :33–42, 2003. [cité p. 8]

[13] J. Mahier, M. Pasquet, C. Rosenberger, and F. Cuozzo. Biometric authentication. *Encyclopedia of Information Science and Technology*, pages 346–354, 2008. [cité p. 8, 9, 158]

[14] N. Rudin, K. Inman, G. Stolovitzky, and I. Rigoutsos. *Biometrics : Personal Identification in Networked Society*, chapter DNA Based Identification, pages 287–309. Kluwer Academic Publishers, 2002. [cité p. 8]

[15] International Biometric Group. http ://www.biometricgroup.com/, 2010. [cité p. 11, 15, 154]

[16] L. O'Gorman. Comparing passwords, tokens, and biometrics for user authentication. In *Proceedings of the IEEE*, volume 91, pages 2021–2040, 2003. [cité p. 11]

[17] A. K. Jain, L. Hong, and S. Pankanti. Biometrics : Promising frontiers for emerging identification market. Technical report, Department of Computer Science, Michigan State University, 2000. [cité p. 15]

[18] Z. Korotkaya. Biometric person authentication : Odor, 2003. [cité p. 15]

[19] M. Hashiyada. Development of biometric dna ink for authentication security. *Tohoku Journal of Experimental Medicine*, pages 109–117, 2004. [cité p. 15]

[20] K. Phua, J. Chen, T. H. Dat, and L. Shue. Heart sound as a biometric. *Pattern Recognition*, 41 :906–919, 2008. [cité p. 15]

[21] T. Artieres, J.-M. Marchand, P. Gallinari, and B. Dorizzi. Multi-modal segmental models for online handwriting recognition. In *15th International Conference on Pattern Recognition (ICPR)*, volume 2, pages 2247–2250, 2000. [cité p. 15]

[22] D. Muramatsu and T. Matsumoto. Effectiveness of pen pressure, azimuth, and altitude features for online signature verification. In *International Conference on Biometrics (ICB'07)*, volume 4642, pages 503–512, 2007. [cité p. 15]

[23] J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Transactions Pattern Analysis and Machine Intelligence (PAMI)*, 28 :316–322, 2006. [cité p. 15]

[24] D. Gafurov, E. Snekkenes, and T. E. Buvarp. Robustness of biometric gait authentication against impersonation attack. In *On the Move to Meaningful Internet Systems : OTM 2006 Workshops*, volume 4278, pages 479–488, 2006. [cité p. 15]

[25] F. Monrose and A. Rubin. Authentication via keystroke dynamics. In *Proceedings of the 4th ACM conference on Computer and communications security*, pages 48–56. ACM, 1997. [cité p. 15]

[26] S. Hocquet, J.Y. Ramel, and H. Cardot. User classification for keystroke dynamics authentication. In *International Conference on Biometrics (ICB)*, pages 531–539, 2007. [cité p. 15]

[27] R. Giot, M. El Abed, and C. Rosenberger. Keystroke dynamics with low constraints SVM based passphrase enrollment. In *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, pages 425–430, 2009. [cité p. 15]

[28] H. Harb and L. Chen. Voice-based gender identification in multimedia applications. *Journal of Intelligent Information Systems (JIIS)*, 24(2-3) :179–198, 2005. [cité p. 15]

[29] R. Giot, M. El Abed, and C. Rosenberger. Greyc keystroke : a benchmark for keystroke dynamics biometric systems. In *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, pages 1–6, 2009. [cité p. 15, 37, 43]

[30] S. Hocquet, J. Y. Ramel, and H. Cardot. Authentification par la dynamique de frappe. *15e congrès francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, 2005. [cité p. 16]

[31] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces : Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 :711–720, 1997. [cité p. 17]

[32] F. Perronnin, J. L. Dugelay, and K. Rose. Deformable face mapping for person identification. In *IEEE International Conference on Image Processing (ICIP)*, pages 14–17, 2003. [cité p. 17]

[33] B. Ben Amor, M. Ardabilian, and L. Chen. Enhancing 3D face recognition by mimics segmentation. In *Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA)*, pages 150–155, 2006. [cité p. 17]

[34] C. Samir, A. Srivastava, and M. Daoudi. 3D face recognition using shapes of facial curves. *IEEE Transactions Pattern Analysis and Machine Intelligence (PAMI)*, 28 :1858–1863, 2006. [cité p. 17]

[35] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle. An identity-authentication system using fingerprints. *Proceedings of the IEEE*, 85 :1365–1388, 1997. [cité p. 17]

[36] Y. Chen and A.K. Jain. Beyond minutiae : A fingerprint individuality model with pattern, ridge and pore features. In *International Conference on Biometrics (ICB)*, pages 523–533, 2009. [cité p. 17]

[37] A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain. Personal verification using palmprint and hand geometry biometric. In *Audio- and Video-based Biometric Person Authentication (AVBPA)*, pages 668–678, 2003. [cité p. 17]

[38] A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain. Personal authentication using hand images. *Pattern Recognition Letters*, 27 :1478–1486, 2006. [cité p. 17]

[39] R.P. Wildes, J.C. Asmuth, G.L. Green, S.C. Hsu, R.J. Kolczynski, J.R. Matey, and S.E. McBride. A system for automated iris recognition. In *IEEE Workshop on Applications of Computer Vision*, pages 121–128, 1994. [cité p. 17]

[40] J. Cui, Y. Wang, J. Huang, T. Tan, and Z. Sun. An iris image synthesis method based on PCA and super-resolution. In *17th International Conferenceon on Pattern Recognition (ICPR)*, pages 471–474, 2004. [cité p. 17]

[41] C. Rosenberger and L. Brun. Similarity-based matching for face authentication. *International Conference on Pattern Recognition (ICPR'08)*, pages 1–4, 2008. [cité p. 17]

