ROYAUME DU MAROC
Université Mohammed V - RABAT -
جامعة محمد الخامس – الرباط –
Faculté des sciences
كلية العلوم
CENTRE D'ETUDES DOCTORALES - SCIENCES ET TECHNOLOGIES

N° d'ordre : 3578

THESE

En vue de l'obtention du : ***DOCTORAT***

Structure de Recherche *: Laboratoire de Recherche en Informatique et Télécommunications*
Discipline *: Sciences de l'ingénieur*
Spécialité *: Informatique et Télécommunications*

Présentée et soutenue le 25-12-2021 par :

Ayoub LAHMIDI

Protection des Systèmes de Vérification Biométrique : Contributions à la Protection des Modèles des Empreintes Digitales.

Rachid OULAD HAJ THAMI	*PES*	*Ecole Nationale Supérieure d'Informatique et d'Analyse des Systèmes, Université Mohammed V.*	*Président*
Hicham LAANAYA	*PH*	*Faculté des Sciences de Rabat, Université Mohammed V.*	*Rapporteur/Examinateur*
Khadija RHOULAMI	*PH*	*Faculté des Lettres et Sciences Humaines de Rabat, Université Mohamed V*	*Rapporteur/Examinateur*
Mounia MIKRAM	*PH*	*Ecole d'Ingénieurs en Sciences de l'Information de Rabat.*	*Rapporteur/Examinateur*
Chouaib MOUJAHDI	*PH*	*Institut Scientifique de Rabat, Université Mohammed V.*	*Examinateur*
Khalid MINAOUI	*PES*	*Faculté des Sciences de Rabat, Université Mohammed V.*	*Co-Directeur de thèse*
Mohammed RZIZA	*PES*	*Faculté des Sciences de Rabat, Université Mohammed V.*	*Directeur de thèse*

Année Universitaire : 2020/2021

À ma mère et mon père bien aimés Lhajja Zahra et Lhajj Mustapha
Je dédie ce travail de thèse aux 2 personnes
qui me sont le plus chères au monde
pour avoir fait de moi ce que je suis.

À mon cher frère : Mohamed
À ma chère petite soeur : Amira

À toute ma famille

Remerciements

Les travaux présentés dans ce mémoire ont été effectués au Laboratoire de Recherche en Informatique et Télécommunications (LRIT) de la Faculté des Sciences de Rabat (FSR)-Université Mohammed V au Maroc sous la direction de **M. Mohammed RZIZA** et la co-direction de **M. Khalid MINAOUI**.

En premier lieu, je tiens à remercier **M. Mohammed RZIZA** et **M. Khalid MINAOUI**, Professeurs d'Enseignement Supérieur à la Faculté des Sciences de Rabat. Dans une ambiance toujours décontractée mais néanmoins studieuse, ils m'ont bien guidé dans le monde de la recherche. Je réitère mes remerciements pour leur grande disponibilité, leur aide précieuse, ainsi que pour leurs efforts qu'ils ont prodigués pour l'accomplissement de ce travail de thèse. J'espère avoir été digne de leur confiance qu'ils m'ont accordée et que ce travail est finalement à la hauteur de leurs espérances.

Je tiens à remercier **M. Rachid OULAD HAJ THAMI**, Professeur d'Enseignement Supérieur à l'Ecole Nationale Supérieure d'Informatique et d'Analyse des systèmes de l'université Mohammed V de Rabat, d'avoir accepté de présider le jury de ma thèse.

Je tiens aussi à remercier **M. Hicham LAANAYA**, Professeur Habilité à la Faculté des Sciences de Rabat, d'avoir accepté de juger la qualité de mon travail en tant que rapporteur/examinateur.

Je tiens aussi à remercier **M. Khadija RHOULAMI**, Professeur Habilité à la Faculté des Lettres et Sciences Humaines de Rabat, d'avoir accepté de juger la qualité de mon travail en tant que rapporteur/examinateur.

Mes remerciements vont également à **M. Mounia MIKRAM**, Professeur Habilité à l'École d'Ingénieurs en Sciences de l'Information de Rabat, d'avoir accepté de rapporter et d'examiner ce mémoire de thèse.

Mes sincères remerciements vont également à **M. chouaib MOUJAHDI**, qu'il me soit permis de saluer sa bienveillance à mon égard. C'est avec honneur que j'ai pu bénéficier de ses connaissances et de ses remarques et critiques aiguisées. Je le remercie, particulièrement, pour son empathie dont il a fait preuve depuis que nous travaillons ensemble. Je le remercie aussi d'avoir accepté d'examiner mon travail.

Je remercie mes amis qui m'ont apporté leur soutien et leurs encouragements tout au long de cette aventure. Dans le désordre, un merci spécial à Marouane NAZIH, Khalid BEKKAOUI, Mehdi ZOUITINI et Mohamed OUNJAR pour être de vrais amis, merci pour vos beaux cœurs et votre soutien continu, merci pour les discussions inspirantes et tous les moments que nous avons passés ensemble.

Je garde le meilleur pour la fin, ma famille qui a supporté toutes les difficultés morales et matérielles pour me soutenir au terme de mes études. J'adresse ma profonde gratitude et mon immense reconnaissance à mes raisons d'être, ma mère et mon père, qui m'ont éduqué et orienté. Merci de m'avoir encouragé et soutenu dans mes choix. Nul mot et nulles expressions refléteront le grand amour et la profonde gratitude que je porte pour vous. À mon cher frère et ma chère petite sœur pour leurs soutiens et attentions. Ils étaient pour moi, une vraie source d'inspiration et ont été toujours à mes côtés. Que Dieu le Tout Puissant vous garde et vous procure santé et bonheur.

Table des matières

Liste des abréviations

MoC	*Match-on-Card*
SoD	*System-on-Device*
MoD	*Match-on-Device*
MoB	*Match-on-Board*
SoC	*System-on-a-Chip or System-on-Card*
FTE	*failure-to-enroll*
FTA	*failure-to-acquire*
FNMR	*False Non-Match Rate*
FMR	*False Match Rate*
FAR	*False Acceptance Rate*
FRR	*False rejection rate*
ERR	*Equal Error Rate*
FVC	*Fingerprint Verification Competition*
ROC	*Receiver Operating Characteristic*
GAR	*Genuine Accept Rate*
DET	*Detection Error Tradeoff*
K-S	*Kolmogorov-Smirnov*
ARM	*Attack via Record Multiplicity*

Liste des figures

Liste des tableaux

Liste des algorithmes

Résumé

La biométrie est aujourd'hui l'une des technologies les plus émergentes. Elle a été déployée avec succès dans divers projets gouvernementaux et organisationnels en étant une excellente solution pour l'authentification des individus, en raison de plusieurs avantages inhérents qu'elle offre par rapport aux systèmes traditionnels d'authentification qui utilisent les mots de passe et les cartes ID. Avec le déploiement généralisé des systèmes biométriques, la sécurité et la confidentialité de la technologie biométrique font l'objet de préoccupations croissantes. L'une des vulnérabilités potentielles est la compromission des modèles biométriques, qui peut entraîner de lourdes atteintes à la sécurité et à la vie privée. La plupart des techniques de protection des modèles existantes ne répondent pas à toutes les exigences d'un système biométrique pratique, telles que la révocabilité, la diversité, la sécurité, la confidentialité et la précision de la correspondance.

Dans le cadre de cette thèse, nous nous sommes intéressés à la conception et au développement de nouvelles approches dédiées spécifiquement à la protection des modèles stockés dans les systèmes de reconnaissance par empreintes digitales. En effet, Nous avons proposé deux approches de protection. La première approche consiste à exploiter les propriétés géométriques des minuties pour réaliser des transformations à sens unique. La mise en correspondance entre les modèles des empreintes digitales est effectuée dans l'espace transformé. Quant à la deuxième approche, nous avons adopté une nouvelle méthodologie pour la conception des schémas de protection des modèles biométriques. Nous nous sommes basés sur la spécification d'un système de vérification des empreintes digitales non protégé pour construire un schéma de protection spécifique et adapté, et qui fournit un excellent compromis entre sécurité et précision de reconnaissance.

Mots clés : Modèle biométrique, Empreinte digitale, Révocabilité, Diversité, Sécurité.

Abstract

Biometrics is one of the most emerging technologies today. It has been successfully deployed in various government and organizational projects, being an excellent solution for authenticating individuals, due to several inherent advantages it offers over traditional authentication systems that use passwords and ID cards. With the widespread deployment of biometric systems, there are growing concerns about the security and privacy of biometric technology. One of the potential vulnerabilities is the compromise of biometric templates, which can lead to serious security and privacy breaches. Most existing template protection techniques do not meet all the requirements of a practical biometric system, such as revocability, divisibility, security, confidentiality and matching accuracy.

In this thesis, we have focused on the design and development of new approaches specifically dedicated to stored template protection in fingerprint recognition systems. Indeed, we have proposed two protection approaches. The first approach consists in exploiting the geometric properties of minutiae to perform one-way transformations. The mapping between fingerprint templates is performed in the transformed space. In the second approach, we have adopted a new methodology for the design of biometric template protection schemes. Based on the specification of an unprotected fingerprint verification system, we built a specific and suitable protection scheme, which provides an excellent compromise between security and recognition accuracy.

Keywords : Biometric template, Fingerprint, Revocability, Diversity, Security.

Chapitre

1 Introduction

1.1 Contexte et problématique

Alors que le monde poursuit sa marche inexorable vers tout ce qui est numérique, les informations personnelles qui se transmettent à travers les nouvelles technologies de l'information et de la communication risquent d'être de plus en plus exposées. Des préoccupations se sont donc fait jour en ce qui concerne l'usurpation d'identité et l'accès frauduleux aux données à caractère personnel. Ces risques peuvent avoir des répercussions très lourdes pour la victime, tels que l'atteinte à l'intimité de la vie privée, la divulgation de secrets d'affaires et de savoir-faire, où même pire, des problèmes financières (Endettement, interdiction bancaire).

L'usurpation d'identité connait aujourd'hui une hausse fulgurante surtout avec la pandémie mondiale de Covid-19 en plein vigueur. Selon le *Digital Journal*, les américains dépensent plus de 3,5 milliards de dollars par an dans les services de protection de l'identité, mais ils sont encore nombreux à être victimes. La Commission fédérale du commerce (FTC) qui suit les plaintes de fraude et de vol d'identité des consommateurs qui ont été déposées auprès des organismes fédéraux, étatiques et locaux chargés de l'application de la loi et des organisations privées, a reçu 4,8 millions de rapports de vol d'identité et de fraude en 2020, soit une augmentation de 45 % par rapport aux 3,3 millions de(en) 2019, principalement due à l'augmentation de 113 % des plaintes pour vol d'identité. En 2020, 1,4 million de plaintes concernaient le vol d'identité, contre 651 000 en 2019, ce qui montre que l'usurpation d'identité gagne rapidement en gravité et continuera de prospérer tant que de nouvelles méthodes d'authentification et de validation ne seront pas mises en place.

Il est impératif de disposer de techniques qui permettent de déterminer avec précision l'identité des individus. Une telle action peut être nécessaire pour diverses raisons, mais l'intention première, dans la plupart des systèmes numériques, est d'empêcher les imposteurs d'accéder à des ressources confidentielles. Les méthodes traditionnelles de détermination d'identité reposent sur des dispositifs basés sur la connaissance (par exemple, les codes PIN et les mots de passe) ou sur la possession (par exemple, les clés et les cartes d'identité), ces représentations substitutives ne permettent pas enfaite de garantir, avec précision, le rapport direct entre l'utilisation du service et l'utilisateur réel puisqu'elles

peuvent être volées, perdues, partagées ou manipulées, ce qui nuit à la sécurité prévue étant donné que le système ne peut pas faire dans cette situation la distinction entre un fraudeur et un utilisateur légitime. Un autre enjeu lié à ces mécanismes d'authentification est la répudiation. Un utilisateur peut délibérément partager ses informations d'identification et prétendre ensuite qu'elles ont été volées. Ainsi, un tel système peut facilement être trompé.

Étant dans l'ère des nouvelles technologies de communication, les systèmes d'authentification traditionnels basés sur les titres de compétences ne suffisent plus à vérifier avec fiabilité l'identité des individus, ce qui a conduit à un besoin prononcé en matière de techniques d'authentification qui offrent une meilleure sécurité, une plus grande efficacité et, dans de nombreux cas, un confort accru pour l'utilisateur.

La biométrie a pu offrir une solution naturelle et fiable à la gestion de l'identité en utilisant des schémas entièrement automatisés ou semi-automatisés pour reconnaître les individus sur la base de leurs caractéristiques morphologiques (Empreinte digitale, visage, iris, géométrie des mains, etc.), de leurs caractéristiques comportementales (voix, signature, frappe au clavier, etc.) où bien de leurs caractéristiques biologiques (Salive, Urine, ADN, etc.), appelées identifiants ou traits biométriques. Grâce à la biométrie, il est possible de fonder une identité sur ce que l'on est, plutôt que sur ce que l'on possède (par exemple, une carte d'identité), ou sur ce dont on se souvient (par exemple, un mot de passe). Aujourd'hui, la biométrie tend à devenir un atout essentiel pour les solutions visant à identifier efficacement les personnes, car les identifiants biométriques sont à la fois universels, permanents, uniques et à caractère distinctif. Ils ne peuvent donc ni être partagés, ni perdus, ni oubliés contrairement aux mots de passes et aux cartes d'identités, et ils représentent de plus intrinsèquement l'identité corporelle de l'individu. Les systèmes d'authentifications basés sur la biométrie ont donc été en mesure de surmonter les limites des systèmes traditionnelles tout en offrant une bonne sécurité et une grande efficience. Outre le renforcement de la sécurité, les systèmes biométriques améliorent également le confort des utilisateurs en leur évitant d'avoir à concevoir et à retenir des mots de passe longues où complexes.

Bien que la biométrie appartienne à la catégorie des technologies d'authentification forte, et s'avère plus prometteuse par rapport aux systèmes traditionnelles, elle n'est pas non plus hors danger. Enfaite, la biométrie présente plusieurs vulnérabilités et défis qui peuvent entraîner de nombreuses violations de la sécurité et de la vie privée, et donc à dégrader considérablement leur intérêt.

Aujourd'hui, on ne peut pas garantir la robustesse des systémes biométriques en pratique dans un contexte d'utilisation spécifique. De plus, il existe plusieurs facteurs affectant la performance de ces systèmes [Jain et Ross, 2004] en terme de précision 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 des différents individus 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).

Une vulnérabilité dans la sécurité biométrique se traduit généralement par une reconnaissance incorrecte ou une incapacité à reconnaître correctement les individus. Cette définition inclut des méthodes permettant d'accepter faussement un individu (régénération de gabarit), d'impacter les performances globales du système (déni de service), ou d'attaquer un autre système via des données fuitées (vol d'identité). Les vulnérabilités sont mesurées par rapport à des revendications de conception explicites ou implicites.

De nombreuses études [Ratha *et al.*, 2001b, Bolle *et al.*, 2002, Jain *et al.*, 2007] ont fait état de problèmes de sécurité dans les systèmes biométriques, et qui peuvent être soit (i) inhérentes à la technologie biométrique soit (ii) intentionnellement causées par des attaques adverses. Généralement, les aspects les plus vulnérables de la biométrie sont décrits ci-dessous :

1. L'usurpation d'identité : le fraudeur peut en effet tenter de se faire passer pour un utilisateur légitime. En compromettant par exemple le modèle biométrique stocké, il est ainsi possible de reconstruire un signal artificiel proche de l'original qui peut toutefois passer le seuil de décision de vérification.
2. L'irrévocabilité : Le grand inconvénient de la biométrie est que si le modèle biométrique est utilisé abusivement ou compromis, il ne peut généralement pas être révoqué, remplacé ou mis à jour.
3. La violation de la vie privée : Le recours au corps humain en tant que support d'identification et son stockage dans des bases de données, soulèvent un véritable problème éthique. Par ailleurs, étant donné que la biométrie met en jeu des données personnellement identifiables et donc très sensibles [Information *et al.*, 2008], sa récolte, son stockage et son usage doivent être réglementés par des juridictions légales. Bien que le caractère unique de la biométrie soit perçu comme un atout, il peut être aussi considéré comme une possibilité de profilage et de surveillance d'une

personne, constituant ainsi une menace pour sa liberté individuelle. Par conséquent, la mise en œuvre d'un système biométrique doit être établie sur la base d'un fort impératif de respect de la vie privée.

Compte tenu des risques encourus, un système biométrique doit être subordonné à diverses restrictions en matière de sécurité et de respect de la vie privée avant d'être déployé. Il est donc essentiel d'assurer la sécurité des systèmes biométriques et de protéger l'identifiant biométrique par des contre-mesures efficaces.