[42] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV)*, 60 :91 – 110, 2004. [cité p. 17, 18, 49, 65, 67, 155]

[43] P.-O. Ladoux, C. Rosenberger, and B. Dorizzi. Palm vein verification system based on sift matching. In *the 3rd IAPR/IEEE International Conference on Biometrics (ICB'09)*, pages 1290–1298, 2009. [cité p. 18, 65]

[44] ANSI. Information technology - BioAPI Specification (version 1.1), 2002. [cité p. 20]

[45] ISO/IEC 19794-1. BioAPI 2.0, 2005. [cité p. 20]

[46] Fernando L. Podio, Jeffrey S. Dunn, Lawrence Reinert, Catherine J. Tilton, Lawrence O'Gorman, M. Paul Collier, Mark Jerde, and Brigitte Wirtz. Common biometric exchange file format (CBEFF) - NISTIR 6529, 2001. [cité p. 20]

[47] ANSI/NIST-ITL 1. Data format for the interchange of fingerprint, facial & other biometric information, 2011. [cité p. 20]

[48] ISO/IEC FCD 19792. Information technology – security techniques – security evaluation of biometrics, 2008. [cité p. 20, 21, 23, 58, 104, 109]

[49] ISO/IEC JRC 1/SC 37. Biometric data interchange formats - part 2 : Fingerprint minutiae data, 2004. [cité p. 20]

[50] ISO/IEC JRC 1/SC 37. Biometric Data Interchange Formats - Part 3 : Fingerprint Pattern Spectral Data, 2004. [cité p. 20]

[51] ISO/IEC JRC 1/SC 37. Biometric Data Interchange Formats - Part 4 : Finger Image Data, 2004. [cité p. 20]

[52] ISO/IEC JRC 1/SC 37. Biometric Data Interchange Formats - Part 5 : Face Image data, 2004. [cité p. 20]

[53] ISO/IEC JRC 1/SC 37. Biometric Data Interchange Formats - Part 6 : Iris Image Data, 2004. [cité p. 20]

[54] D. Dessimoz, C. Champod, J.Richiardi, and A. Drygajlo. Multimodal biometrics for identity documents. *Forensic Science International*, 167 :154–159, 2007. [cité p. 20]

[55] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of artificial "gummy" fingers on fingerprint systems. In *The International Society for Optical Engineering*, volume 4677, 2002. [cité p. 22]

[56] M. Theofanos, B. Stanton, and C. A. Wolfson. *Usability & Biometrics : Ensuring Successful Biometric Systems*. National Institute of Standards and Technology (NIST), 2008. [cité p. 23]

[57] P. J. Phillips, A. Martin, C. l. Wilson, and M. Przybocki. An introduction to evaluating biometric systems. *Computer*, pages 56 – 63, 2000. [cité p. 26]

[58] James P. Egan. Signal detection theory and ROC-analysis. by Academic Press, New York, 1975. [cité p. 26, 30]

[59] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The DET curve in assessment of detection task performance. In *the 5th European Conference on Speech Communication and Technology*, pages 1895 – 1898, 1997. [cité p. 26]

[60] J. Bhatnagar and A. Kumar. On estimating performance indices for biometric identification. *Pattern Recognition*, 42 :1803–1815, 2009. [cité p. 26, 33, 34, 37]

[61] F. Cherifi, B. Hemery, R. Giot, M. Pasquet, and C. Rosenberger. Performance evaluation of behavioral biometric systems. In *Behavioral Biometrics for Human Identification : Intelligent Applications*, pages 57–74, 2009. [cité p. 31, 37]

[62] D. Faraggi and B. Reiser. Estimation of the area under the ROC curve. *Statistics in medicine*, 21 :3093–3106, 2002. [cité p. 34]

[63] R. Tronci, G. Giacinto, and F. Roli. Designing multiple biometric systems : Measures of ensemble effectiveness. *Engineering Applications of Artificial Intelligence*, 22 :66–78, 2009. [cité p. 34, 37]

[64] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 1947. [cité p. 34]

[65] R. M. Bolle, N. K. Ratha, and S. Pankanti. Error analysis of pattern recognition systems : the subsets bootstrap. *Computer Vision and Image Understanding*, 93 :1 – 33, 2004. [cité p. 36]

[66] L. Allano. *La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles.* PhD thesis, Institu National des Télécommunications dans le cadre de l'école doctorale SITEVRY en co-accréditation avec l'Université d'Evry-Val d'Essonne, 2009. [cité p. 36]

[67] University of Essex. Faces94 database, face recognition data, 1994. [cité p. 37, 38, 154]

[68] A.M. Martinez and R. Benavente. The AR face database. *CVC Tech. Report*, 1998. [cité p. 37, 38, 154]

[69] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face recognition algorithms. *Journal of Image and Vision Computing*, 16 :295–306, 1998. [cité p. 37, 39, 155]

[70] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(10) :1094–1104, 2000. [cité p. 37]

[71] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. Fvc2002 : Second fingerprint verification competition. In *International Conference on Pattern Recognition (ICPR'02)*, volume 3, pages 811 – 814, 2002. [cité p. 37, 39, 155]

[72] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and 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)*, volume 1, pages 947–954, 2005. [cité p. 37]

[73] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer. The humanID gait challenge problem : data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(2) :162–177, 2005. [cité p. 37]

[74] B. Hemery, C. Rosenberger, and H. Laurent. The ENSIB database : a benchmark for face recognition. In *International Symposium on Signal Processing and its Applications (ISSPA), special session "Performance Evaluation and Benchmarking of Image and Video Processing"*, 2007. [cité p. 37, 40, 155]

[75] K. Messer, J. Matas, J.V. Kittler, J. Luettin, and G. Maitre. XM2VTSDB : The Extended M2VTS Database. In *Proc. Second International Conference on Audio-*

and Video-based Biometric Person Authentication (AVBPA'99), pages 72–77, 1999. [cité p. 37]

[76] V. Popovici, J. Thiran, E. Bailly-Bailliere, S. Bengio, F. Bimbotand M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messerand B. Ruiz, and F. Poiree. The BANCA database and evaluation protocol. In *4th International Conference on Audio- and Video-Based Biometric Person Authentication*, volume 2688, pages 625–638, 2003. [cité p. 37]

[77] BIOSECURE. Biosecure Multimodal Biometric Database. http ://www.biosecure.info/, 2008. [cité p. 38, 42, 150]

[78] R. Cappelli. Use of synthetic data for evaluating the quality of minutia extraction algorithms. In *Second NIST Biometric Quality Workshop*, 2007. [cité p. 40, 155]

[79] D.Y. Yeung, H. Chang, Y.M. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll. SVC2004 : First International Signature Verification Competition. In *International Conference on Biometric Authentication (ICBA'04)*, pages 16 – 22, 2004. [cité p. 41, 150]

[80] J. P. Phillips, T. W. Scruggs, A. J. O'toole, P. J. Flynn, K.W. Bowyer, C. L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale results. Technical report, National Institute of Standards and Technology, 2007. [cité p. 41, 150]

[81] A. Mayoue, B. Dorizzi, L. Allano, G. Chollet, J. Hennebert, D. Petrovska-Delacrétaz, and F. Verdet. *Guide to biometric reference systems and performance evaluation*, chapter The BioSecure multimodal evaluation campaign 2007 (BMEC'2007), pages 327–372. 2009. [cité p. 42]

[82] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'Toole, D. S. Bolme, K. W. Bowyer, B. A. Draper, G. H. Givens, Y. M. Lui, H. Sahibzada, J. A. Scallan, and S. Weimer. Overview of the multiple biometrics grand challenge. In *International Conference on Biometrics (ICB'09)*, pages 705 – 714, 2009. [cité p. 42]

[83] D. Petrovska and A. Mayoue. Description and documentation of the biosecure software library. Technical report, BioSecure, 2007. [cité p. 42]

[84] ISO/IEC FCD 19794-2. Information technology – biometric data interchange format – part 2 : Finger minutiae data, 2004. [cité p. 43]

[85] R. Cappelli, D. Maio, D. Maltoni, J. L. Wayman, and A. K. Jain. Performance evaluation of fingerprint verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28 :3–18, 2006. [cité p. 44]

[86] F. A. Fernandez. *Biometric Sample Quality and its Application to Multimodal Authentication Systems*. PhD thesis, Technical University of Madrid, 2008. [cité p. 45, 47, 48, 158]

[87] ISO/IEC 29794-1. Information technology – biometric sample quality – part 1 : Framework, 2009. [cité p. 45]

[88] J. Fierrez Aguilar, Y. Chen, J. Ortega Garcia, and A.K. Jain. Incorporating image quality in multi-algorithm fingerprint verification. In *International Conference on Biometrics (ICB)*, volume 3832, pages 213–220, 2006. [cité p. 45]

[89] N. Poh, T. Bourlai, J. Kittler, L. Allano, F. Alonso-Fernandez, O. Ambekar, J. Baker, B. Dorizzi, O. Fatukasi, J. Fierrez, H. Ganster, J. Ortega-Garcia, D. Maurer, A. A. Salah, T. Scheidat, and C. Vielhauer. Benchmarking quality-dependent and cost-sensitive score-level multimodal biometric fusion algorithms. *Transactions on Information Forensics and Security*, 4(4) :849–866, 2009. [cité p. 45]

[90] N. Poh, T. Bourlai, and J. Kittler. A multimodal biometric test bed for quality-dependent, cost-sensitive and client-specific score-level fusion algorithms. *Pattern Recognition*, pages 1094–1105, 2010. [cité p. 45]

[91] T. Scheidat, A. Makrushin, and C. Vielhauer. Automatic template update strategies for biometrics. Technical report, Advanced Multimedia and Security Lab, 2007. [cité p. 46]

[92] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. A comparative study of fingerprint image-quality estimation methods. *IEEE Transactions on Information Forensics and Security*, 2 :734–743, 2007. [cité p. 48]

[93] L. Shen, A. C. Kot, and W. M. Koo. Quality measures of fingerprint images. In *Proceedings of the 3rd International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 266–271, 2001. [cité p. 48]

[94] Y. Chen, S. C. Dass, and A. K. Jain. Fingerprint quality indices for predicting authentication performance. In *5th International Conference Audio- and Video-Based Biometric Person Authentication (AVBPA)*, volume 3546, pages 160–170, 2005. [cité p. 48]

[95] S. Lee, C. Lee, and J. Kim. Model-based quality estimation of fingerprint images. In *IAPR/IEEE International Conference on Biometrics (ICB'06)*, pages 229–235, 2006. [cité p. 48]

[96] E. Tabassi and C.L. Wilson. A novel approach to fingerprint image quality. In *International Conference on Image Processing (ICIP)*, pages 37–40, 2005. [cité p. 48, 75]

[97] E. Krichen, S. Garcia Salicetti, and B. Dorizzi. A new probabilistic iris quality measure for comprehensive noise detection. In *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, pages 1–6, 2007. [cité p. 48]

[98] Y. Chen, S.C. Dass, and A.K. Jain. Localized iris image quality using 2-d wavelets. In *International Conference on Biometrics (ICB)*, 2006. [cité p. 48]

[99] N. D. Kalka, J. Zuo, N. A. Schmid, and B. Cukic. Image quality assessment for iris biometric. In *Proc. SPIE 6202*, 2006. [cité p. 48]