L'une des vulnérabilités potentielles d'un système biométrique est la divulgation d'informations sur les modèles biométriques, qui peut entraîner de graves menaces pour la sécurité et la confidentialité. La plupart des techniques de protection des gabarits disponibles ne parviennent pas à répondre à toutes les exigences souhaitées d'un système biométrique pratique, comme la révocabilité, la sécurité, la confidentialité et une grande précision de correspondance. En particulier, la protection des modèles des empreintes digitales a été un problème assez complexe en raison des grandes variations intra-classe (par exemple, la rotation, la translation, la déformation non linéaire et les empreintes partielles). Il y a deux défis fondamentaux dans tout schéma de protection de modèle d'empreinte digitale. Premièrement, il faut sélectionner un schéma de représentation approprié qui capture la plupart des informations discriminatoires, mais qui soit suffisamment invariant aux changements de placement des doigts et qui puisse être sécurisé à l'aide des algorithmes de protection des modèles disponibles. Deuxièmement, il faut aligner ou enregistrer automatiquement les empreintes digitales obtenues lors de l'inscription et de la mise en correspondance, sans utiliser des informations qui pourraient révéler les caractéristiques, qui permettent de distinguer de façon unique une empreinte digitale. En fonction des caractéristiques utilisées pour la reconnaissance, les solutions existantes pour la sécurité des gabarits des empreintes digitales peuvent être classées en approches basées sur les minuties ou sur les motifs.

La protection des données devrait veiller à ce que le déploiement des systèmes biométriques se fasse de manière à ce que l'accès aux données biométriques soit restreint aux personnes habilitées, en respectant certaines conditions. Toutefois, plutôt que de se fonder simplement sur les principes des bonnes pratiques, il serait plus judicieux de veiller à ce que la vie privée soit préservée dès la conception du système biométrique, et ce, avant même son déploiement. Aujourd'hui, un autre axe de recherche a vu le jour, qui vise à intégrer la protection de la vie privée comme une contrainte fonctionnelle du système biométrique. L'objectif est de trouver des solutions techniques qui permettraient de renforcer la technologie biométrique afin de protéger les données de référence. Cependant, il est apparu qu'en améliorant la préservation de la vie privée, les performances de reconnaissance se détériorent de manière notable. Les objectifs de sécurité et de confidentialité sont souvent contradictoires et l'un des principaux défis consiste à assurer un équilibre positif entre les deux.

Dans cette thèse, nous nous intéressons principalement à la protection des modèles

des empreintes digitales stockés dans les bases de donnés, et qui présentent des menaces de violation de vie privée, potentiellement envahissantes. Le but est d'apporter des solutions approuvées qui puissent réconforter les appréhensions dans les systèmes biométriques sans pour autant en diminuer les fonctionnalités. Le prochain paragraphe présente nos objectifs de recherche et les contributions apportées .

1.2 Objetifs et contributions

Les solutions les plus pertinentes pour la protection du modèle biométrique sont basées sur des approches de transformation où une version transformée du modèle original est enregistrée dans la base de données. Pour préserver la vie privée, le modèle transformé ou sécurisé doit être :

- Non-inversible, ce qui permet de déterminer la complexité du retour au modèle d'origine depuis le modèle transformé et donc de protéger les données personnelles.
- Diversifié, ce qui évitera que des modèles transformés soient mis en relation, même s'ils proviennent de la même personne. Ceci rendra impossible le suivi des individus dans plusieurs bases de données.

Parallèlement, les critères de sécurité suivants sont à respecter :

- Des performances de reconnaissances fiables, afin que le modèle transformé ne nuise pas aux performances du système.
- Une méthode de révocabilité et de substitution du modèle transformé du même trait biométrique, pour surmonter le problème de la compromission du modèle de référence.

Le défi scientifique relevé est de trouver une approche de transformation qui puisse garantir ces exigences en présence d'un signal biométrique naturellement variable. C'est dans cette problématique que s'inscrit l'objectif global de cette thèse.

La modalité biométrique qui fait l'objet de cette thèse est **l'empreinte digitale** et ce pour les raisons suivantes : L'empreinte digitale occupe une place prépondérante sur le marché de la biométrie tant au niveau privé que public ou gouvernemental. Le rapport 2009-2014 fournit par le groupe *"International Biometric Group"* le confirme [Group, 2013].

Les empreintes digitales sont considérées comme invasives en ce qui concerne la violation de la vie privée. En fait, les différentes modalités biométriques ne présentent pas les mêmes risques de violation de la vie privée. Une étude réalisée dans [Thieme, 2003] révèle que le visage et l'empreinte digitale sont des modalités à risque élevé en raison de leur grande accessibilité (capture à l'insu de la personne ou récupération sur supports ou avec une caméra). L'empreinte est une modalité à trace et le visage peut aisément être capturé par une caméra. En outre, leur haute compatibilité avec les bases de données existantes peut favoriser le suivi et la surveillance des personnes. L'iris et la rétine sont

considérées comme moyennement risquées alors que la signature, la voix et la dynamique de frappe sont considérées comme faiblement risquées. C'est donc bien la question relative à la protection des empreintes digitales qui est au centre de nos recherches.

Il existe de nombreuses solutions visant à protéger les empreintes digitales, mais le déploiement d'une solution adaptée au respect de la vie privée tout en maintenant les performances de reconnaissance demeure un sujet de recherche ouvert. Dans plusieurs de travaux de recherche, nous constatons que les propositions ne sont pas pour autant résistantes aux diverses attaques éventuelles. En matière de biométrie révocable, un constat de faiblesse porte sur la gestion de la clé utilisée comme paramètre dans les fonctions de transformation. Lorsque le fraudeur est au courant de cette clé, les performances en termes de taux de fausse acceptation diminuent. Dans certaines situations, si le modèle transformé est également compromis, le processus de réversibilité deviendra envisageable. Dans cette thèse, nous avons abordé le problème de la protection des empreintes digitales avec des contributions qui peuvent être présentées comme suit :

1. Contribution 1 : La conception d'un nouveau schéma de protection des modèles des empreintes digitales qui améliore la sécurité du système protégé tout en préservant ces performances. L'approche proposée est une technique basée sur les propriétés des minuties et qui effectue la correspondance des empreintes digitales dans un espace transformé en utilisant des tétraèdres de minuties irréversibles. Une analyse de sécurité rigoureuse et des expériences approfondies sont élaborées pour évaluer l'approche proposée. En utilisant le protocole original *Fingerprint Verification Competition* (FVC), les résultats expérimentaux fournis sur les bases de données d'empreintes digitales FVC2002 DB1, DB2 et DB3 ont montré des taux de reconnaissance satisfaisants. Nos résultats sont comparés à certaines techniques dans l'état de l'art qui utilisent le même protocole de test. Nous avons également prouvé que le système proposé répond aux exigences de révocabilité, de diversité et de non-inversibilité.

2. Contribution 2 : Étant donné que chaque système a ses propres particularités, allant de l'acquisition de l'empreinte digitale jusqu'au processus de correspondance, la majorité des techniques proposés comme solutions génériques, ne sont pas suffisamment matures pour un déploiement à grande échelle. Par conséquent, nous pensons que la méthodologie de conception des schémas de protection des empreintes digitales devrait être orientée pour construire des schémas de protection spécifiques pour chaque système non protégé, ce qui permettra d'obtenir le meilleur compromis entre performance et sécurité comparé à toute solution de protection générique. En adoptant cette méthodologie, nous proposons un nouveau schéma de protection pour les modèles des empreintes digitales qui est bien adapté à un système non protégé bien connu. Nos résultats expérimentaux, obtenus en utilisant des benchmarks standards tels que FVC 2002 DB1 et DB2, ont prouvé que la technique proposée répond aux exigences de révocabilité, de diversité, de non-inversibilité et de haute précision de reconnaissance.

1.3 Organisation du rapport

Ce manuscrit de thèse est organisé selon les six chapitres suivants :

- Le chapitre 1 introduit le problème de la sécurité dans les systèmes biométriques, clarifie le contexte de ce travail et la perspective suivie lors de l'élaboration de la thèse et précise les objectifs et les principales contributions.
- Le chapitre 2 a été consacré à la présentation de la biométrie de manière générale, les nombreuses applications de la biométrie et les notions utiles à la bonne compréhension de notre travail. L'architecture générale et le fonctionnement des systèmes biométriques sont aussi détaillées, ainsi que les vulnérabilités et les attaques qui menacent ce genre de système. Une description est donnée de manière plus spécifique sur la biométrie des empreintes digitales, sur les systèmes de reconnaissance par empreinte digitale et leurs défis, et finalement sur les problèmes liés à la sécurité des modèles de cette biométrie.
- Le chapitre 3 adresse dans un premier temps toutes les exigences d'un schéma de protection idéal, puis introduit les techniques de protection des modèles biométriques dans l'état de l'art, en mettant l'accent sur ceux qui visent uniquement les modèles des empreintes digitales. Les différents travaux qui s'accentuent autour des exigences de révocabilité et de confidentialité de la donnée d'empreinte digitale sont bien détaillés.
- Le chapitre 4 aborde les différentes méthodologies statistiques d'évaluation de performance et de sécurité des systèmes biométriques qui seront utilisées ultérieurement dans notre expérimentation. Des méthodologies d'évaluation spécifiques aux schémas de protection des modèles biométriques sont également présentées.
- Le chapitre 5 présente deux nouvelles approches de protection des modèles biométriques proposées dans le cadre de cette thèse, l'évaluation expérimentale de chaque approche est présentée également.
- Le chapitre 6 fait le point de cette recherche et liste les différentes perspectives de cette thèse.

Chapitre

2

Les systèmes biométriques

Sommaire

2.1 Introduction

Ce chapitre vise avant tout à donner un aperçu global de la biométrie et à poser les grandes idées et les principales notions qui seront utilisées tout au long de cette thèse. Une grande attention sera accordée à l'empreinte digitale puisqu'elle fait l'objet de toutes nos contributions. Le chapitre 2 est organisé de la manière suivante : Dans la section 2.2, nous commençons d'abord par définir la biométrie, son contexte et ses champs d'application. Nous présentons ensuite les grands concepts qui y sont rattachés. Nous discutons notamment des différentes modalités biométriques existantes, des caractéristiques particulières d'une information biométrique. Dans la section 2.3, nous abordons l'architecture générale ainsi que le fonctionnement d'un système biométrique. Les limites, les défaillances et les enjeux liés à la sécurité des systèmes biométriques seront consacrés à la section 2.4 en détaillant les types d'attaques qui peuvent être lancées contre ce genre de système. Nous nous focalisons ensuite dans la section 2.5 sur la biométrie des empreintes digitales qui a été traitée plus particulièrement dans ce travail de thèse. Cette section n'a pas pour objectif de présenter un état de l'art complet sur la biométrie des empreintes digitales. Nous visons surtout deux objectifs. Le premier est de présenter les principaux éléments théoriques liés aux empreintes digitales et à la biométrie associée, utiles à la bonne compréhension de notre problématique. Le deuxième est de discuter des problèmes rencontrés lors de la reconnaissance par empreintes digitales et les attaques qui menacent les modèles de cette biométrie. Enfin, le bilan du chapitre est présenté dans la section 2.6.

2.2 Biométrie

La biométrie est la science qui consiste à déterminer l'identité d'un individu sur la base de ses propriétés physiques ou comportementales. La pertinence de la biométrie a été consolidée par le besoin de systèmes de vérification et de gestion d'identité à grande échelle dont la fonctionnalité repose sur la détermination précise de l'identité d'un individu dans une variété d'environnements, allant de l'application de la loi et de la santé aux services financiers et à la sécurité des entreprises.

La propagation des services basés sur le web (par exemple, les services bancaires en ligne) et le déploiement de centres de service client décentralisés (par exemple, les cartes de crédit) ont encore accentué la nécessité de disposer de systèmes de gestion d'identité fiables, précis et capables de prendre en charge un grand nombre d'individus.

2.2.1 Modalités biométriques

Dans la littérature biométrique, toute caractéristique physiologique ou comportementale humaine qui peut être utilisée pour reconnaître une personne est appelée trait, indicateur, identifiant ou modalité biométrique, à condition qu'elle réponde aux sept exigences décrites par [Jain *et al.*, 2004], à savoir :

1. L'universalité : suppose que chaque personne dispose du trait biométrique (sauf en cas de mutilation ou d'handicap) ;
2. Le caractère distinctif : indique que deux personnes doivent être suffisamment différentes par rapport à leurs identifiants biométriques ;
3. La permanence : implique que la biométrie doit être suffisamment invariante sur une période donnée ;
4. La collectabilité : indique que le trait biométrique peut être mesuré quantitativement.

Cependant, dans un système pratique qui s'appuie sur les identifiants biométriques, un certain nombre d'autres enjeux doivent être prises en compte, notamment :

5. La performance : fait référence à la précision, à la vitesse et à la robustesse de reconnaissance réalisables, aux ressources nécessaires pour atteindre la précision et la vitesse de reconnaissance souhaitées, ainsi qu'aux facteurs opérationnels ou environnementaux qui affectent la précision et la vitesse de reconnaissance ;
6. L'acceptabilité : indique dans quelle mesure les utilisateurs sont prêts à accepter un identifiant biométrique particulier dans leur vie quotidienne ;
7. Le contournement : reflète la facilité avec laquelle il est possible de tromper le système par des méthodes frauduleuses.

Un système biométrique pratique doit présenter une précision et une vitesse de reconnaissance acceptables avec des besoins en ressources raisonnables, être inoffensif pour les utilisateurs, être accepté par la population visée et être suffisamment résistant aux diverses méthodes frauduleuses.

Un certain nombre d'identifiants biométriques sont utilisés dans diverses applications (figure 2.1). Chaque biométrie a ses avantages et ses inconvénients et, par conséquent, le choix d'un trait biométrique pour une application particulière dépend d'une variété de questions en plus de ses performances de correspondance.

Lors du choix d'un identifiant biométrique pour une application particulière, les questions suivantes doivent être abordées :

- L'application nécessite-t-elle une vérification ou une identification ? Si une application nécessite l'identification d'un sujet à partir d'une grande base de données, elle a besoin d'une biométrie évolutive et relativement plus distinctive (par exemple, empreinte digitale, iris ou ADN).
- Quels sont les modes de fonctionnement de l'application ? Par exemple, si l'application est surveillée (semi-automatique) ou non (entièrement automatique), si les utilisateurs sont habitués (ou prêts à s'habituer) à la biométrie donnée, si l'application est secrète ou ouverte, si les sujets sont coopératifs ou non, etc.
- Quelles sont les exigences de l'application en matière de stockage ? Par exemple, une application qui effectue la reconnaissance sur un serveur distant peut exiger une petite taille de modèle.

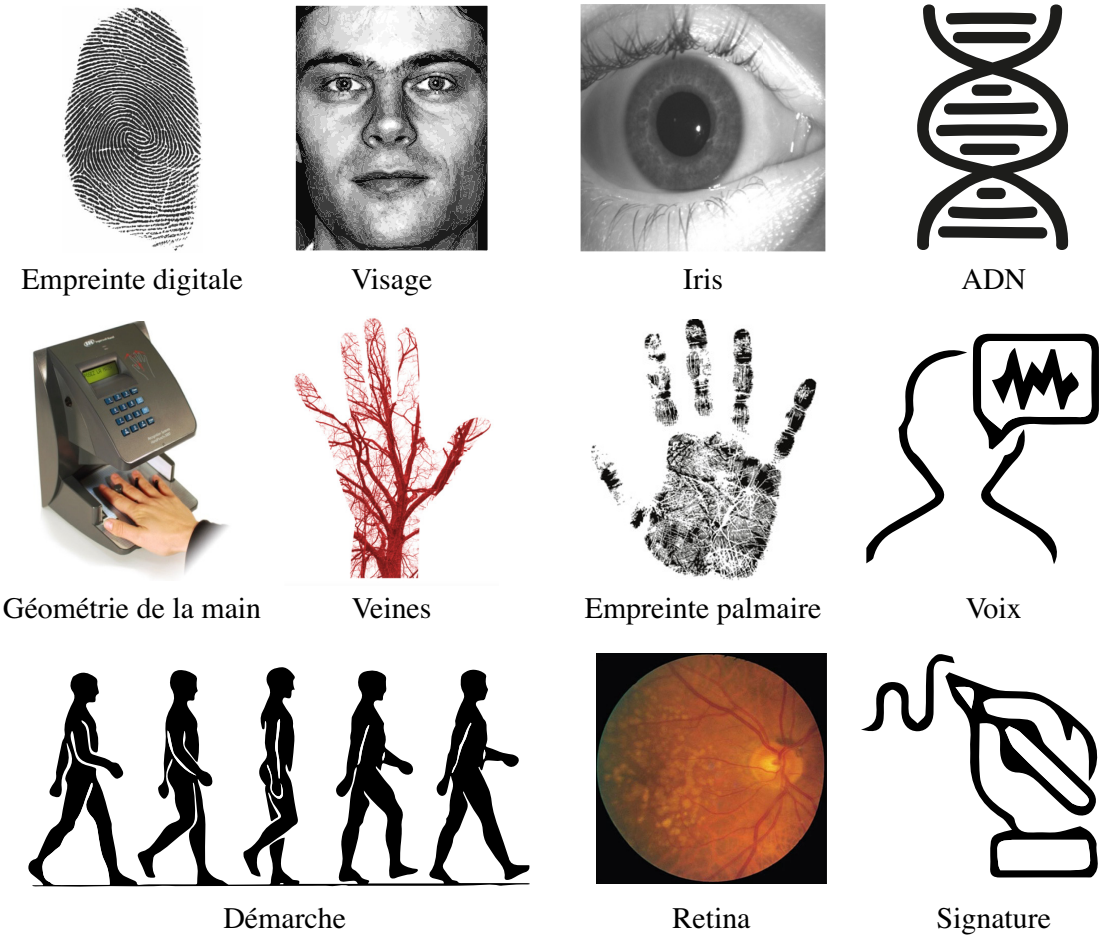


FIGURE 2.1 – Quelques modalités biométriques.