[100] Q. He, Z.A. Sun, T.N. Tan, and Y. Zou. A hierarchical model for the evaluation of biometric sample quality. In *International Conference on Pattern Recognition (ICPR)*, pages 1–4, 2008. [cité p. 49]

[101] G. Zhang and Y. Wang. Asymmetry-based quality assessment of face images. In *Proceedings of the 5th International Symposium on Advances in Visual Computing (ISVC)*, volume 5876, pages 499–508, 2009. [cité p. 49]

[102] X.F. Gao, S.Z. Li, R. Liu, and P.R. Zhang. Standardization of face image sample quality. In *International Conference on Biometrics (ICB'07)*, pages 242–251, 2007. [cité p. 49]

[103] J. Sang, Z. Lei, and S. Z. Li. Face image quality evaluation for ISO/IEC standards 19794-5 and 29794-5. In *Proceedings of the Third International Conference on Advances in Biometrics (ICB)*, pages 229–238, 2009. [cité p. 49]

[104] F. Deane, K. Barrelle, R. Henderson, and D. Mahar. Perceived acceptability of biometric security systems. *Computers & Security*, 14 :225–231, 1995. [cité p. 51]

[105] ORC. Public Attitudes Toward the Uses of Biometric Identification Technologiesby Government and the Private Sector. Technical report, Opinion Research Corporation International (ORC), 2002. [cité p. 51]

[106] L. Coventry, A. De Angeli, and G. Johnson. Biometric verification at a self service interface. In *Contemporary ergonomics*, pages 247–252, 2003. [cité p. 51]

[107] L. Coventry, A. De Angeli, and G. Johnson. Honest it's me ! self service verification. In *The ACM Conference on Human Factors in Computing Systems (CHI)*, pages 1–4, 2003. [cité p. 51]

[108] L. Coventry, A. De Angeli, and G. Johnson. Usability and biometric verification at the ATM interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 153–160, 2003. [cité p. 51]

[109] M. Theofanos, B. Stanton, S. Orandi, R. Micheals, and N.F. Zhang. Usability testing of ten-print fingerprint capture. Technical report, National Institute of Standards and Technology (NIST), 2007. [cité p. 52, 53]

[110] M. Theofanos, B. Stanton, C. Sheppard, R. Micheals, N. Zhang, J. Wydler, L. Nadel, and W. Rubin. Usability testing of height and angles of ten-print fingerprint capture. Technical report, National Institute of Standards and Technology (NIST), 2008. [cité p. 52, 53]

[111] A. P. Pons and P. Polak. Understanding user perspectives on biometric technology. *Communications of the Association for Computing Machinery (ACM)*, 51(9) :115–118, 2008. [cité p. 52]

[112] L. A. Jones, A. I. Antón, and J. B. Earp. Towards understanding user perceptions of authentication technologies. In *ACM Workshop on Privacy in the Electronic Society*, pages 91–98, 2007. [cité p. 52]

[113] S. J. Elliott, S. A. Massie, and M. J. Sutton. The perception of biometric technology : A survey. *Automatic Identification Advanced Technologies*, pages 259–264, 2007. [cité p. 52]

[114] J. Moody. Public perceptions of biometric devices : The effect of misinformation on acceptance and use. In *the Informing Science and Information Technology Education*, volume 1, pages 753–761, 2004. [cité p. 52]

[115] R. R. Heckle, A. S. Patrick, and A. Ozok. Perception and acceptance of fingerprint biometric technology. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 153–154, 2007. [cité p. 52, 53]

[116] S. Schimke, C. Vielhauer, P.K. Dutta, T.K. Basu, A. De Rosa, J. Hansen, J. Dittmann, and B. Yegnanarayana. Cross cultural aspects of biometrics. In *Workshop Proceedings Biometric Challenges arising from Theory to Practice*, pages 27–30, 2004. [cité p. 52]

[117] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40 :614 – 634, 2001. [cité p. 54, 55, 104, 105, 155, 156]

[118] B. Schneier. The uses and abuses of biometrics. *Communications of the ACM*, 1999. [cité p. 55]

[119] T. V. der Putte and J. Keuning. Biometrical fingerprint recognition : Don't get your fingers burned. In *Proceedings of the Fourth Working Conference on Smart Card Research and Advanced Applications*, volume 31, pages 289–306, 2000. [cité p. 56]

[120] V. Ruiz-Albacete, P. Tome-Gonzalez, F. Alonso-Fernandez, J. Galbally, J. Fierrez, and J. Ortega-Garcia. Direct attacks using fake images in iris verification. In *Biometrics and Identity Management*, pages 181–190, 2008. [cité p. 56, 57]

[121] J. Galbally, R. Cappelli, A. Lumini, G. Gonzalez-de Rivera, D. Maltoni, J. Fierrez, J. Ortega-Garcia, and D. Maio. An evaluation of direct attacks using fake fingers generated from iso templates. *Pattern Recognition Letters*, 31 :725–732, 2010. [cité p. 56]

[122] A. Adler. Sample images can be independently restored from face recognition templates. *Electrical and Computer Engineering*, 2 :1163–1166, 2003. [cité p. 56, 57, 155]

[123] U. Uludag and A. K. Jain. Attacks on biometric systems : A case study in fingerprints. In *Proc. SPIE-EI 2004, Security, Seganography and Watermarking of Multimedia Contents VI*, volume 5306, pages 622–633, 2004. [cité p. 56, 111, 113]

[124] C. J. Hill. Risk of masquerade arising from the storage of biometrics. Master's thesis, The Department of Computer Science Australian National University, 2001. [cité p. 56]

[125] M. Espinoza, C. Champod, and P. Margot. Vulnerabilities of fingerprint reader to fake fingerprints attacks. *Forensic Science International*, 204 :41–49, 2010. [cité p. 57, 114]

[126] B. Schneier. Inside risks : the uses and abuses of biometrics. *Communications of the ACM*, 42, 1999. [cité p. 56, 109]

[127] B. Schneier. Attack trees. *Dr. Dobb's Journ. of Softw. Tools*, 1999. [cité p. 57]

[128] O. Henniger, D. Scheuermann, and T. Kniess. On security evaluation of fingerprint recognition systems. In *Internation Biometric Performance Testing Conference (IBPC)*, pages 1–10, 2010. [cité p. 57, 109, 114]

[129] V. Matyás and Z. Ríha. Biometric authentication - security and usability. In *Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security*, pages 227–239, 2002. [cité p. 57]

[130] C. Dimitriadis and D. Polemi. Application of multi-criteria analysis for the creation of a risk assessment knowledgebase for biometric systems. In *international conference on biometric authentication (ICB)*, volume 3072, pages 724–730, 2004. [cité p. 58]

[131] EBIOS. Expression des besoins et identification des objectifs de sécurité (EBIOS). Technical report, L'Agence nationale de la sécurité des systèmes d'information (ANSSI), 2004. [cité p. 58, 59, 104]

[132] CC. Common criteria for information technology security evaluation, 1999. [cité p. 59, 104]

[133] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment : From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4) :600–612, 2004. [cité p. 63]

[134] L. Qiang and Z. Wang. Reduced-reference image quality assessment using divisive normalization-based image representation. *IEEE Journal of Selected Topics in Signal Processing*, 3 :202–211, 2009. [cité p. 63]

[135] Z. Wang, H. R. Sheikh, and A. C. Bovik. No-reference perceptual quality assessment of jpeg compressed images. In *IEEE International Conference on Image Processing*, volume 1, pages 477–480, 2002. [cité p. 63]

[136] Z. Wang, A. C. Bovik, and B. L. Evans. Blind measurement of blocking artifacts in images. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 981–984, 2000. [cité p. 63]

[137] Z. M. Parvez Sazzad, Y. Kawayoke, and Y. Horita. No-reference image quality assessment for JPEG2000 based on spatial features. *Signal Processing : Image Communication*, 23 :257–268, 2008. [cité p. 63]

[138] M. Jung, D. Lger, and M. Gazalet. Univariant assessment of the quality of images. *Journal of Electronic Imaging*, 11(3) :354–364, 2002. [cité p. 63]

[139] C. Charrier, G. Lebrun, and O. Lezoray. A machine learning-based color image quality metric. In *Third European Conference on Color Graphics, Imaging, and Vision*, pages 251–256, 2006. [cité p. 63]

[140] M. Saad, A. C. Bovik, and C. Charrier. A DCT statistics-based blind image quality index. *IEEE Signal Processing Letters*, 17(6) :583–586, 2010. [cité p. 63, 65]

[141] S. Gabarda and G. Cristbal. Blind image quality assessment through anisotropy. *Journal of Optical Society of America*, pages B42–B51, 2007. [cité p. 64]

[142] P. Parisot. *Suivi d'objets dans des séquences d'images de scènes déformables*. PhD thesis, Institut de recherche en Informatique de Toulouse, 2009. [cité p. 65]

[143] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In *International Conference on Computer Vision*, pages 454 – 461, 2001. [cité p. 65]

[144] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110 :346 – 359, 2008. [cité p. 65]

[145] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27 :1615–1630, 2005. [cité p. 65]

[146] S. Berretti, A. D. Bimbo, P. Pala, B. B. Amor, and M. Daoudi. A set of selected sift features for 3D facial expression recognition. In *Proceedings of the 20th International Conference on Pattern Recognition (ICPR)*, pages 4125–4128, 2010. [cité p. 65]

[147] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–151, 1988. [cité p. 66]

[148] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995. [cité p. 68]

[149] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM : a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`. [cité p. 69]

[150] G. Saporta. *Probabilités, Analyse des données et Statistiques*. Editions Technip, 1990. [cité p. 75]

[151] R. Giot, M. El Abed, and C. Rosenberger. Keystroke dynamics authentication for collaborative systems. *Collaborative Technologies and Systems, International Symposium*, pages 172–179, 2009. [cité p. 80]

[152] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22 :1–55, 1932. [cité p. 80]

[153] J. J. Higgins. An introduction to modern nonparametric statistics. *The American Statistician*, 2003. [cité p. 83]

[154] R.R. Bouckaert. *Bayesian Belief Networks : from Construction to Inference*. PhD thesis, University of Utrecht, 1995. [cité p. 83]

[155] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 1997. [cité p. 83]

[156] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. Classification and regression trees. Wadsworth International Group, 1984. [cité p. 83, 86, 88]

[157] C. Fauré. *Découvertes de motifs pertinents par l'implémentation d'un réseau bayésien : application à l'industrie aéronautique*. PhD thesis, Institut national des sciences appliquées de Lyon, 2007. [cité p. 84]

[158] R. R. Bouckaert, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald, and D. Scuse. Weka manual. Technical report, Department of Computing Science, University of Waikato, New Zealand, 2009. [cité p. 84]

[159] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1) :1–38, 1977. [cité p. 86]

[160] J. R. Quinlan. *C4.5 : Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning)*, volume 16. Morgan Kaufmann, 1993. [cité p. 88, 89]

[161] J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1986. [cité p. 89]

[162] R. Rakotomalala. *Graphes d'induction*. PhD thesis, Université Claude Bernard - Lyon 1, 1997. [cité p. 90, 91]

[163] C. X. Ling, J. Huang, and H. Zhang. AUC : A better measure than accuracy in comparing learning algorithms. In *Canadian Conference on Artificial Intelligence*, volume 2671, pages 329–341, 2003. [cité p. 90]