- Quelles sont les exigences en matière de performances ? Par exemple, une application qui exige une très grande précision a besoin d'un élément biométrique plus distinctif.
- Quels types d'éléments biométriques sont acceptables pour les utilisateurs ? Différents types de biométrie sont acceptables dans les applications déployées auprès de différents groupes démographiques, en fonction des normes culturelles, éthiques, sociales, religieuses et hygiéniques de la société concernée. L'acceptabilité d'une biométrie dans une application est souvent un compromis entre la sensibilité d'une communauté à diverses perceptions/tabous et la valeur/convenance offerte par la reconnaissance biométrique.

Pratiquement, aucune biométrie n'est censée répondre efficacement aux exigences de toutes les applications. La correspondance entre un élément biométrique et une application est déterminée en fonction des caractéristiques de l'application et des propriétés de l'élément biométrique, donc le choix de la modalité a des implications sur la conception du système et, potentiellement, sur ses performances.

Nous présentons ci-dessous une brève introduction sur les modalités biométriques

les plus répandues, qui ont été adoptées dans des systèmes commerciaux, des systèmes de contrôle d'accès physique ou qui font l'objet de recherches récentes. Nous abordons rapidement les empreintes digitales dans cette liste, car elles sont largement couvertes dans le reste de ce rapport.

2.2.1.1 Le visage

Le visage est l'un des éléments biométriques les plus acceptés car il s'agit de l'une des méthodes de reconnaissance les plus courantes que les humains utilisent dans leurs interactions visuelles. Bien que la performance d'authentification des systèmes de reconnaissance faciale disponibles soit acceptable dans certaines applications, ils imposent un certain nombre de restrictions sur la façon dont les images faciales sont obtenues, nécessitant souvent un arrière-plan fixe et simple avec un éclairage contrôlé. Ces systèmes ont des difficultés à faire correspondre des images de visages capturées à partir de deux vues différentes, dans des conditions d'éclairage différentes, et à des moments différents. La méthode d'acquisition des images du visage est non intrusive. Le déguisement des visages est un problème dans les applications de reconnaissance sans surveillance. Il est très difficile de développer des techniques de reconnaissance des visages qui peuvent tolérer les effets des expressions faciales, les légères variations de l'environnement d'imagerie et les variations de la pose du visage par rapport à la caméra. Le temps qui s'écoule entre l'inscription dans un système et la tentative de reconnaissance peut également constituer un défi, car l'apparence du visage change avec le temps surtout en cas de vieillissement.

2.2.1.2 L'iris

L'iris est la région annulaire de l'œil délimitée par la pupille et la sclérotique (blanc de l'œil) de chaque côté. L'image de l'iris est généralement obtenue par un procédé d'imagerie sans contact. La capture d'une image de l'iris implique la coopération de l'utilisateur, à la fois pour enregistrer l'image de l'iris dans la zone centrale d'imagerie et pour s'assurer que l'iris se trouve à une distance prédéterminée du plan focal de la caméra. La texture complexe de l'iris porte des informations très distinctives utiles à la reconnaissance des personnes. La précision et la vitesse des systèmes de reconnaissance basés sur l'iris actuellement déployés sont prometteuses et peuvent prendre en charge l'identification à grande échelle. Bien que les systèmes de reconnaissance basés sur l'iris de première génération nécessitassent une participation considérable des utilisateurs et étaient coûteux, les systèmes plus récents sont devenus plus conviviaux et plus rentables.

2.2.1.3 La voix

La voix est une combinaison de caractéristiques biométriques physiques et comportementales. Les caractéristiques physiques de la voix d'un individu sont basées sur la forme et la taille des appendices (par exemple, le conduit vocal, la bouche, les cavités nasales et les lèvres) qui sont utilisés dans la synthèse du son. Ces caractéristiques physiques de la

parole humaine sont invariables pour un individu, mais les aspects comportementaux de la parole changent avec le temps en raison de l'âge, d'une condition médicale (comme un rhume), d'un état émotionnel, etc. La voix n'est pas non plus très distinctive et peut ne pas convenir à une identification à grande échelle. Un système de reconnaissance vocale dépendant du texte est basé sur l'énonciation d'une phrase prédéterminée. Un système de reconnaissance vocale indépendant du texte reconnaît le locuteur indépendamment de ce qu'il dit. Un système à guidage textuel invite l'utilisateur à répéter une phrase générée dynamiquement, ce qui offre une meilleure protection contre la fraude. Un inconvénient de la reconnaissance vocale est que les caractéristiques de la parole sont très sensibles à des facteurs tels que le bruit de fond et les caractéristiques du microphone. La reconnaissance du locuteur est la plus appropriée dans les applications téléphoniques, mais la qualité du signal vocal est généralement dégradée par le canal de communication.

2.2.1.4 L'empreinte digitale

L'empreinte digitale est la modalité la plus ancienne des sciences biométriques pour identifier ou vérifier l'identité des individus. Elle est la plus couramment déployée, utilisée dans un large éventail d'applications d'accès physique et logique. Toutes les empreintes digitales ont des caractéristiques distinctives et des motifs uniques. Il a été déterminé empiriquement que même les empreintes digitales des vrais jumeaux sont différentes, de même que les empreintes de chaque doigt d'une même personne [Maltoni *et al.*, 2009]. La précision des systèmes de reconnaissance par empreintes digitales actuellement disponibles est très élevée [Wilson *et al.*, 2004] pour les systèmes d'authentification dans plusieurs applications, notamment la criminalistique. D'autre part, l'arrivée de scanners des empreintes digitales peu coûteux et compacts a donné naissance à un grand nombre d'applications commerciales au cours des dernières années. Les principaux inconvénients de la modalité des empreintes digitales sont leur caractère intrusif et leur lien avec l'identification criminelle, ainsi que la production d'erreurs avec des doigts vieux, secs et sales. Il faut noter que l'empreinte digitale constitue la modalité principale de nos recherches.

2.2.1.5 La géométrie de la main

L'idée fondamentale de l'authentification géométrique basée sur la main est construite autour de l'hypothèse que tout le monde possède une main et que sa forme est unique. Cette technique est très simple, relativement facile à utiliser et peu coûteuse [Ma *et al.*, 2004]. Les systèmes biométriques de reconnaissance de la main mesurent et analysent la structure globale, la forme et les proportions de la main, par exemple la longueur, la largeur et l'épaisseur de la main, des doigts et des articulations, les caractéristiques de la surface de la peau. Les facteurs environnementaux opérationnels tels que le temps sec, ou les anomalies individuelles telles que la peau sèche, n'ont généralement pas d'effets négatifs sur la précision de l'identification. Le principal inconvénient de cette technique est sa faible capacité de discrimination. Les informations relatives à la géométrie de la main peuvent ne pas être invariables tout au long de la vie d'un individu, en

particulier pendant l'enfance.

2.2.1.6 La signature

Chaque personne a un style d'écriture unique, et les signatures de deux personnes différentes ne sont pas identiques. Cependant, les variations d'une signature typique dépendent également de l'état physique et émotionnel d'une personne. La précision d'identification des systèmes basés sur cette biométrie hautement comportementale est raisonnable mais ne semble pas suffisamment élevée pour permettre une reconnaissance à grande échelle. Il existe deux approches de l'identification basée sur la signature : statique et dynamique [Patel *et al.*, 2015]. L'identification statique de la signature utilise uniquement les caractéristiques géométriques (forme) d'une signature, tandis que l'identification dynamique (en ligne) de la signature utilise à la fois les caractéristiques géométriques (forme) et les caractéristiques dynamiques telles que les profils d'accélération, de vitesse, de pression et de trajectoire de la signature. Un avantage inhérent à un système biométrique basé sur la signature est que la signature a été établie comme une forme acceptable d'identification personnelle et peut être incorporée de manière transparente dans les processus commerciaux existants nécessitant des signatures, comme les transactions par carte de crédit. Cette modalité nécessite un faible temps de vérification et un équipement relativement peu coûteux [Duta, 2009] .

2.2.1.7 Les Veines

La structure des veines est l'une des caractéristiques les plus stables de la vie d'une personne [Bleumer, 1999]. L'authentification par les veines permet de capturer des images des motifs uniques des veines sur la rétine de l'œil, le doigt ou bien la main en faisant passer une lumière infrarouge en enregistrant l'effet via un capteur. Cette technique est plus pratique et moins intrusive que de nombreuses autres méthodes et fournit des résultats très précis, proches de ceux des systèmes d'identification par iris [Ye *et al.*, 2016]. Les systèmes de capture des veines utilisent des diodes électroluminescentes (DEL) infrarouges peu coûteuses, ce qui constitue un facteur inhibant l'utilisation généralisée de cette technique.

2.2.1.8 La rétine

La vasculature rétinienne est riche en structure et est censée être distinctive pour chaque individu et chaque œil. La biométrie de la rétine est généralement considérée comme la méthode biométrique la plus sûre [Cavoukian, 1999] car il n'est pas facile de modifier ou de répliquer la vascularisation rétinienne. Les scanners de la rétine comparent les vaisseaux sanguins de l'œil. Un dispositif de balayage qui utilise une faible lumière compare des motifs uniques sur la rétine. La présence de lunettes a un effet négatif sur le balayage de la rétine. Un scanner de la rétine produit au moins le même volume de données qu'une image d'empreinte digitale. En pratique, le balayage de la rétine est surtout utilisé pour la vérification. Le balayage rétinien n'est que rarement utilisé aujourd'hui car il n'est

pas convivial et reste très coûteux. Le balayage de la rétine convient aux applications qui exigent une sécurité élevée et pour lesquelles l'acceptation de l'utilisateur n'est pas un aspect majeur [Matyáš et Rıha, 2000].

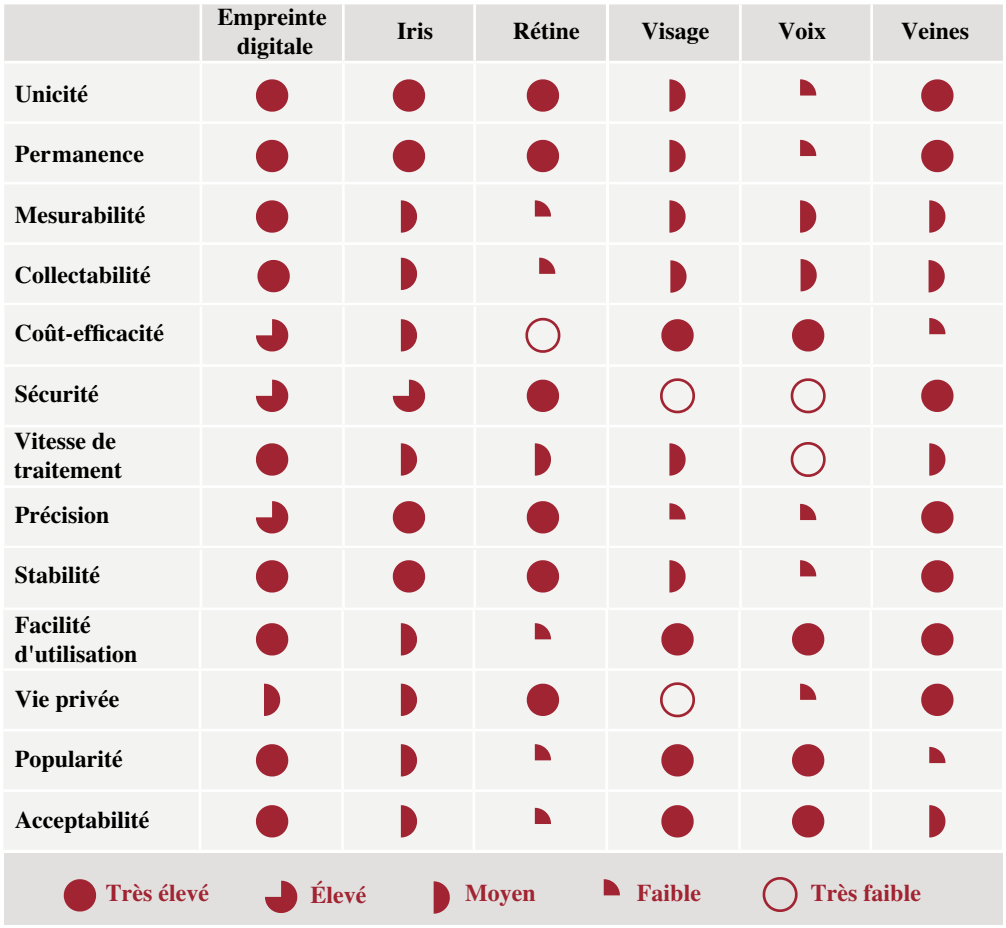


FIGURE 2.2 – Comparaison de quelques modalités biométriques

D'après la figure 2.2 qui décrit une comparaison de quelques modalités biométriques, il est clair que la reconnaissance par empreintes digitales constitue un très bon équilibre de toutes les propriétés souhaitables. Les empreintes digitales humaines sont détaillées, presque uniques, difficiles à modifier et durables tout au long de la vie d'un individu, ce qui en fait des marqueurs à long terme de l'identité. Il est donc naturel qu'il existe une large gamme de systèmes de reconnaissance par empreintes digitales destinés à être utilisés dans des applications de haute sécurité et pour l'identification automatique des personnes. Les caractéristiques de commodité, de sécurité et de performance ont fait de la reconnaissance des empreintes digitales la technologie d'authentification biométrique la plus utilisée aujourd'hui. Les capteurs des empreintes digitales à balayage direct peuvent facilement capturer des images de haute qualité et ils ne souffrent pas du problème de

la segmentation de l'empreinte digitale par rapport à l'arrière-plan (contrairement à la reconnaissance des visages, par exemple). La reconnaissance par empreintes digitales est l'une des technologies biométriques les plus matures et convient à un grand nombre d'applications de reconnaissance. Cela se reflète également dans le marché de la biométrie aussi bien sur le plan privé que public ou gouvernemental.

2.2.2 Modéles biométriques

Un modèle biométrique (également appelé Gabarit et Template) est une référence numérique de caractéristiques distinctes qui ont été extraites d'un échantillon biométrique d'un utilisateur. Une fois les données biométriques brutes acquises, elles sont analysées et converties utilisant un algorithme propriétaire en un fichier numérique qui décrit de manière optimale l'identité de l'utilisateur. Ce dernier peut être aussi bien un fichier mathématique binaire qu'un modèle statistique. Ce sont ces fichiers mathématiques qui sont connus sous le nom de modèles biométriques, et non les images qui ont été extraites et créées. Pour le stockage des modèles, il existe quatre emplacements principaux que sont la clé USB, la base centralisée, la machine individuelle de travail ou 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.

Bien évidemment, la conception du modèle biométrique dépend entièrement de la nature de la modalité biométrique concernée, ainsi que de la méthode d'authentification employée (Vérification / Identification). Cependant, nous pouvons identifier plusieurs variantes communes de création de modèles biométriques. Nous pouvons distinguer trois grandes catégories de types de modèles biométriques :

- Modèle biométrique à une seule référence : Une seule capture de bonne qualité est nécessaire lors de la phase d'enregistrement (l'enrôlement). Cette unique capture sert de référence durant la vérification de l'identité.
- Modèle biométrique à plusieurs références : Plusieurs captures de bonne qualité sont nécessaires lors de l'enrôlement afin de tenir compte des variations intra-classes associées à l'utilisateur. Nous parlons de galerie pour désigner l'ensemble des captures de références stockées dans le modèle biométrique.
- Modèle biométrique à grappes de références [Lumini et Nanni, 2006]. Il s'agit d'un cas particulier du type précédent. Les références (ou groupes de références) sont organisées sous forme hiérarchique. Chaque branche de l'arbre correspond à une contrainte particulière. Cette contrainte peut être explicite (profile, visage, luminosité, . . .) ou implicite (qualité de 0.1, qualité de 0.5, . . .).

Cependant, cette distinction n'est pas encore totalement parfaite. Pour la plupart des méthodes d'authentification biométrique, ce ne sont pas les données brutes qui sont stockées dans le modèle biométrique, mais le résultat d'un calcul (cf. transformée de fourrier, détection de points d'intérêt, information de texture, modèle statistique sur l'ensemble des

éléments de la galerie. . .). Dans le cas des systèmes à plusieurs références, le calcul peut être fait sur l'ensemble des références. Dans ce cas, le modèle biométrique ne correspond plus à une galerie, mais à une référence unique calculée grâce aux différents éléments de la galerie.