[164] A. Benaiss, U. Saeed, J. L. Dugelay, and M. Jedra. Impostor detection using facial stereoscopic images. In *17th European Signal Processing Conference (EUSIPCO)*, pages 1–4, 2009. [cité p. 101]

[165] G. Stoneburner, A. Goguen, and A. Feringa. Risk management guide for information technology system. Technical report, National Institute of Standards and Technology (NIST), 2002. [cité p. 106]

[166] R. Ortalo, Y. Deswarte, and M. Kaaniche. Experimenting with quantitative evaluation tools for monitoring operational security. *IEEE Transactions on Software Engineering*, 25 :633–650, 1999. [cité p. 107]

[167] C. Alberts, A. Dorofee, J. Stevens, and C. Woody. Introduction to the OCTAVE approach. Technical report, U.S. Department of Defense, 2003. [cité p. 107]

[168] Z. Yazar. A qualitative risk analysis and management tool - CRAMM. Technical report, SANS Institute, 2002. [cité p. 107]

[169] COBRA. Consultative, Objective and Bi-functional Risk Analysis (COBRA). http ://www.security-risk-analysis.com/, 2010. [cité p. 107]

[170] A. Rot. IT risk assessment : Quantitative and qualitative approach. In *the World Congress on Engineering and Computer Science (WCECS)*, pages 1–6, 2008. [cité p. 107]

[171] MCA. *Multi-criteria analysis : a manual*. Department for Communities and Local Government : London, 2009. [cité p. 107]

[172] J. Ashbourn. Vulnerability with regard to biometric systems. http ://www.eetimes.com/, 2010. [cité p. 109]

[173] Common Criteria Biometric Evaluation Methodology Working Group. *Common Criteria for Information Technology Security Evaluation*, 2002. [cité p. 109]

[174] C. Roberts. Biometric attack vectors and defences. *Computers & Security*, 2007. [cité p. 109]

[175] M. El Abed, R. Giot, B. Hemery, and C. Rosenberger. A study of users' acceptance and satisfaction of biometric systems. In *International Carnahan Conference on Security Technology (ICCST)*, pages 170–178, 2010. [cité p. 109, 119]

[176] G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds. Sheep, goats, lambs and wolves : A statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation. In *International Conference on Spoken Language Processing (ICSLP)*, pages 1–4, 1998. [cité p. 112]

[177] ISO/IEC 27000. Information technology – security techniques – information security management systems – overview and vocabulary, 2009. [cité p. 113]

[178] N. Ratha, J. Connell, R. M. Bolle, and S. Chikkerur. Cancelable biometrics : A case study in fingerprints. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 370–373, 2006. [cité p. 113]

[179] A. K. Jain, S. Prabhakar, and S. Pankanti. Can identical twins be discriminated based on fingerprints ? Technical report, Department of Computer Science, Michigan State University, 2000. [cité p. 114]

[180] M. El Abed, R. Giot, C. Charrier, and C. Rosenberger. Evaluation of biometric systems : A SVM-based quality index. In *The third Norsk Information Security Conference (NISK)*, pages 1–12, 2010. [cité p. 118]

[181] M. El Abed, R. Giot, B. Hemery, C. Charrier, and C. Rosenberger. A SVM-based model for the evaluation of biometric sample quality. In *IEEE International Workshop on Computational Intelligence in Biometrics and Identity Management*, pages 115–122, 2011. [cité p. 118]

[182] M. El Abed, B. Hemery, C. Charrier, and C. Rosenberger. Evaluation de la qualité de données biométriques. *Revue des Nouvelles Technologies de l'Information (RNTI), numéro spécial "Qualité des Données et des Connaissances / Evaluation des méthodes d'Extraction de Connaissances dans les Données"*, pages 1–18, 2011. [cité p. 118]

[183] M. El Abed, B. Hemery, C. Charrier, and C. Rosenberger. Un modèle SVM pour l'évaluation de la qualité des données biométriques. In *XXIII Colloque GRETSI*, pages 1–4, 2011. [cité p. 118]

[184] M. El Abed, R. Giot, B. Hemery, and C. Rosenberger. Evaluation of biometric systems : A study of users' acceptance and satisfaction. *Inderscience International Journal of Biometrics*, pages 1–26, 2011. [cité p. 119]

[185] M. El Abed, R. Giot, B. Hemery, J.J. Shwartzmann, and C. Rosenberger. Towards the security evaluation of biometric authentication systems. In *IEEE International Conference on Security Science and Technology (ICSST)*, pages 167–173, 2011. [cité p. 120]

[186] M. El Abed, P. Lacharme, and C. Rosenberger. Security evabio : An analysis tool for the security evaluation of biometric authentication systems. In *the 5th IAPR/IEEE International Conference on Biometrics (ICB)*, pages 1–6, 2012. [cité p. 120]

[187] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko. Users's Guide to NIST Biometric Image Software (NBIS). Technical report, National Institute of Standards and Technology (NIST), 2007. [cité p. 120]

[188] M. Dacier and Y. Deswarte. Privilege graph : an extension to the typed access matrix model. In *European Symposium in Computer Security (ESORICS'94)*, volume 875 of *Lecture Notes in Computer Science*, pages 319–334, 1994. [cité p. 121]

[189] A. J. O'Toole, J. Harms, S. L. Snow, D. R. Hurst, M. R. Pappas, J. H. Ayyad, and H. Abdi. A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27 :812 – 816, 2005. [cité p. 150]

Annexes

Annexe A

Rappels statistiques

Nous rappelons brièvement la définition de quelques outils statistiques utiles dans le cadre de cette thèse.

Soit X une série statistique prenant les valeurs x_1, x_2, ..., x_p avec les effectifs n_1, n_2, ..., n_p, et $n = n_1 + n_2 + ...+ n_p$ l'effectif total de cette série.