2.2.3 Application et usage de la biométrie

Le besoin de disposer de mécanismes d'authentification fiable des utilisateurs s'est accru au long des dernières décennies en raison des préoccupations croissantes en matière de sécurité et des progrès rapides dans le contexte des réseaux ouverts, les communications et les déplacements. Ainsi, la biométrie a été progressivement incorporée dans la majorité des domaines qui nécessitent de vérifier ou de déterminer les identités des individus. En effet, elle couvre aujourd'hui un champ d'application très vaste. Les applications biométriques se répartissent désormais en trois grandes catégories :

- Les applications commerciales (commerce électronique, utilisation de guichet automatique bancaires ou de cartes de crédit, le contrôle physique).
- Les applications gouvernementales (carte d'identité nationale, passeport électronique, permis de conduire, contrôle des frontières, la gestion des détenus dans les établissement pénitentiaires).
- Les applications médico-légales (les tests de paternité et la recherche de la relation familiale, l'identification de cadavres, les enquêtes criminelles).

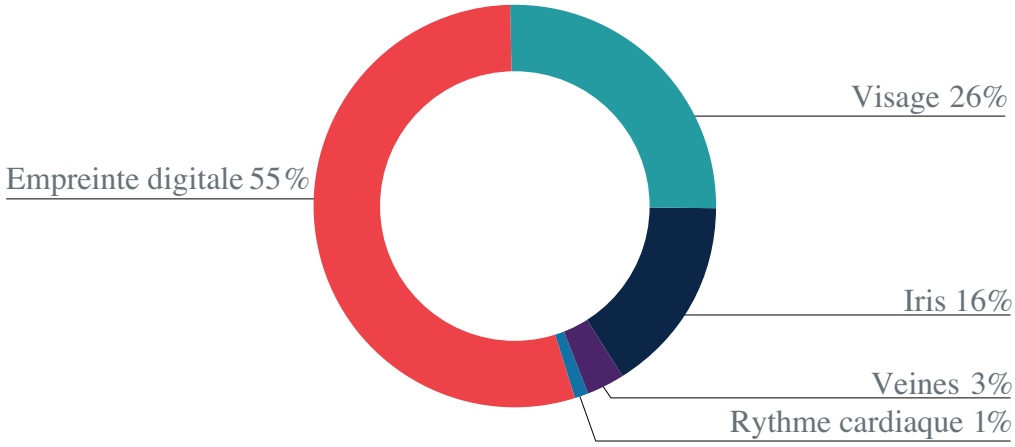


FIGURE 2.3 – Part de marché des revenus par type, y compris les marchés des consommateurs, des entreprises, des banques, services financiers et assurances, santé et gouvernement et sécurité.

À l'ère d'Internet et des progrès des technologies des smartphones, l'utilisation de la biométrie pour faciliter le processus d'authentification des utilisateurs devient très courante. Selon les prévisions, tous les smartphones seront équipés de fonctionnalités biométriques. Dans un contexte de numérisation de l'économie et de mise en relation des identités des consommateurs, il n'est pas étonnant de prédire que la biométrie deviendra un véritable allié pour vérifier les identités numériques dans le monde mobile. Ce tournant n'est pas uniquement inspiré par les besoins des utilisateurs mais il est également impulsé par des institutions publiques et des gouvernements qui passent à la biométrie pour développer des systèmes d'identification à grande échelle, à la fois fiables et pratiques.

La figure 2.3, réalisée d'après les chiffres de *ABI Research 2019* (une société d'études de marché et de renseignements commerciaux basée à New York.), montre les parts de marché des principales technologies biométriques en 2019. 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.

2.3 Généralités sur les systèmes biométriques

Tout système de reconnaissance de formes permettant d'authentifier un utilisateur en déterminant la validité d'une caractéristique physiologique ou comportementale spécifique est fondamentalement un système biométrique. Avec autant de modalités biométriques différentes, il semblerait que chaque système biométrique prenant en charge ces modalités soit unique. Cependant, les systèmes biométriques ont beaucoup de points communs entre eux. Les composants biométriques sont généralement similaires en termes de fonction. En outre, tous les systèmes biométriques partagent les mêmes préoccupations en ce qui concerne l'acceptation, la fraude, le stockage des données et la confidentialité.

2.3.1 Architecture

Un système biométrique se compose d'éléments matériels et logiciels. Le matériel comprend généralement des composants électroniques et des capteurs capables de lire des données à partir des échantillons biométriques, tandis que la partie logicielle utilise des algorithmes pour améliorer et reconnaître ces données afin de générer un modèle unique pour l'individu dont elles proviennent. L'architecture d'un système biométrique est la représentation d'un système dans son ensemble, y compris les fonctionnalités sur les composants matériels et logiciels, et l'interaction humaine avec ces composants. Les composants des systèmes biométriques peuvent varier d'un système à l'autre, mais un système biométrique généralisé est une combinaison fonctionnelle de cinq principaux composants à savoir : 1) le capteur, 2) le traitement du signal, 3) la base de données (stockage des modèles), 4) la classification, 5) la prise de décision, comme l'illustre la figure 2.4 :

Un système biométrique peut être considéré comme un système de traitement de signal avec une architecture de reconnaissance de formes. Il capte le signal biométrique, le

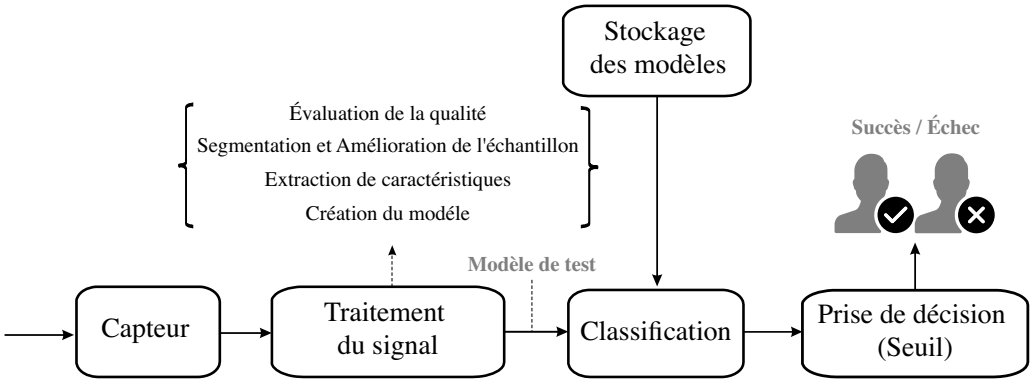


FIGURE 2.4 – Architecture générique d'un système biométrique

traite puis extrait un ensemble de caractéristiques représentatives (modéles biométriques) qui seront ensuite comparées au modèle préalablement stocké sur une base de données.

2.3.1.1 Capteur

Le capteur est le seul point d'interaction entre l'utilisateur et le système biométrique. Il existe généralement un bon nombre de capteurs biométriques, ils se distinguent, notamment par : leur type (sans ou avec contact), leur technologie, leur coût, leur qualité d'acquisition et leur facilité d'intégration. Ce composant peut effectuer différentes tâches, en fonction du système biométrique et des données qu'il recueille. Dans le cas d'une authentification, le capteur biométrique est le responsable de l'acquisition de l'échantillon biométrique (sous forme d'image ou de signal) de l'individu. La qualité de l'échantillon biométrique et la manière dont l'utilisateur présente les caractéristiques biométriques au capteur ont un grand impact sur les performances à long terme du système biométrique. Des données d'acquisition de mauvaise qualité se propageront dans le reste du système et entraîneront des taux d'erreur élevés. En toute équité, le capteur est considéré comme le composant le plus pertinent d'un système biométrique.

2.3.1.2 Traitement du signal

Le traitement du signal est chargé d'extraire les caractéristiques de l'échantillon biométrique afin de générer un modèle qui représente l'unicité de l'échantillon. Habituellement, les données biométriques brutes provenant du capteur sont soumises à des opérations de pré-traitement avant que les caractéristiques n'en soient extraites. Le processus de traitement du signal comprend généralement : 1) La phase de pré-traitement qui consiste à évaluer la qualité de l'acquisition, à appliquer une segmentation et enfin à améliorer l'échantillon, et 2) La phase d'extraction de caractéristiques.

La première étape du pré-traitement permet d'évaluer la qualité des échantillons biométriques acquis afin de déterminer s'ils conviennent à un traitement ultérieur. Si les données brutes ne sont pas d'une qualité suffisamment bonne, soit une autre tentative d'acquisition des données auprès de l'utilisateur est demandée, soit une exception (alarme

d'échec) est déclenchée. L'étape suivante du pré-traitement est la segmentation, dont l'objectif est de séparer les données biométriques requises du bruit de fond. Par la suite, les données biométriques segmentées sont soumises à un algorithme d'amélioration de la qualité du signal afin d'améliorer leur qualité et de réduire d'avantage le bruit avant de passer à la phase d'extraction de caractéristiques. Le traitement du signal est extrêmement important pour la précision d'un système biométrique car la qualité de l'extraction des caractéristiques a un effet sur le processus de génération du modèle biométrique.

2.3.1.3 Base de données

La base de données du système biométrique fait office de référentiel des données biométriques. Lorsqu'il s'agit d'un système de vérification biométrique, l'ensemble des modèles générés sont stockés dans la base de données avec certaines informations d'identité personnelles (telles que le nom, le numéro d'identification personnel (PIN), l'adresse, etc.) caractérisant l'utilisateur. L'une des décisions clés dans la conception d'un système biométrique est de savoir s'il faut utiliser une base de données centralisée ou décentralisée, ou même des dispositifs portables (jeton) tel qu'une carte à puce, un support de stockage personnel, etc. Le stockage de tous les gabarits dans une base de données centrale peut être bénéfique du point de vue de la sécurité du système, car les données peuvent être sécurisées par une isolation physique et par des mécanismes de contrôle d'accès stricts. D'autre part, la compromission d'une base de données centrale aurait des implications bien plus importantes que la compromission d'un des sites de la base de données décentralisée. En effet, des personnes malveillantes (administrateurs corrompus ou pirates informatiques) peuvent abuser des informations biométriques stockées dans la base de données pour compromettre la vie privée des utilisateurs.

2.3.1.4 Classification

L'objectif d'une classification biométrique est de comparer les caractéristiques biométriques d'une requête avec un ou plusieurs modèles stockés dans la base de données afin de générer des scores de correspondance. Un score de correspondance est une mesure de similarité entre un modèle de test et un modèle de référence. Un score plus élevé indique une plus grande similarité entre les deux modèles. Dans le cas où il s'agit de mesurer la dissimilarité (au lieu de la similarité) entre les deux ensembles de caractéristiques, le score est appelé score de distance. Un score de distance plus petit indique une plus grande similarité. Le score de correspondance peut être modéré en fonction de la qualité des données biométriques présentées. Une fois le score est calculé, il est transféré directement au composant de prise de décision.

2.3.1.5 Prise de décision

Ce composant reçoit le score de correspondance en entrée du composant de classification pour le comparer au seuil de verification ou d'identification. Le seuil est une valeur

prédéfinie, normalement choisie par l'administrateur du système biométrique. Si le score de similarité dépasse le seuil, les modèles comparés correspondent, si le score est inférieur de la valeur seuil, les modèles comparés ne correspondent pas [Modi, 2011]. Le composant de décision produit le résultat, également appelé décision, de la comparaison entre le score similarité et la valeur seuil. Le résultat du composant de décision peut être une correspondance, une non-correspondance ou un résultat non concluant. Ces résultats sont liés à la valeur seuil et au score de comparaison. Une correspondance peut conduire à une authentification réussie, une non-correspondance peut conduire à une authentification infructueuse, tandis qu'une décision non concluante peut exiger du sujet qu'il présente un autre échantillon au système.

2.3.2 Fonctionnement

Les systèmes de reconnaissance biométrique comprennent deux étapes clés de fonctionnement à savoir : 1) L'étape d'inscription connue sous le nom de « Enrôlement » et 2) l'étape d'authentification qui peut se manifester soit sous forme d'un processus de vérification ou d'identification.

2.3.2.1 Enrôlement

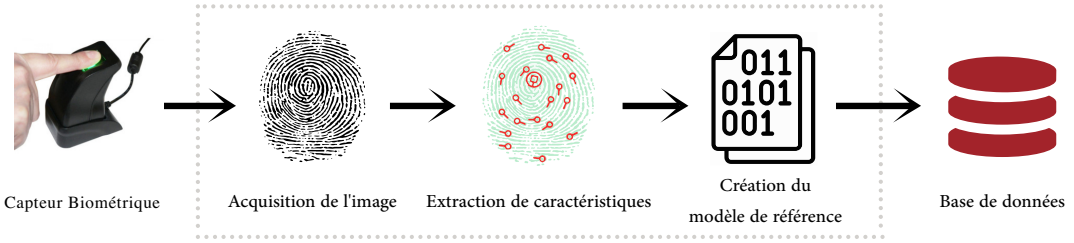


FIGURE 2.5 – Les étapes de la phase d'enrôlement.

L'enrôlement est la première phase de tout système biométrique. Il s'agit de l'étape pendant laquelle un utilisateur est inscrit dans le système pour la première fois. Au cours de ce processus, l'utilisateur présente ses données biométriques auprès d'un dispositif d'acquisition (capteur biométrique), elles sont évaluées puis traitées pour extraire une forme numérique. Cette représentation est ensuite réduite à travers des algorithmes d'extraction bien défini afin de réduire la quantité de données à stocker pour ainsi faciliter les authentifications futures. Dans un système biométrique pratique, la phase d'enrôlement comprend les étapes suivantes (Figure 2.5) : 1) L'acquisition des données biométriques, 2) le pré-traitement qui consiste à améliorer l'échantillon biométrique, à vérifier la qualité de l'échantillon capturé (le système peut le rejeter ou l'accepter en fonction d'un score de qualité) et à extraire les caractéristiques saillantes et distinctives. 3) La création du modèle de référence (qui peut nécessiter plusieurs échantillons). 4) La conversion éventuelle du modèle dans un format d'échange de données et de stockage. L'enrôlement est le processus le plus critique du système biométrique. Rien d'autre ne peut affecter l'utilisation

réussie des technologies biométriques plus que l'enrôlement. La qualité de cette phase est un facteur critique pour la précision à long terme des technologies biométriques. Plus les enrôlements sont de mauvaise qualité (qualité des modèles), moins le système sera précis en général et plus les taux d'erreur seront élevés, notamment le taux de fausses acceptations et le taux de faux rejets. Le fait d'éviter les images détériorées générées pendant le processus d'enrôlement devrait en fait améliorer la précision du système biométrique [Jain *et al.*, 2007, Jain *et al.*, 2006a] .

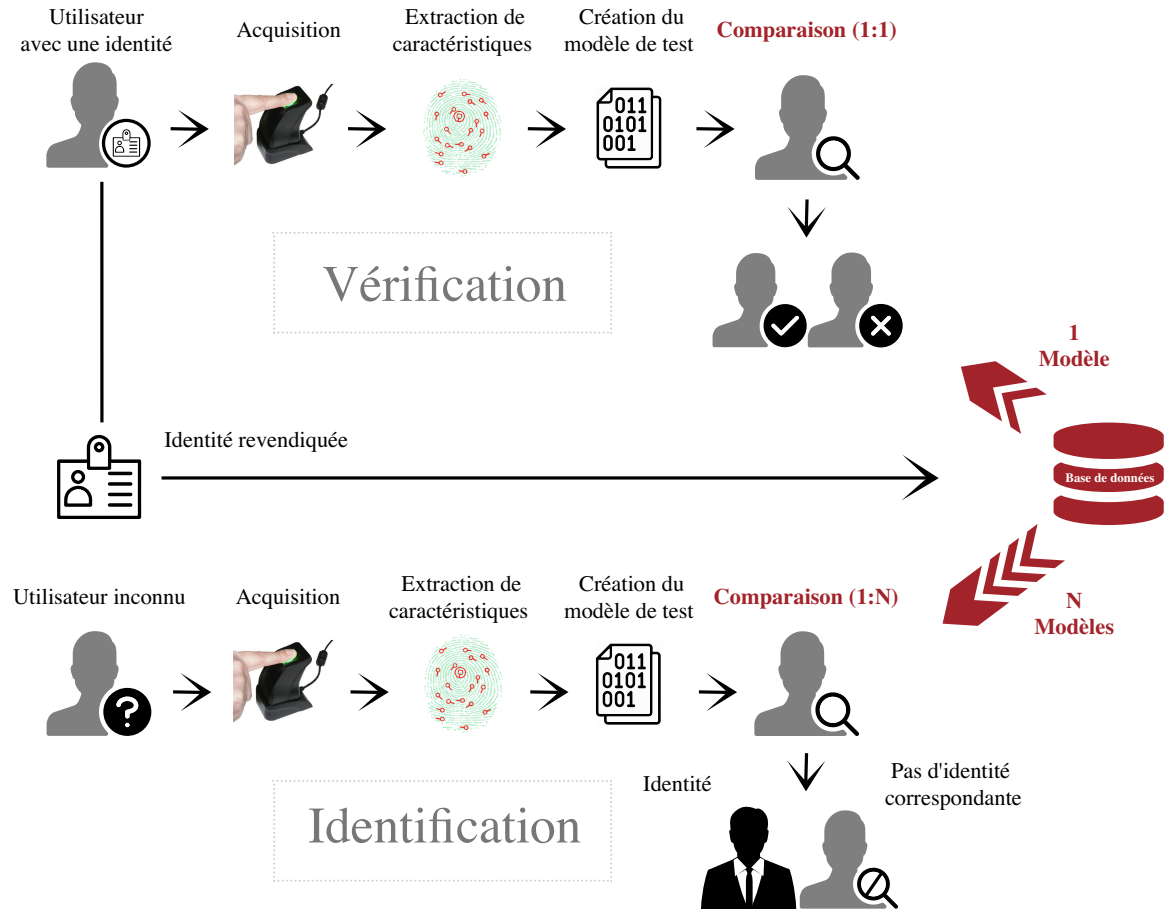


FIGURE 2.6 – Différence entre le processus de vérification et d'identification.