- **Moyenne**
  La moyenne est une mesure statistique caractérisant les éléments d'un ensemble de quantités. Elle est donnée par :

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{i=p} n_i x_i \tag{A.1}$$

- **Variance**
  La variance est une mesure arbitraire servant à caractériser la dispersion d'une distribution ou d'un échantillon. Elle est donnée par :

$$V = \frac{1}{n} \sum_{i=1}^{i=p} n_i x_i^2 - (\overline{x})^2 \tag{A.2}$$

- **Coefficient de corrélation linéaire de Pearson**
  Le coefficient de corrélation entre deux variables aléatoires, $X(x_1, ..., x_p)$ et $Y(y_1, ..., y_p)$, permet de quantifier la relation de dépendance qui peut exister entres ces variables. Elle est donnée par :

$$r_p = \frac{\sum_{i=1}^{i=p}(x_i - \overline{x}).(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{i=p}(x_i - \overline{x})^2}.\sqrt{\sum_{i=1}^{i=p}(y_i - \overline{y})^2}} \tag{A.3}$$

Le coefficient de corrélation linéaire est compris entre -1 et 1. Plus le coefficient est proche des valeurs extrêmes -1 et 1, plus la corrélation linéaire entre les variables est forte. Les valeurs intermédiaires renseignent sur le degré de dépendance linéaire entre les deux variables. Une corrélation égale à 0 signifie que les variables sont linéairement indépendantes.

– **Test de Kruskal-Wallis (KW)**
C'est un test non paramétrique utilisé pour tester si k échantillons ($k \geq 2$) peuvent être considérés comme similaires (en terme de moyenne μ). Les deux hypothèses du test KW sont :

– H_0 : la loi ayant généré les données est la même pour tous les échantillons contre
– H_1 : certains échantillons présentent systématiquement des valeurs plus élevées que d'autres.

$$\begin{cases} H0 : \ \mu_1 \ = \ \mu_2 \ = ... = \ \mu_k \\ H1 : \ \mu_i \ \neq \ \mu_j \quad \exists \ (i,j) \text{ avec } i \neq j \end{cases} \tag{A.4}$$

Avec des échantillons de taille respective $n_1, n_2, \cdots, n_k$, soit au total N mesures ($N = n_1 + n_2 + \cdots + n_k$), on calcule les rangs sur la réunion de tous les échantillons, $r_{11}, \ldots, r_{n1,1}, r_{21}, \ldots, r_{n2,1}, \ldots, r_{1,k}, \ldots, r_{nk,k}$. Le test statistique de Kruskal-Wallis (KW) est ainsi défini par :

$$H \ = \ \frac{12}{N(N+1)} \ \left(\sum_{i=1}^{k} \ \frac{W_i^2}{n_i} \right) \ - \ 3 \ (N+1) \tag{A.5}$$

où

$$W_i = n_i \bar{r_{i.}} \qquad \text{et} \qquad \bar{r_{i.}} \ = \ \frac{\sum_{j=1}^{n_i} r_{ij}}{ni}$$

Le critère de décision est défini par :

$$\begin{cases} p - valeur \ \geq \ 0.05 & \text{accepter } H_0 \\ sinon & \text{rejeter } H_0 \end{cases} \tag{A.6}$$

où p-valeur est la valeur estimée, pour k échantillons, en utilisant la distribution de probabilité du χ^2 avec k-1 degrés de liberté.

Annexe B

Performance des systèmes biométriques

	Intitulé	**Métrique**
Fondamentale	Taux d'échec à l'acquisition	FTA
	Taux d'échec à l'enrôlement	FTE
	Taux de fausse non-correspondance	FNMR
	Taux de fausse correspondance	FMR
	Capacité d'un système biométrique	C_s
	Capacité d'un utilisateur m	C_m
	Capacité d'un système biométrique multimodale	C_ψ
Vérification	Taux de faux rejets	FRR
	Taux de fausses acceptations	FAR
	Taux d'égale erreur	EER
	Taux d'erreur moyenne	HTER
	zone sous la courbe ROC	AUC
	courbe robustesse	RC
Identification	Taux d'identification	IR
	Taux de faux-négatif d'identification	FNIR
	Taux de faux-positif d'identification	FPIR
	Taux de pénétration	PR

TAB. B.1: Tableau de métriques.

Base	**Modalités**	**Participants / nombre d'échantillons par participant**	**Accessibilté**	**Commentaires**
FACES94	visage	152 / 20	`http://cswww.essex.ac.uk/mv/allfaces/faces94.html`	la base est téléchargeable gratuitement
AR	visage	120 / 26	`http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html`	la base est téléchargeable gratuitement
FERET	visage	725 / 5 à 91	`http://face.nist.gov/colorferet/`	la base est téléchargeable gratuitement
*FVC*2002 DB_2	empreinte digitale	100 / 8	`http://bias.csr.unibo.it/fvc2002/databases.asp`	la base est distribuée sur un DVD inclus dans [11]
ENSIB	visage	100 / 40	Disponible sur demande [a]	la base est téléchargeable gratuitement

TAB. B.2: Bases de données biométriques utilisées.

a. christophe.rosenberger@ensicaen.fr

Compétition	**Modalités**	**Bases de données**	**Commentaires**
SVC	dynamique de signature	deux bases de données sont utilisées et téléchargeable gratuitement sur `http://www.cse.ust.hk/svc2004/download.html`	les résultats de la compétition sont donnés dans [79]
FVC	empreinte digitale	quatre bases dont trois sont réelles et une synthétique générée par le logiciel *SFinGe*	les résultats de la compétition et les bases de données employées sont accessibles sur `http://bias.csr.unibo.it/fvc2006/`
FRVT et *ICE*	visage et iris, respectivement	les bases sont propres à la compétition	les résultats de la compétition sont donnés dans [80]
SRE	voix	données collectées par *Linguistic Data Consortium (LDC)* [a]	les résultats de ces compétitions sont accessibles sur `http://www.itl.nist.gov/iad/mig//tests/sre/`
BMEC	visage, empreinte, signature et iris	la base multimodale BioSecure [77]	les résultats sont accessibles sur `http://biometrics.it-sudparis.eu/BMEC2007/`
MBGC	visage et iris	données collectées par *U. of Notre Dame* et *U. of Texas at Dallas* [189]	les résultats sont accessibles sur `http://www.nist.gov/itl/iad/ig/mbgc-presentations.cfm`