Généralement, les systèmes d'authentification biométriques sont en mesure de mener deux fonctions principales dans lesquelles les données biométriques sont différemment exploitées, à savoir la vérification et l'identification. En effet, selon le contexte de l'application, un système d'authentification biométrique peut être qualifié de système de vérification ou de système d'identification (Figure 2.6).

2.3.2.2 Vérification

Un système de vérification valide l'identité d'une personne en comparant les données biométriques capturées avec son ou ses propres modèles biométriques de référence préala-

blement enrôlés et stockés dans la base de données. Au cours de ce processus, le système fournit la réponse à la question "Suis-je celui que je prétends être ?" en exigeant qu'un individu revendique une identité généralement par le biais d'un code PIN, d'un nom d'utilisateur ou d'une carte à puce pour qu'une comparaison biométrique (correspondance) soit effectuée. Le système de vérification biométrique acquiert les données biométriques d'un individu, puis extrait les caractéristiques de l'échantillon biométrique afin de générer le modèle de test. Une comparaison (1 : 1) est effectuée entre le modèle biométrique de test et uniquement le modèle biométrique de référence stocké (sélectionné par rapport à l'identité réclamé) pour déterminer en fonction de la politique de décision (valeur seuil) si la revendication est vraie (l'individu est considéré comme "authentique") ou non (l'individu est considéré comme "imposteur"). Le processus de vérification peut être formalisé comme suit :

Soit le vecteur V_U représentant les caractéristiques biométriques de l'utilisateur U extraites par le systéme, et T_U son modèle biométrique stocké dans la base de données, le système renvoie une valeur booléenne suite au calcul de la fonction f définie par :

$$f(V_U, T_U) = \begin{cases} 1 & \text{si } S(V_U, T_U) \geq \tau, \\ 0 & \text{sinon.} \end{cases} \tag{2.1}$$

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

2.3.2.3 Identification

Un système d'identification permet de répondre à la question "Qui suis-je ?" sans revendiquer une identité, c'est le système qui révèle plutôt l'identité associée aux caractéristiques biométriques présentées, avant même que la comparaison ne soit établie. L'individu dans ce cas est supposé fournir que ses données biométriques. Le système reconnaît un individu en parcourant les modèles biométriques de références de tous les utilisateurs de la base de données pour chercher une correspondance. Le système effectue une comparaison un à plusieurs (1 : N) entre le modèle de test et tous les modèles de références stockés, si l'un de ces derniers correspond relativement au modèle de test, le système indique cette identité, sinon l'individu est rejeté. Le processus d'identification peut être formalisé comme suit :

Soit le vecteur V_U représentant les caractéristiques biométriques extraites par le système lorsqu'un utilisateur U se présente devant celui-ci, l'identification revient à déterminer l'identité de $\{I_t \mid t \in 0, 1, ..., N\}$ où $I_1, ..., I_N$ sont les N 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 :

$$f(V_U) = \begin{cases} I_k & \text{si } max_{1 \leq k \leq N} S(V_U, T_U) \geq \tau, \\ I_0 & \text{sinon.} \end{cases} \tag{2.2}$$

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

2.4 Vulnérabilités des systèmes biométriques

2.4.1 Limites

Bien que les technologies biométriques aient été de plus en plus utilisées pour la reconnaissance automatique des personnes en raison des avantages inhérents qu'elles offrent par rapport aux méthodes de reconnaissance traditionnelles, elles souffrent enfaite de plusieurs limitations qui peuvent dégrader considérablement leur intérêt. En effet, contrairement aux systèmes d'authentification traditionnels, les systèmes d'authentification basés sur la biométrie sont moins précis, avec un pourcentage de similarité compris entre 0% et 100%, le 100% n'étant presque jamais atteint. Ce manque de précision est lié à plusieurs raisons : la variabilité lors de la capture (le bruit d'acquisition, l'utilisation de plusieurs capteurs d'acquisition, etc.), la variabilité intra-classe (variabilité des données biométriques pour un individu) et la similarité inter-classe (c'est-à-dire la similarité des données biométriques de plusieurs individus).

L'utilisation de données biométriques soulève aussi de nombreux problèmes de sécurité qui sont propres aux systèmes de reconnaissance basés sur la biométrie et qui n'affectent pas les autres approches employées pour la reconnaissance automatique des personnes. Dans ce contexte, certaines données biométriques telles que la voix, le visage, les empreintes digitales et bien d'autres sont des traits exposés, ils ne sont pas secrets et peuvent donc être discrètement acquis ou volés par un attaquant et utilisés à mauvais escient. Cela peut conduire par exemple à une usurpation d'identité. De plus, les données biométriques brutes ne peuvent pas être révoquées, annulées ou réémises si elles sont compromises, puisqu'il s'agit de caractéristiques intrinsèques de l'utilisateur et qu'elles sont en nombre limité. Ainsi, si une biométrie est compromise, toutes les applications utilisant cette biométrie le sont aussi, et comme les identifiants biométriques sont permanents, le fait de devoir les changer soulève un sérieux problème.

En outre, le recours à la biométrie suscite également plusieurs questions relatives à la confidentialité, en effet, lorsqu'un individu donne ses biométries, volontairement ou non, il divulgue des informations uniques sur lui-même. Il a également été démontré que les données biométriques peuvent contenir des informations pertinentes sur la santé des personnes. Ces informations peuvent être utilisées, par exemple, pour discriminer des personnes à l'embauche ou pour refuser une assurance à ceux qui ont des problèmes de santé latents. L'utilisation des données biométriques peut également susciter des préoccupations d'ordre culturel, religieux ou ethnique. Dans une certaine mesure, la perte de l'anonymat

peut être directement perçue par les utilisateurs comme une perte d'autonomie.

Une autre limite de la biométrie est celle relative à son usage. La biométrie et en particulier les empreintes digitales ont une mauvaise réputation et sont associées à la surveillance des personnes et à l'identification des criminels. Selon la modalité utilisée, l'acquisition des données biométriques se fait sans ou avec contact avec le capteur biométrique. Ce contact est une préoccupation pour certains utilisateurs pour des raisons d'hygiène et d'intrusion physique.

2.4.2 Défaillance d'un système biométrique

Un système biométrique est vulnérable à différents types d'attaques qui peuvent compromettre la sécurité offerte par le système, entraînant ainsi sa défaillance. Les modes de défaillance d'un système biométrique peuvent être classés en deux catégories : défaillance intrinsèque et défaillance due à une attaque adverse. Les défaillances intrinsèques sont dues à des limitations inhérentes aux technologies de détection, d'extraction de caractéristiques ou de mise en correspondance, ainsi qu'au caractère discriminatoire limité du trait biométrique utilisé. Dans les attaques adverses, un attaquant ingénieux (ou éventuellement un groupe organisé) tente de contourner le système biométrique à des fins personnelles.

2.4.2.1 Défaillance intrinsèque

La défaillance intrinsèque est un échec de sécurité due à une décision incorrecte prise par le système biométrique. Un système de vérification biométrique peut commettre deux types d'erreurs dans la prise de décision, à savoir une fausse acceptation et un faux rejet. Un utilisateur légitime peut être faussement rejeté par le système biométrique en raison des grandes différences entre le modèle de référence de l'utilisateur et les ensembles de caractéristiques biométriques d'interrogation. Ces variations intra-classe peuvent être dues à une interaction incorrecte de l'utilisateur avec le système biométrique (par exemple, des changements de pose et d'expression dans une image de visage) ou au bruit introduit sur le capteur (par exemple, des empreintes résiduelles laissées sur un capteur des empreintes digitales). Les fausses acceptations sont généralement causées par le manque d'individualité ou d'unicité du trait biométrique, ce qui peut entraîner une grande similarité entre les ensembles de caractéristiques des différents utilisateurs (par exemple, la similarité des images de visage de jumeaux ou de frères et sœurs). Les variations intra-classe et la similarité inter-classe peuvent également être causées par l'utilisation de caractéristiques non saillantes et d'appariements non robustes. Parfois, un capteur peut ne pas réussir à acquérir le trait biométrique d'un utilisateur en raison des limites de la technologie de détection ou de conditions environnementales défavorables. Par exemple, un capteur d'empreinte digitale peut ne pas être en mesure de capturer une empreinte digitale de bonne qualité sur des doigts secs/mouillés. Cela entraîne des erreurs de type *failure-to-enroll* (FTE) ou *failure-to-acquire* (FTA).

Les défaillances intrinsèques peuvent se produire même en l'absence d'effort explicite d'un adversaire pour contourner le système. Ce type de défaillance est donc également connu sous le nom *Attaque à zéro-effort*. Il constitue une menace sérieuse si les probabilités de fausse acceptation et de faux rejet sont élevées. Les recherches en cours visent à réduire la probabilité de défaillance intrinsèque, principalement par la conception de nouveaux capteurs capables d'acquérir les caractéristiques biométriques d'un individu de manière plus fiable, plus pratique et plus sûre, par le développement de schémas de représentation invariants et d'algorithmes de correspondance robustes et fiables, et par l'utilisation de systèmes multibiométriques [Ross *et al.*, 2006].

2.4.2.2 Défaillance due à une attaque adverse

Un adversaire organise intentionnellement une attaque contre le système biométrique dont le succès dépend des failles dans la conception du système et de la disponibilité de ressources informatiques. Selon [Jain *et al.*, 2008], les attaques adverses peuvent être réparties en trois catégories principales : attaques administratives, infrastructure non sécurisée et dépassement biométrique.

Attaque administrative : Cette attaque, également connue sous le nom de *The insider attack*, fait référence à toutes les vulnérabilités introduites par une mauvaise administration du système biométrique. Il s'agit notamment de l'intégrité du processus d'enrôlement (par exemple, la validité des informations d'identification présentées pendant l'enrôlement), de la collusion entre l'adversaire et l'administrateur du système ou un utilisateur légitime, et de l'abus des procédures de traitement des exceptions.

Infrastructure non sécurisée : L'infrastructure d'un système biométrique comprend le matériel, les logiciels et les canaux de communication entre les différents modules. Un adversaire peut manipuler l'infrastructure biométrique de plusieurs façons, ce qui peut entraîner des failles de sécurité. Une discussion détaillée sur ces types d'attaques est présentée à la sous-section suivante 2.4.4.

Dépassement biométrique : Il est possible pour un adversaire d'acquérir secrètement les caractéristiques biométriques d'un utilisateur authentique (par exemple, des impressions d'empreinte digitale relevées sur une surface) et de les utiliser pour créer des artefacts physiques (Doigts gommeux) du trait biométrique. Par conséquent, si le système biométrique n'est pas capable de faire la distinction entre une présentation biométrique réelle et une usurpation artificielle, un adversaire peut contourner le système en présentant des traits usurpés.

2.4.3 Effets de la défaillance d'un système biométrique

La compromission d'un système biométrique peut engendrer deux effets principaux : 1) le déni de service et 2) l'intrusion.

Le déni de service : Désigne le scénario dans lequel un utilisateur légitime est empêché d'obtenir le service auquel il a droit. Un adversaire peut saboter l'infrastructure (par exemple, endommager physiquement le capteur biométrique) empêchant ainsi les utilisateurs d'accéder au système. Les défaillances intrinsèques telles que le faux rejet, l'échec de la capture et l'échec de l'acquisition conduisent également à un déni de service. Les abus administratifs tels que la modification des modèles ou des paramètres de fonctionnement (par exemple, le seuil de correspondance) du système biométrique peuvent également entraîner un déni de service.

L'intrusion : Fait référence à un imposteur obtenant un accès illégitime au système, ce qui entraîne une atteinte à la vie privée (par exemple, un accès non autorisé à des informations personnelles) et des menaces pour la sécurité (par exemple, des terroristes traversant les frontières). Les quatre facteurs à l'origine de la vulnérabilité des systèmes biométriques, à savoir la défaillance intrinsèque, les abus administratifs, l'infrastructure non sécurisée et l'excès de biométrie, peuvent tous entraîner une intrusion.

2.4.4 Niveaux d'attaque

Comme déjà mentionné, un système biométrique est décrit comme la cascade du capteur d'acquisition, du composant de traitement du signal, du module qui effectue la correspondance entre la sortie de l'extracteur de caractéristiques et les modèles stockés dans la base de données, et finalement du composant de décision. Chaque composant de cette structure peut être un sujet d'une attaque intentionnelle spécifique menée par un adversaire. En littérature, les vulnérabilités et les attaques contre un système biométrique ont été présentées sous plusieurs formes / modèles et à partir de plusieurs points de vue [Ratha *et al.*, 2001a, Ratha *et al.*, 2001b, Ratha *et al.*, 2003, Cukic et Bartlow, 2005, Adler, 2005, Jain *et al.*, 2006b, Jain *et al.*, 2008, Roberts, 2007]. Ratha et al. [Ratha *et al.*, 2001a] ont identifié huit niveaux différents d'attaques qui peuvent être lancées contre un système biométrique (Figure 2.7).

Ces attaques visent soit à contourner la sécurité offerte par le système, soit à dissuader le fonctionnement normal du système. Elles peuvent être classées de la manière suivante :

2.4.4.1 Attaques sur l'interface utilisateur (Capteur biométrique)

Les attaques au niveau de l'interface utilisateur définissent comment un attaquant peut accorder l'accès au système grâce à une soumission biométrique manipulée au capteur, Ces attaques sont le plus souvent dues à la présentation d'une fausse caractéristique

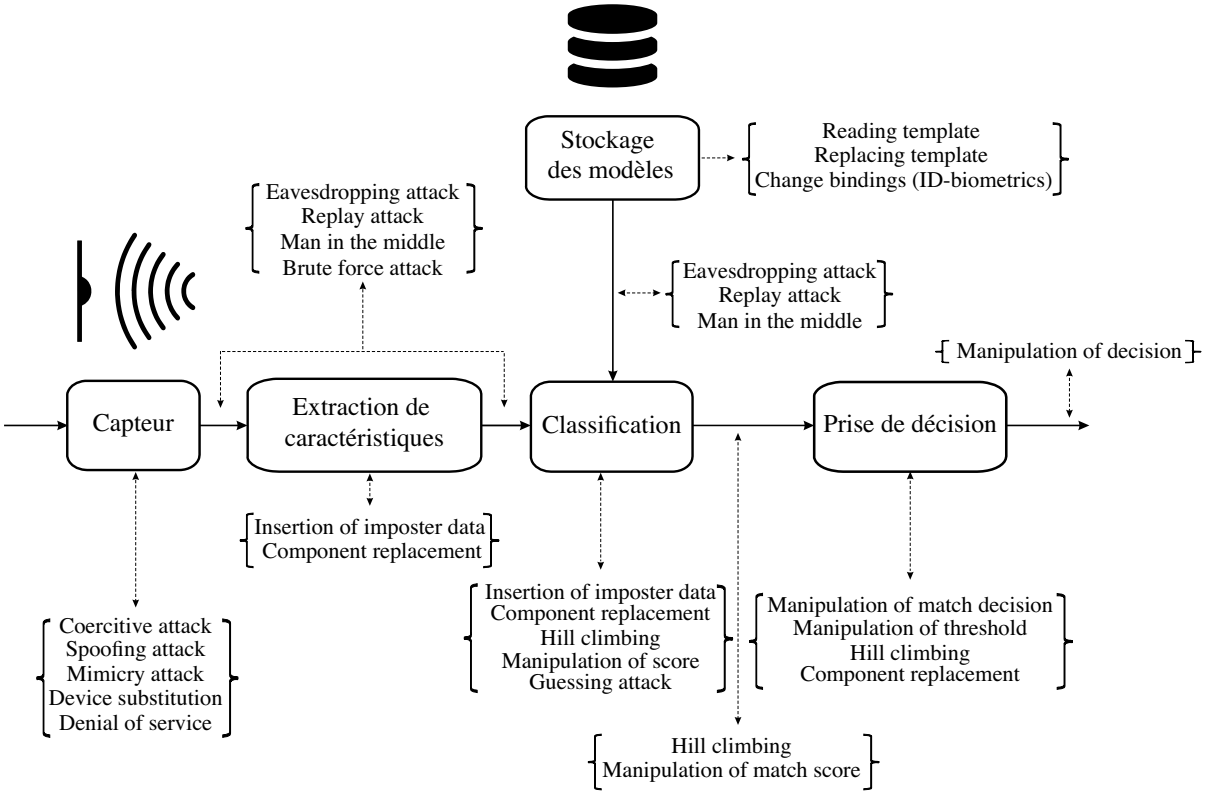


FIGURE 2.7 – Les niveaux d'attaque sur un système biométrique générique

biométrique (figure 2.8) [Eriksson et Wretling, 1997, Matsumoto *et al.*, 2002]. Si le capteur est incapable de distinguer les traits biométriques authentiques des faux, l'adversaire s'introduit facilement dans le système sous une fausse identité. Les mécanismes de sécurité mis en œuvre pour la protection numérique, tels que les cryptosystèmes, sont inefficaces dans de tels scénarios. Parmi les attaques les plus courantes sur l'interface utilisateur :

- *Coercive attack* : La vraie biométrie est présentée mais d'une manière non autorisée, par exemple lorsqu'un imposteur oblige un utilisateur légitime à lui accorder l'accès au système, ou bien l'utilisateur autorisé coopère tout simplement pour aider l'imposteur à obtenir la biométrie valide. En réalité, un tel scénario représente une faible menace dans la plupart des applications.
- *Spoofing attack/Mimicry attack* : Ces attaques sont liées respectivement aux biométries physiologiques et comportementales. Elles consistent à copier, par différentes stratégies, la caractéristique biométrique de l'utilisateur inscrit, et à la présenter au capteur afin de tromper le système (cela peut prendre la forme d'un doigt artificiel, d'un masque sur un visage ou d'une lentille sur un œil).
- *Device substitution* : Le remplacement d'un dispositif de capture biométrique légitime par une unité simulée ou modifiée.
- *Denial of service* : (La déscription de cette attaque est détaillé dans la sous-section 2.4.3).