TAB. B.3: Compétitions en biométrie.

a. http ://ldc.upenn.edu/

Plate-forme	**Modalités**	**Accessibilité**	**Commentaires**
BioSecure Reference and evaluation framework	visage, voix, iris, empreinte digitale, géométrie de la main et signature dynamique	`http://www.biosecure.info/`	la plateforme est accessible gratuitement en ligne
GREYC-Keystroke	dynamique de frappe au clavier	`http://www.epaymentbiometrics.ensicaen.fr/index.php/research-activities/resources`	le logiciel est téléchargeable gratuitement
FVC-onGoing	empreinte digitale	`https://biolab.csr.unibo.it/FVCOnGoing`	la plate-forme est accessible gratuitement en ligne

TAB. B.4: Plateformes en biométrie.

Liste des abréviations

FTA	Taux d'échec à l'acquisition (*failure-to-acquitre*)
FTE	Taux d'échec à l'enrôlement (*failure-to-enroll rate*)
$FNMR$	Taux de fausse non-correspondance (*false non-matche rate*)
FMR	Taux de fausse correspondance (*false matche rate*)
FRR	Taux de faux rejets (*false rejection rate*)
FAR	Taux de fausses acceptations (*false acceptance rate*)
IR	Taux d'identification (*identification rate*)
$FNIR$	Taux de faux-négatif d'identification (*false-negative identification-error rate*)
$FPIR$	Taux de faux-positif d'identification (*false-positive identification-error rate*)
PR	Taux de pénétration (*penetration rate*)
ROC	Courbe représentant les taux d'erreur (*Receiver operating characteristic curve*)
CMC	Courbe représentant les taux d'identification correcte (*Cumulative match characteristic curve*)
RC	Courbe de robustesse (*Robustness curve*)
EER	Taux d'égale erreur (*Equal Error Rate*)
WER	Taux d'erreur pondéré (*Weighted Error Rate*)
$HTER$	Erreur moyenne (*Half Total Error Rate*)
AUC	Aire sous la courbe ROC (*Area Under ROC Curve*)
C	Capacité du système biométrique (*Constrained capacity of biometric system*)
IC	Intervalle de confiance (*Confidence interval*)
GAB	Guichet Automatique Bancaire

Table des figures

Liste des tableaux

Les systèmes biométriques sont de plus en plus utilisés pour vérifier ou déterminer l'identité d'un individu. Compte tenu des enjeux liés à leur utilisation, notamment pour des applications dans le domaine de commerce électronique ou le contrôle d'accès physique, il est particulièrement important de disposer d'une méthodologie d'évaluation de tels systèmes. Le problème traité dans cette thèse réside dans la conception d'une méthodologie générique (*i.e.*, indépendante de la modalité) visant à évaluer un système biométrique. Les enjeux sont nombreux comme la comparaison des systèmes biométriques pour une application particulière, l'évaluation d'un algorithme au cours de son développement ou son paramétrage optimal.

Nous proposons dans cette thèse une méthode d'évaluation de la qualité de données biométriques morphologiques. La méthode proposée possède l'avantage d'être plurimodale (visage, empreinte digitale et veines de la main), et indépendante du système de vérification utilisé. Le comportement de cette méthode d'évaluation a été testé sur cinq bases de données publiques. Nous avons également mis au point deux méthodes d'évaluation portant sur l'usage et la sécurité d'un système biométrique. La première consiste à évaluer l'acceptabilité et la satisfaction des usagers lors de l'utilisation des systèmes biométriques. La seconde consiste à mesurer la robustesse d'un système biométrique (architecture et algorithmes) contre la fraude. Les deux méthodes proposées possèdent l'avantage d'être indépendantes de la modalité biométrique considérée. Trois systèmes d'authentification biométrique ont été utilisés pour montrer l'intérêt de ces deux méthodes.

Evaluation of Biometric Systems

Biometric systems are increasingly used to verify or identify the identity of an individual. Given the challenges related to their use, mainly for e-commerce applications or physical access control (border control), it is important to have an evaluation methodology of such systems. The problematic addressed in this thesis is the design of a modality-independent evaluation methodology of biometric systems. The challenges are many such as the comparison of biometric systems for a particular application, the evaluation of an algorithm during its development or in setting its optimal parameters.

We propose in this thesis a quality assessment method of biometric raw data. The proposed method has the advantage of being plurimodal (Face, Fingerprint and hand veins), and independent from the used verification system. Five public databases are used to validate the proposed method. We have also developed two evaluation methods related to the usability and the security aspects of a biometric system. The first consists of measuring users' acceptance and satisfaction when using biometric systems. The second consists of measuring the system (architecture and algorithms) robustness against attacks. The two presented methods have the advantage of being modality-independent. Three biometric authentication systems are used to show the benefits of both methods.

Indexation Rameau : EVALUATION / RECONNAISSANCE DE FORMES (INFORMATIQUE) / TRAITEMENT D'IMAGES - TECHNIQUES NUMÉRIQUES / CLASSIFICATION

Informatique et applications

Laboratoire GREYC - UMR CNRS 6072 - Université de Caen Basse-Normandie - Ensicaen
6 Boulevard du Maréchal Juin - 14050 CAEN CEDEX