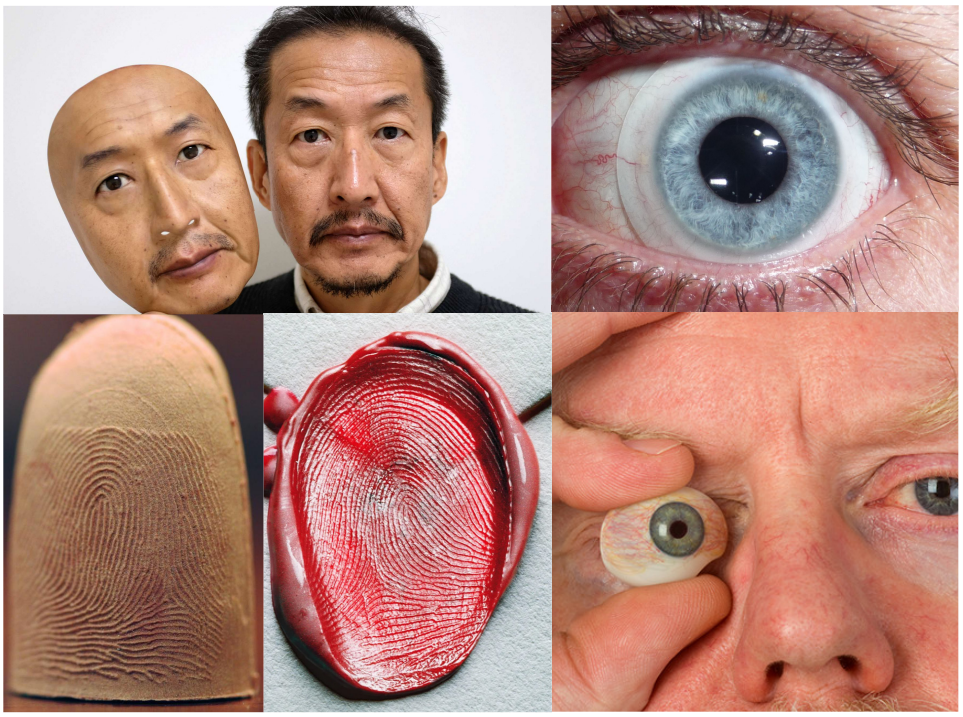


FIGURE 2.8 – Exemples de faux échantillons biométriques

2.4.4.2 Attaques sur l'interface entre les modules

Les canaux qui interconnectent les différentes modules d'un système biométrique, comme le canal entre le capteur et l'extracteur de caractéristiques, entre l'extracteur de caractéristiques et le classifieur, entre la base de données et le classifieur, et entre le classifieur et le module de prise de décision, peuvent être interceptés et contrôlés s'ils ne sont pas sécurisé physiquement ou cryptographiquement. En fait, un adversaire peut intercepter et/ou modifier les données transférées. Dans ce contexte, Juels et al. [Juels *et al.*, 2005] ont décrit les problèmes de sécurité et de confidentialité introduits par des canaux de communication non sécurisés dans une application de passeport électronique qui utilise l'authentification biométrique. Les canaux de communication non sécurisés permettent également à un adversaire de lancer des attaques de type *Replay Attacks* [Syverson, 1994] ou *Hill-Climbing Attacks* [Adler, 2005]. Une façon courante de sécuriser un canal est de coder cryptographiquement toutes les données envoyées par l'interface, par exemple en utilisant une infrastructure à clé publique. Mais même dans ce cas, un adversaire peut mettre en scène une *Replay Attack* en interceptant d'abord les données cryptées passant par l'interface lorsqu'un utilisateur authentique interagit avec le système, puis en envoyant ces données capturées au module souhaité lorsqu'il veut s'introduire dans le système. Parmi les attaques possibles, nous pouvons mentionner :

- *Attaque par écoute clandestine (Eavesdropping attack en Anglais)* : Le fait d'écouter discrètement la transmission des données biométriques. Elle repose sur des commu-

nications non sécurisées pour accéder aux données en transit entre les modules.

- *Man in the middle attack* : Un attaquant est capable de manipuler les messages échangés entre deux parties sans que celles-ci ne sachent que la liaison a été compromise.
- *Brute force attack* : Présentation exhaustive d'un grand nombre d'entrées biométriques au système de reconnaissance pour en trouver une qui correspont ;
- *Replay attack* : l'attaquant intercepte le moyen de communication puis modifie le contenu en injectant des fausses données afin d'accéder au système en tant qu'une personne légitime, puis retransmet les données au traitement. Il s'agit d'un type d'attaque *Man-in-the-middle* et qui exige que l'attaquant ait une bonne connaissance du système et de la base de données.
- *Hill climbing attack* : C'est une attaque dans laquelle le score de similarité donné par le classifieur est utilisé pour modifier itérativement un modèle généré synthétiquement, ou un groupe de gabarits, jusqu'à ce que le seuil de vérification soit atteint [Galbally *et al.*, 2009] (Annexe A.1). L'objectif de l'attaque est d'obtenir une autorisation en raison de la limitation intrinsèque des fausses acceptations du système biométrique.
- *Manipulation of match score* : L'adversaire peut faciliter le processus d'authentification en modifiant le score de correspondance intercepté entre le canal reliant le classifieur et le module de décision. le score généré pour la comparaison de modèles authentiques est remplacé par un score de correspondance supérieur au seuil prédéfini.
- *Manipulation of the decision* : C'est une attaque dans laquelle les résultats finaux de la décision sont capturés et modifiés.

2.4.4.3 Attaques sur les modules logiciels

Le programme exécutable d'un module peut être modifié de telle sorte qu'il émette toujours les valeurs souhaitées par l'adversaire. De telles attaques sont connues sous le nom d'attaques de type *Trojan-horse*. En effet, un attaquant peut concevoir un programme pour reproduire ou remplacer les fonctionnalités exécutées durant la phase d'extraction de caractéristiques ou la phase de classification.

Lors de l'étape d'extraction de caractéristiques, il se peut qu'un attaquant force à produire des caractéristiques pré-sélectionnées en insérant des données imposteurs. Par exemple, un adversaire peut développer un algorithme pour produire des images de test synthétiques qui ont une apparence similaire à celle des images acquises par le capteur utilisé. Cet algorithme peut alors transmettre les images au module d'extraction de caractéristiques en se faisant passer pour le capteur. Dans ce cas, le système peut ne pas être en mesure de différencier si les images proviennent de son capteur ou du logiciel malveillant de l'adversaire.

Concernant le classifieur, il peut être attaqué principalement pour produire de faux scores. Cette tâche peut être réalisée de différentes manières :

- *Manipulation of the match scores :* Le fait de capturer et de modifier la valeur d'un score de correspondance avant qu'il n'affecte la décision.
- *Reply attack :* Une version enregistrée des vraies données est injectée dans le canal.
- *Component replacement :* Substitution d'un des composants logiciels/matériels afin de contrôler son comportement :
- *Hill climbing attack :* C'est une attaque itérative [Adler, 2003a] qui peut être réalisée lorsque l'accès est accordé aux scores de correspondance. Plus précisément, étant donné une entrée, une légère modification de l'entrée est effectuée. Si le score de correspondance est augmenté, la modification est conservée, sinon la modification est rejetée. La procédure est itérée jusqu'à ce que le score de correspondance soit supérieur au seuil.

2.4.4.4 Attaques sur la base de modèles

L'une des attaques les plus sévères pour un système biométrique est celle qui porte sur les modèles biométriques stockés dans la base de données du système. Les attaques contre les modèles peuvent entraîner les vulnérabilités suivantes. 1) Un modèle peut être récupéré puis remplacé par le modèle d'un imposteur pour obtenir un accès non autorisé (*Reading/Replacing template*). 2) Une usurpation physique peut être créée à partir du modèle [Adler, 2004, Ross *et al.*, 2007] afin d'obtenir un accès non autorisé au système (ainsi qu'à d'autres systèmes qui utilisent le même trait biométrique). 3) Le modèle volé peut être rejoué face au module de correspondance. 4) Le changement du lien entre l'identité et le modèle biométrique *(Change bindings ID-biometrics)*.

2.5 Généralités sur l'empreinte digitale

Puisque l'empreinte digitale est la modalité biométrique étudiée dans cette thèse, une présentation détaillée sera abordé dans ce qui suit sur les caractéristiques de cette modalité, la reconnaissance par empreintes digitales et les défis qui font face à cette reconnaissance.

Depuis longtemps utilisées dans les enquêtes médico-légales et criminelles, les empreintes digitales sont les identifiants biométriques les plus répandus. Grâce aux progrès de la détection des empreintes digitales et aux développements rapides dans des domaines tels que le traitement des images et la reconnaissance des formes, les systèmes biométriques basés sur les empreintes digitales sont entrés dans une ère d'applications étendues dans les domaines commercial, civil et financier. L'authentification par empreinte digitale a été largement déployée dans des systèmes à petite et grande échelle pour le contrôle d'accès ou l'identification personnelle.

2.5.1 Description de l'empreinte digitale

Une empreinte digitale est le dessin formé par les lignes de la peau des doigts. Elles sont uniques et immuables, elles ne se modifient donc pas au cours du temps (sauf par accident) mis à part leur qualité qui peut se dégrader. Une empreinte digitale est constituée d'un ensemble de lignes localement parallèles formant un motif (*Ridge Pattern*) unique pour chaque individu. On distingue les crêtes, ce sont les lignes en contact avec une surface au touché et les vallées, ce sont les creux entre deux crêtes (figure 2.9). A l'intérieur de ce motif, il y a un très grand nombre d'éléments qui nous différencient les uns des autres. Ces caractéristiques sont formées par le flux des crêtes formant l'empreinte.

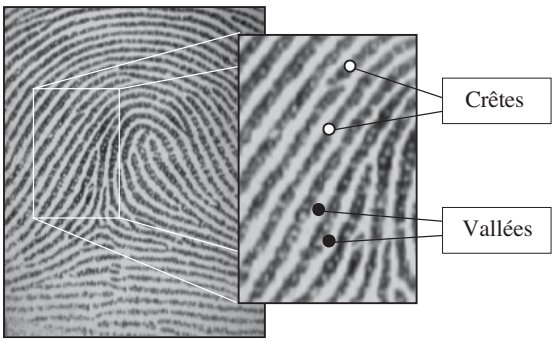


FIGURE 2.9 – Crêtes et vallées sur une image d'empreinte digitale.

Ces éléments sont à leur tour découplés en deux familles : les *Minuties* et les *Singularités*. La minutie est l'arrangement particulier des lignes papillaires (crêtes et vallées) à l'origine de l'individualité des empreintes. Les minuties peuvent être de différents types comme le montre la figure 2.10, mais en pratique, deux types seulement sont utilisés, à savoir les terminaisons (le point où la crête se termine) et les bifurcations (le point de carrefour de plusieurs crêtes). Cela s'explique par le fait que les autres types sont des combinaisons de terminaisons et de bifurcations.

Chacune de ces deux minuties (les bifurcations et les terminaisons) est désignée par sa classe, ses coordonnées x et y et l'angle entre la tangente à la ligne de crête à la position de la minutie et l'axe horizontal (figure 2.11). Dans la pratique, il existe une ambiguïté entre les points caractéristiques de terminaison et de bifurcation ; selon la pression exercée par le doigt sur la surface où l'empreinte est laissée, les terminaisons peuvent apparaître comme des bifurcations et vice versa.

Il existe deux points de singularités (figure 2.12) : le *Core* et le *Delta*. Le Delta est localisé à la confluence de trois différentes crêtes. Le Core est le point de courbure maximale.

2.5.2 Représentation du modèle d'empreinte digitale par minuties

Les détails des minuties constituent la représentation la plus populaire de toutes les représentations existantes. Elles répondent efficacement au problème de taille posé

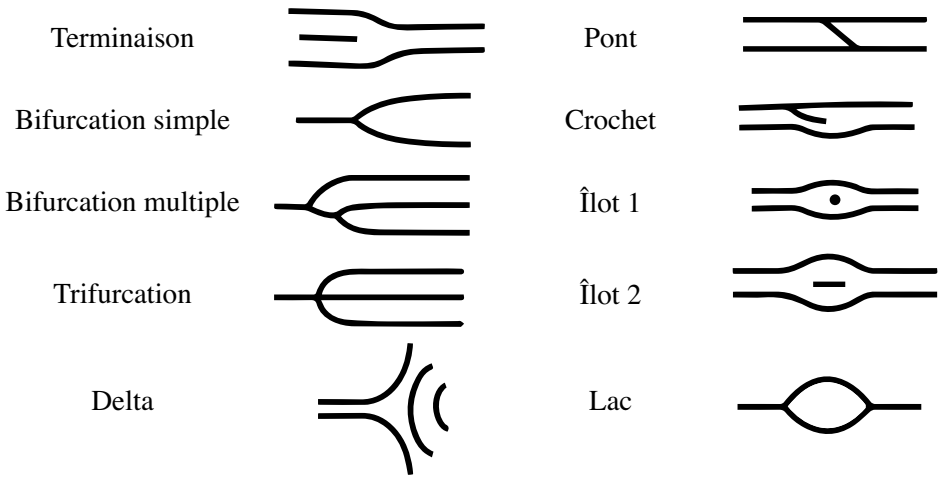


FIGURE 2.10 – Les différents types de minuties.

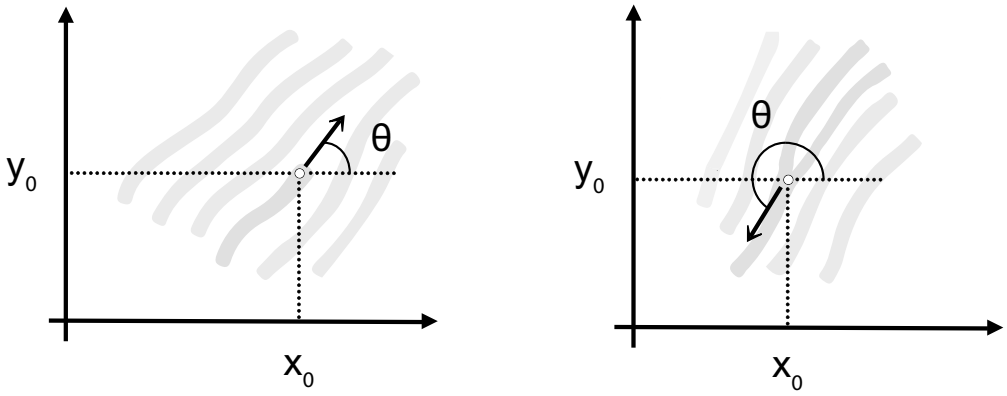


FIGURE 2.11 – Les caractéristiques des minuties de type : terminaison et bifurcation.

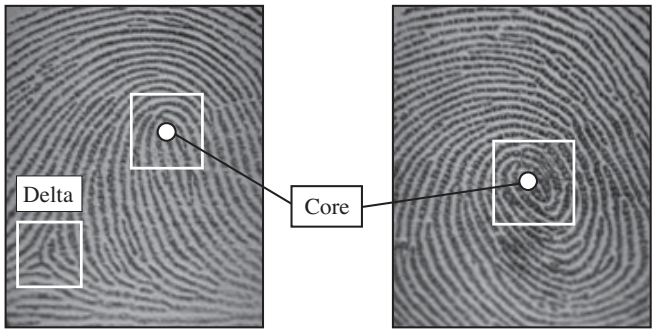


FIGURE 2.12 – Les singularités dans une empreinte digitale.

précédemment. Les minuties peuvent être appariées en considérant le problème comme un problème d'appariement des primitives points (*Point Pattern Matching*).

Dans le cas idéal, nous supposons plusieurs propriétés :

1. La correspondance entre les deux ensembles de minuties est totale.

2. Il n'y a aucune déformation telles que la translation ou la rotation.
3. Chaque minutie présentée sur une empreinte est exactement localisée.

Alors, la vérification par empreintes digitales devient une tâche insignifiante. Cependant, dans la pratique, la qualité de l'empreinte rencontrée durant la vérification est très incertaine, elle varie sur une grande portée :

1. Il peut ne pas y avoir un chevauchement suffisant entre les deux empreintes à comparer. Cela est particulièrement présent pour les capteurs de petite surface.
2. Il y a une translation relative, une rotation et des déformations non-linéaires des minuties dans l'image d'entrée par rapport au descripteur (modèle de référence).
3. Quelques minuties peuvent manquer dans l'image d'entrée et de fausses minuties peuvent apparaître.
4. Il peut y avoir des erreurs de localisation (position ou orientation) par l'extracteur, particulièrement sur les images de très mauvaise qualité.

Les performances des algorithmes d'extraction de minuties et d'autres techniques de reconnaissance des empreintes digitales dépendent fortement de la qualité des images d'empreintes digitales d'entrée. Dans une image d'empreinte digitale idéale, les crêtes et les vallées alternent et s'écoulent dans une direction localement constante. Dans de telles situations, les crêtes peuvent être facilement détectées et les points caractéristiques peuvent être localisés avec précision dans l'image. La figure 2.13(a) montre un exemple d'image d'empreinte digitale de bonne qualité. Cependant, dans la pratique, en raison de l'état de la peau (humide ou sèche, coupures et contusions), du bruit du capteur, de la pression incorrecte du doigt et de la mauvaise qualité inhérente des doigts (personnes âgées, travailleurs manuels), un pourcentage important des images d'empreintes digitales (environ 10%) est de mauvaise qualité, comme celles des figures 2.13(b) et 2.13(c).

Le processus de détection automatique des minuties est un processus extrêmement critique, particulièrement pour les empreintes de mauvaise qualité. Le processus traditionnel d'extraction des minuties suit les étapes suivantes tel qu'il est illustré sur la figure 2.14.

- **Amélioration** (*Image Enhancement*) : le but de cette étape est d'améliorer la qualité des régions récupérables dans l'image. Les procédures d'amélioration orientées pixel (comme la légalisation d'histogramme, la normalisation, le filtrage ou l'adoucissement des frontières) améliorent la lisibilité de l'empreinte mais ne sont pas suffisantes pour traiter ce type d'images car ils n'agissent pas sur la structure globale des crêtes. En effet, le bruit dans une image d'empreinte s'exprime par une cassure dans le flux directionnel des crêtes. Généralement, une image d'empreinte digitale contient les trois catégories de régions suivantes (figure 2.15) : *(i)* région bien définie (les crêtes et les vallées sont visibles pour une extraction possible), *(ii)* région récupérable (les crêtes et les vallées sont corrompues mais un algorithme

(a) (b) (c)

FIGURE 2.13 – a) Une empreinte digitale de bonne qualité ; b) une empreinte digitale de qualité moyenne caractérisée par des rayures et des cassures de crêtes ; c) une empreinte digitale de mauvaise qualité contenant beaucoup de bruit.

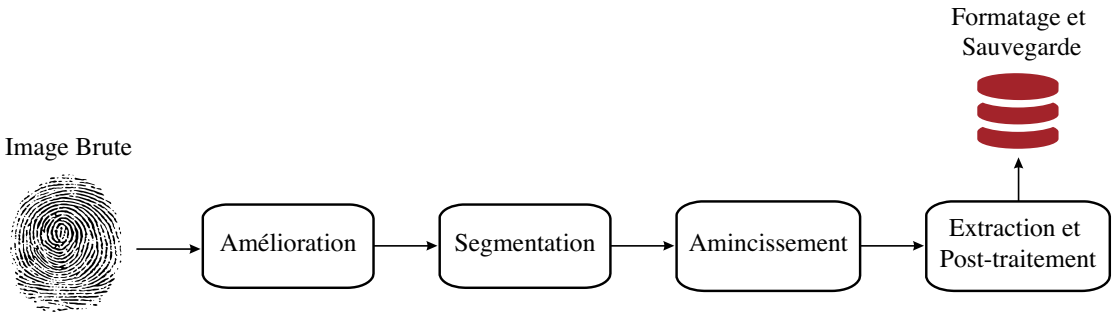


FIGURE 2.14 – Processus usuel d'extraction des minuties

d'amélioration peut les récupérer), *(iii)* région irrécupérable (les régions sont très touchées par le bruit). Un algorithme d'amélioration a comme but de récupérer la région d'intérêt et de l'améliorer et de masquer la région irrécupérable. Pour ce faire, les meilleurs résultats ont été obtenus en utilisant des filtres contextuels où les paramètres du filtre sont localement adaptés. Ces paramètres dépendent essentiellement de la fréquence et de la direction locales des crêtes. Un exemple de ces filtres sont les filtres de Gabor [Hong *et al.*, 1998]. Ils opèrent comme des filtres passe-bandes, en augmentant le contraste entre crêtes et vallées (filtrage de différentiation) dans la direction normale à l'orientation des crêtes tout en effectuant un adoucissement dans la direction des crêtes (filtrage passe-bas) pour combler les impuretés et lier les trous.

- **Segmentation** : l'image en niveaux de gris est convertie en image binaire pour distinguer les crêtes des vallées. A cause de son caractère non stationnaire, une binarisation adaptative est souvent préférée. Le seuil de binarisation est déterminé localement en considérant les propriétés du voisinage local [Jain *et al.*, 1997, Ratha *et al.*, 1995, Domeniconi *et al.*, 1998]. Généralement, cette étape de binari-

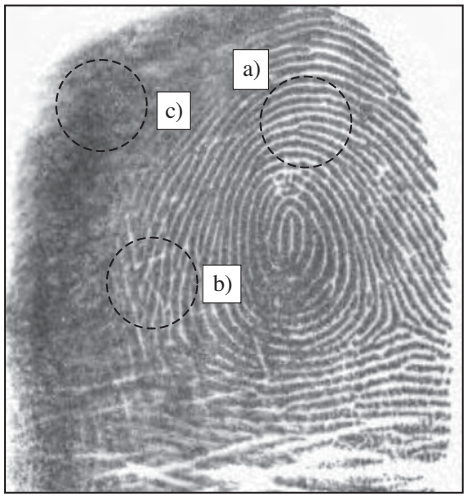


FIGURE 2.15 – Une image d'empreinte digitale contenant des régions de qualité différente : a) une région bien définie ; b) une région récupérable ; c) une région non récupérable.

sation fournit de bons résultats à condition qu'elle soit appliquée à des images de bonne qualité ou après une phase d'amélioration.

- **Amincissement** : l'image binaire, en utilisant des opérateurs morphologiques, est soumise à une étape d'amincissement (l'épaisseur des lignes de crêtes est réduite à un pixel). Quelques algorithmes comme le MINDTCT [Garris *et al.*, 2001] développé par le NIST (*National Institute of Standards end Technology*) pour le FBI (*Federal Bureau of Investigation*), ne requièrent pas cette étape.
- **Extraction et post-traitement** : Un simple calcul du nombre de connexions d'un pixel crête sur l'image amincie peut informer si le pixel concerné est une minutie ou bien non. Un post-traitement s'avère toujours utile pour éliminer les fausses alarmes. Celui-ci est généralement basé sur le choix d'heuristiques structurelles.
- **Formatage et sauvegarde** : les minuties requièrent une représentation très compacte qui dépasse rarement le $1KO$. Chaque minutie peut être décrite par un nombre d'attributs telles que la position (x, y), l'orientation et d'autres informations susceptibles d'aider à l'appariement comme son type. Cependant, *la plupart des algorithmes considèrent seulement sa position et son orientation.*

2.5.3 Reconnaissance par empreintes digitales

La comparaison fiable d'images d'empreintes digitales est un problème extrêmement difficile, principalement en raison de la grande variabilité des différentes empreintes d'un même doigt (c'est-à-dire les grandes variations intra-classes). Les principaux facteurs responsables des variations intra-classes sont : le déplacement, la rotation, le chevauchement partiel, la distorsion non linéaire, la pression variable, l'état changeant de la peau, le bruit et les erreurs d'extraction de caractéristiques. Par conséquent, les empreintes digi-

tales d'un même doigt peuvent parfois sembler très différentes, alors que les empreintes digitales de doigts différents peuvent sembler très similaires.

La comparaison automatique d'empreintes digitales ne suit pas nécessairement les mêmes directives. En fait, bien que la comparaison automatique d'empreintes digitales basée sur les points caractéristiques s'inspire de la procédure manuelle, un grand nombre d'approches ont été conçues au cours des 40 dernières années, et bon nombre d'entre elles ont été explicitement conçues pour être mises en œuvre sur un système informatique. Les approches de la comparaison d'empreintes digitales sont classées en trois catégories :

— Correspondance basée sur la corrélation : deux images d'empreintes digitales sont superposées et la corrélation (au niveau de l'intensité) entre les pixels correspondants est calculée pour différents alignements (par exemple, divers déplacements et rotations) ;

— Correspondance basée sur minuties : les minuties sont extraits des deux empreintes digitales et stockés comme des ensembles de points dans le plan bi-dimensionnel. La correspondance des minuties consiste essentiellement à trouver l'alignement entre le modèle et les ensembles de minuties d'entrée qui donne le nombre maximal de paires de minuties ;

— Correspondance basée sur les caractéristiques des crêtes : l'extraction des minuties est difficile dans les images d'empreintes digitales de très faible qualité, alors que d'autres caractéristiques du motif des crêtes des empreintes digitales (par exemple, l'orientation et la fréquence locales, la forme des crêtes, les informations sur la texture) peuvent être extraites de manière plus fiable que les minuties, même si leur caractère distinctif est généralement plus faible. Les approches appartenant à cette famille comparent les empreintes digitales en fonction des caractéristiques extraites du motif des crêtes.

Compte tenu d'un environnement opérationnel complexe, il est essentiel d'identifier un ensemble d'hypothèses valables sur lesquelles la conception de l'appareil de comparaison d'empreintes digitales pourrait être basée. Il faut souvent choisir s'il est plus efficace d'exercer plus de contraintes en incorporant une meilleure conception technique ou de construire une fonction de similarité plus sophistiquée pour la représentation donnée. Par exemple, dans un dispositif de mise en correspondance d'empreintes digitales, il est possible de limiter complètement la distorsion élastique et de concevoir le dispositif de mise en correspondance sur la base d'une hypothèse de transformation rigide ou d'autoriser des distorsions arbitraires et de s'adapter aux variations des images d'entrée en utilisant un dispositif de mise en correspondance intelligent. À la lumière des environnements opérationnels mentionnés ci-dessus, la conception de l'algorithme d'appariement doit établir et caractériser un modèle réaliste des variations entre les représentations des paires appariées.

2.5.4 Défis de la reconnaissance par empreintes digitales

Dans la pratique, les variations intra-classe rendent la reconnaissance d'empreintes digitales une tâche extrêmement difficile, parce que plusieurs acquisitions du même doigt sont très peu probables d'aboutir à un ensemble identique de minuties. Les principaux facteurs responsables des variations intra-classe sont dus à la distorsion non-linéaire (en raison de l'élasticité de la peau), la translation et la rotation d'une impression (Figure 2.16). Un doigt peut être placé et/ou mis en rotation sur le capteur différemment pendant plusieurs authentifications (selon [Maltoni *et al.*, 2009], le déplacement de 2 mm correspond à 40 pixels de translation et $\pm 20°$ de rotation peut être observé). Par conséquent, la reconnaissance des empreintes digitales est très sensible à l'orientation et à la translation des impressions, ce qui rend la protection des modèles de cette modalité encore plus compliquée.

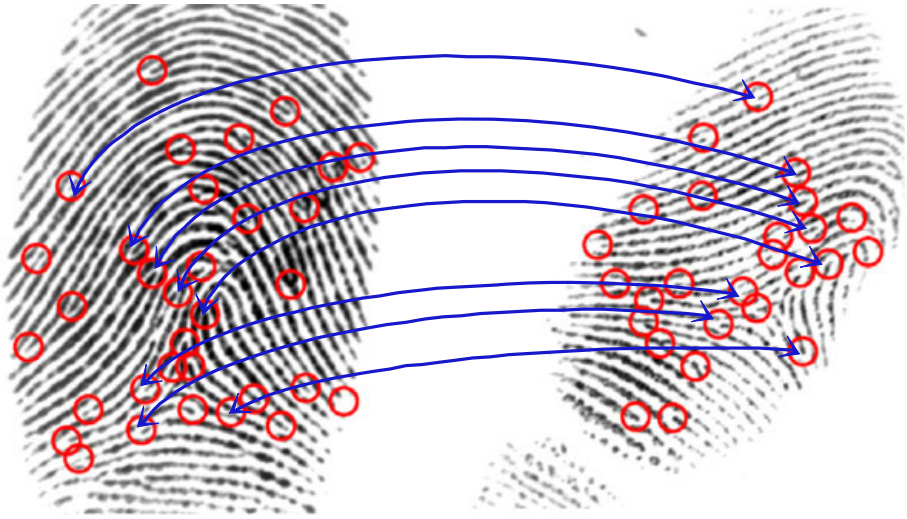


FIGURE 2.16 – Exemples de translation et rotation de deux impressions d'un même doigt.

2.5.5 Attaques contre les modèles d'empreinte digitale

Les données biométriques sont généralement stockées sous la forme de modèles biométriques constitués de caractéristiques saillantes et pouvant être efficacement appariées, extraites du signal ou de l'image biométrique capturée lors de l'enrôlement de l'utilisateur. Si l'image biométrique originale (par exemple, une empreinte digitale ou une image de visage) à partir de laquelle le modèle est dérivé, peut être reconstruite à partir d'un modèle volé, une réplique physique de la biométrie, appelée biométrie d'usurpation, peut être construite, compromettant ainsi la sécurité du système. La récupération de l'image biométrique est également une préoccupation en raison de la possibilité de dériver des informations personnelles sensibles à partir d'une image biométrique. Ainsi, du point de vue de la sécurité du système, le concepteur du système doit s'assurer qu'il est extrêmement difficile de reconstruire l'image biométrique à partir de son modèle stocké. En

revanche, du point de vue de la précision de la correspondance, le modèle doit contenir autant d'informations individuelles que l'image biométrique associée. En d'autres termes, la perte des informations discriminantes doit être très faible lors de l'extraction du modèle d'une image biométrique.

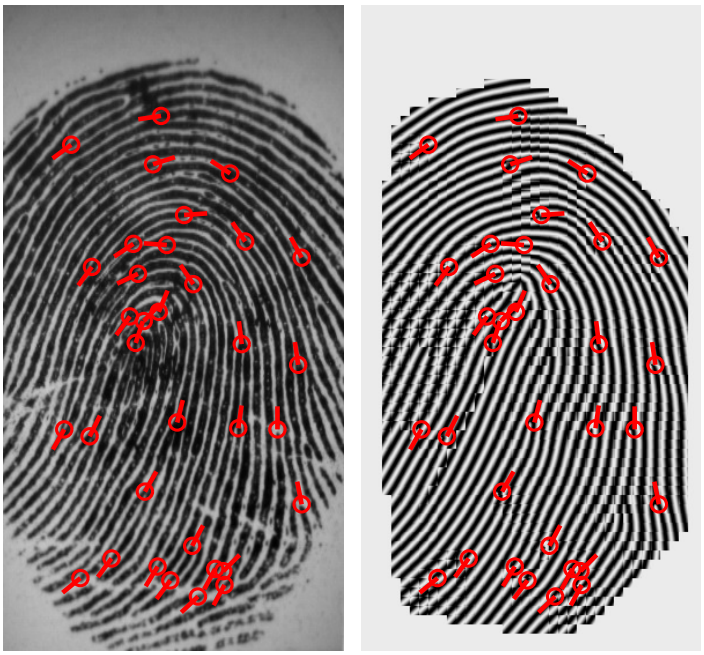
Reference	Biometric trait	Input representation	Recovered output	Technique
[Pötzsch *et al.*, 1996]	Face	Elastic Bunch Graph	Face image	Template inversion
[Hill, 2001]	Fingerprint	Minutiae	Fingerprint image	Template inversion
[Ross *et al.*, 2007]	Fingerprint	Minutiae	Fingerprint image	Template inversion
[Chang *et al.*, 2006]	Fingerprint	Minutiae	Fingerprint image	Template inversion
[Testoni et Kirovski, 2010]	Iris	Iriscode	Iris	Template inversion
[Feng et Jain, 2010]	Fingerprint	Minutiae	Fingerprint image	Template inversion
[Adler, 2003b]	Face	Matching System	Face image	Hill climbing
[Uludag et Jain, 2004a]	Fingerprint	Matching System	Minutiae	Hill climbing
[Yamazaki *et al.*, 2005]	Signature	Matching System	Time series data	Hill climbing
[Mohanty *et al.*, 2007]	Face	Matching System	Face image	Hill climbing
[Muramatsu, 2008]	Signature	Matching System	Time series data	Hill climbing
[Galbally *et al.*, 2010]	Face	Matching System	Face image	Hill climbing
[Martinez-Diaz *et al.*, 2011]	Fingerprint	Matching System	Minutiae	Hill climbing

TABLE 2.1 – Techniques existantes pour récupérer des données biométriques, à partir d'un modèle stocké et d'un système de correspondance.

Ces deux exigences sont concurrentes et il est important de trouver une manière de concevoir un modèle qui réponde le mieux possible à ces deux attentes. Un certain nombre de techniques ont été proposées pour permettre à un imposteur de récupérer l'image biométrique à partir de son modèle stocké (Tableau 2.1). Ces techniques peuvent être classées en deux grandes catégories : i) *Template inversion* et ii) *Hill climbing*. Dans la technique *Template inversion*, les caractéristiques de l'image biométrique sont identifiées à partir d'un modèle volé. Ces caractéristiques sont ensuite utilisées pour reconstruire l'image biométrique. Dans la méthode *Hill climbing* [Soutar, 1999], l'adversaire commence avec une estimation initiale de l'image biométrique qui est affinée itérativement sur la base du score obtenu en faisant correspondre l'image biométrique estimée avec le modèle stocké. À noter que les techniques *Hill climbing* ne nécessitent pas nécessairement l'accès au modèle stocké, mais uniquement aux scores de correspondance obtenus lorsqu'une image biométrique reconstruite est mise en correspondance avec le modèle stocké (Annexe A.1). Malgré sa nature générique, l'inconvénient de la technique *Hill climbing* est qu'il s'agit d'une procédure itérative dont le nombre des itérations dépend fortement des caractéristiques de l'algorithme de comparaison. De plus, il n'y a aucune garantie que l'image biométrique récupérée dans un système correspondra bien à une autre instance de la même biométrie. Dans le cas des empreintes digitales, par exemple, une approche de type *Hill Climbing* peut générer de nombreuses minuties parasites en dehors du domaine de l'ensemble de minuties original ou dans la région périphérique de l'image d'empreinte digitale. Un tel modèle reconstruit peut ne pas conduire à un score de correspondance élevé avec une

autre empreinte du même doigt. Les techniques de type *Template inversion*, en revanche, ne nécessitent pas l'utilisation du comparateur.

Les empreintes digitales sont les modalités biométriques les plus utilisées. Traditionnellement, il était entendu qu'il n'était pas possible de récupérer une image d'empreinte digitale à partir de son ensemble de minuties, c'est-à-dire que les minuties étaient considérées comme non inversibles.



(a) (b)

FIGURE 2.17 – Reconstruction d'une empreinte digitale à partir des minuties. (a) Une empreinte digitale avec des minuties marquées, et (b) l'empreinte digitale reconstruite à partir de l'ensemble de minuties (modèle) de (a), en utilisant la technique proposée dans [Feng et Jain, 2010].

Hill [Hill, 2001] a proposé la première technique de type *Template inversion* pour les empreintes digitales. Dans cette technique, les configurations possibles des points Core ont été sélectionnées et le champ d'orientation de l'empreinte digitale a été construit pour chaque configuration. Le champ d'orientation qui correspond le mieux aux minuties donnés est sélectionné et, à partir de chaque minutie, des lignes de crête sont tracées le long du champ d'orientation estimé pour obtenir une image d'empreinte digitale. Par la suite, un certain nombre d'autres approches efficaces ont été proposées pour reconstruire des images d'empreinte digitale à partir de minuties qui correspondent aux empreintes digitales originales avec une grande précision. Dans [Ross *et al.*, 2007], le champ d'orientation de l'empreinte digitale a été reconstruit sur la base de la direction des minuties voisines et les lignes de crête ont été simulées d'une manière similaire à [Hill, 2001]. Dans [Chang *et al.*, 2006], une technique plus sophistiquée, proposée dans

[Vizcaya et Gerhardt, 1996], a été utilisée pour reconstruire le champ d'orientation à partir des minuties. L'estimation du champ d'orientation est suivie par une étape de filtrage qui utilise des filtres de Gabor locaux orientés selon l'orientation de la crête pour générer le motif de l'empreinte digitale. Bien que cette approche ait produit des empreintes digitales très réalistes par rapport aux approches précédentes, elle a également donné lieu à un grand nombre de minuties erronées. Feng et Jain [Feng et Jain, 2010] ont proposé une procédure de reconstruction d'images d'empreintes digitales en adaptant un modèle AM-FM aux empreintes digitales. Cette approche permet non seulement de générer des empreintes digitales très réalistes, mais aussi de réduire le nombre de fausses minuties. La figure 2.17(b) montre l'image reconstruite de l'image d'empreinte digitale originale de la figure 2.17(a).

2.6 Bilan du chapitre

Nous avons présenté dans ce chapitre les principales notions en biométrie, et en particulier en biométrie des empreintes digitales, nécessaires pour la suite du travail. Après une introduction générale sur la biométrie, ses propriétés, ses technologies et son usage, nous avons présenté l'architecture d'un système biométrique générique, ses limitations et ses défaillances en soulignant l'importance de protéger les modèles biométriques. Nous avons ensuite listé les niveaux d'attaques qui peuvent être menées contre un système biométrique, et par la suite, nous avons introduit les généralités sur l'empreinte digitale puisqu'elle représente l'objet de cette thèse, en précisant d'un coté les problèmes critiques liés à la reconnaissance par empreintes digitales, à savoir les principaux facteurs responsables des variations intra-classe, et d'un autre coté, les attaques de reconstruction à partir d'un modèle d'empreinte digitale volé. Nous présentons dans le chapitre suivant les différentes attaques contre les modèles des empreintes digitales, et ensuite nous abordons un état de l'art des travaux sur la protection des systèmes de sécurité biométriques en mettant l'accent sur les méthodes de la protection des modèles des empreintes digitales.

Chapitre

3 État de l'art de la protection des modèles biométriques

Sommaire

3.1 Introduction

Ce chapitre présente dans un premier lieu les exigences d'un schéma idéal de protection des modèles biométriques, puis fait un tour d'horizon sur les travaux relatifs à la protection générale des systèmes de sécurité biométriques et particulièrement ceux qui sont liée à la protection des modèles des empreintes digitales. Le but principal de ce chapitre n'est pas d'aborder une étude exhaustive de toutes les publications existantes pour la protection des modèles biométriques, mais plutôt notre objectif est de présenter une vue générale sur ces approches avec une attention particulière aux techniques de transformation révocable. Dans ce chapitre, la section 3.2 est consacrée à la présentation des exigences qui doivent être respecté pour tout schéma de protection des modèles biométriques. Dans la section 3.3, nous abordons les solutions existantes pour sécuriser les systèmes biométriques en mettant l'accent sur les approches de transformation révocable. La section 3.4 présentent les méthodes qui relèvent la question de la protection des gabarits d'empreintes digitales. Enfin, le bilan du chapitre est présenté dans la section 3.5.

3.2 Exigences des approches de protection des modèles biométriques

L'accès non autorisé aux modèles biométriques est l'une des menaces les plus dangereuses pour la sécurité et la vie privée des utilisateurs. En fait, bien qu'il soit communément admis qu'il n'est pas possible de reconstruire les caractéristiques biométriques originales à partir du modèle extrait correspondant, certains contre-exemples concrets, qui contredisent cette hypothèse, ont été présentés dans la littérature. Prenons à titre exemple le travail proposé dans [Adler, 2003a] où il est démontré que la connaissance du modèle biométrique du visage et du score de correspondance peut conduire à la reconstruction du visage, ou alors dans [Ross *et al.*, 2007] où un algorithme efficace a été proposé pour générer une empreinte digitale à partir des caractéristiques du modèle de référence correspondant. Le fait donc de conserver les modèles biométriques sans protection ne serait pas assez sûr et dans le cas où le modèle est compromis, il est hautement souhaitable de le révoquer ou de le remplacer, et aussi d'obtenir à partir de la même biométrie différentes références pour accéder à différents emplacements, physiques ou logiques, afin d'éviter un suivi non autorisé. Ces défis ont incité les développeurs et les chercheurs à étudier et proposer différentes techniques, approches et schémas pour assurer un stockage sécurisé des modèles biométriques.

En résumé, quatre propriétés sont essentielles pour tout système de protection des modèles biométriques :

- *Révocabilité* : Il devrait être possible de révoquer un modèle de référence compromis et d'en émettre un nouveau, différent du précédent et basé sur les mêmes données biométriques originales. En outre, le nouveau modèle émis ne doit pas correspondre au modèle précédemment compromis. Il faut noter que la biométrie seule ne peut pas fournir cette propriété car les caractéristiques biométriques ne peuvent pas être

modifiées alors que les systèmes utilisant les mots de passe et les jetons ont une excellente révocabilité.

- *Diversité* : Le modèle protégé ne doit en aucun cas permettre une correspondance croisée dans la base de données. Cette propriété reflète la possibilité de produire un très grand nombre de modèles protégés à partir des mêmes données biométriques originales de telle sorte qu'ils soient suffisamment différents les uns des autres. Cela permettra d'éviter la poursuite et la surveillance des utilisateurs à travers différentes bases de données.
- *Sécurité* : Il devrait être impossible ou difficile sur le plan informatique d'obtenir le modèle biométrique original à partir du modèle stocké et protégé. Cette propriété est nécessaire pour empêcher un adversaire de fabriquer une contrefaçon physique de la caractéristique biométrique à partir des modèles volés.
- *Performance* : Les taux d'erreur de reconnaissance biométrique en termes de taux de faux rejet ou de taux de fausse acceptation ne doivent pas se dégrader de manière significative avec l'introduction d'un système de protection des modèles, par rapport à une approche non protégée.

La conception d'un schéma de protection de modèles capable de satisfaire correctement chacune des propriétés susmentionnées est un véritable défi, principalement en raison de la variabilité intra-classe inévitable que présente chaque trait biométrique. Ces dernières années, de nombreuses solutions ont déjà été proposées pour la génération de gabarits sûrs et révocables. Dans ce chapitre, nous présentons une variété de classifications possibles pour ces algorithmes de protection.

3.3 Sécurité des systèmes biométriques

Les techniques de protection des modèles biométriques peuvent être classées de manière générale comme deux types d'approches : approches matérielles et approches logicielles. Il s'agit de deux types de solutions, l'une au niveau architectural et concerne principalement les systèmes fermés et l'autre au niveau algorithmique ou fonctionnel, et qui peuvent être utilisées séparément ou conjointement (Figure 3.1).

3.3.1 Les systèmes fermés

la sécurité peut être renforcée en déplaçant autant de modules que possible sur du matériel sécurisé (inviolable) auquel un pirate ne peut accéder, même s'il a un accès physique ou à distance. Il existe quelques approches populaires qui ont conduit à des solutions commercialement viables :

- Déplacer uniquement le module de stockage (qui contient les modèles d'enrôlement) et le module de correspondance sur une carte à puce qui peut être en possession de l'utilisateur final. Cette technique est connue sous le nom de *Match-on-Card* (MoC).

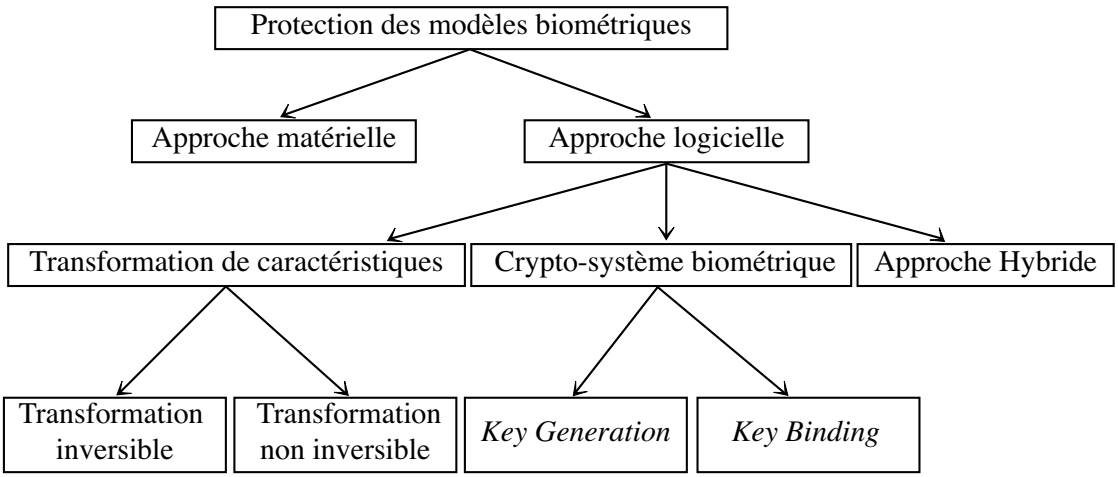


FIGURE 3.1 – Catégorisation des approches de protection des modèles biométriques

- Déplacer tous les modules (y compris l'extraction de caractéristiques ainsi que le capteur d'acquisition) sur les plateformes matérielles sécurisées. En fonction de la plate-forme matérielle, différentes solutions ont été proposées, à savoir :
  - *System-on-Device* (SoD) : désigne les solutions dans lesquelles les modules d'extraction et de comparaison des caractéristiques sont intégrés à la carte matérielle du scanner, ces solutions sont également disponibles dans le commerce sous les noms de *Match-on-Device* (MoD) ou *Match-on-Board* (MoB).
  - *System-on-Card* ou *System-on-a-Chip* (SoC) : désigne les solutions dans lesquelles tous les modules sont mis en œuvre sur une plate-forme compacte, telle qu'une carte à puce spéciale ou une puce sécurisée. Dans ce cas, la plate-forme est généralement mise à la disposition des utilisateurs.

Dans une plate-forme matérielle sécurisée, le traitement s'effectue dans un environnement sécurisé, isolé du système d'exploitation du système client (c'est-à-dire le système d'exploitation hôte). Outre l'avantage de la sécurité (résistance aux attaques par déni de service et aux intrusions [Bolle *et al.*, 2002, Cooper *et al.*, 2007]), les solutions MoC et SoC présentent également des avantages en matière de confidentialité. En effet, la plate-forme est généralement mise à la disposition des utilisateurs qui ont le plein contrôle de leurs propres données biométriques.

3.3.1.1 Match-on-card

L'une des plateformes matérielles sécurisées les plus populaires pour la vérification biométrique est la carte à puce, également connue sous le nom de carte à circuit intégré. Une carte à puce est généralement de la taille d'une carte de crédit ou plus petite et contient un processeur inviolable qui peut généralement exécuter des fonctions cryptographiques. Bien que le stockage des cartes à puce soit très limité, les applications n'envisagent pas le partage des cartes à puce et il n'est donc pas nécessaire d'y stocker de nombreux modèles.

Le dispositif de comparaison sur carte effectue une comparaison entre le(s) modèle(s) stocké(s) sur la carte et le modèle produit qui lui est envoyé par l'ordinateur hôte.

L'avantage de cette approche est que le module de comparaison et le stockage des modèles sont entièrement sécurisés comme le montre la figure 3.2. Les modèles ne peuvent ni être modifiés par des pirates malveillants ni être espionnés. Une fois que le modèle est transféré à la carte à puce, il n'est plus nécessaire de communiquer les modèles à l'hôte. Seul le résultat du processus de comparaison qui doit être communiqué à l'hôte. Enfin, les clés cryptographiques sont également stockées sur la carte à puce. La gestion des clés est donc simplifiée et sécurisée, ce qui renforce considérablement la sécurité du système. Les solutions *Match-on-Card* sont parfois considérées comme plus sûres que les solutions *System-on-Device* en raison de la plus grande isolation du modèle. Toutefois, il convient de noter que même si le modèle enregistré ne peut pas être obtenu par un pirate, un modèle "suffisamment similaire" peut être obtenu en écoutant le module d'extraction de caractéristiques fonctionnant sur l'hôte non sécurisé du système MoC. Le risque d'une telle attaque ne doit pas être sous-estimé car il n'est pas difficile de mettre l'hôte sur écoute.

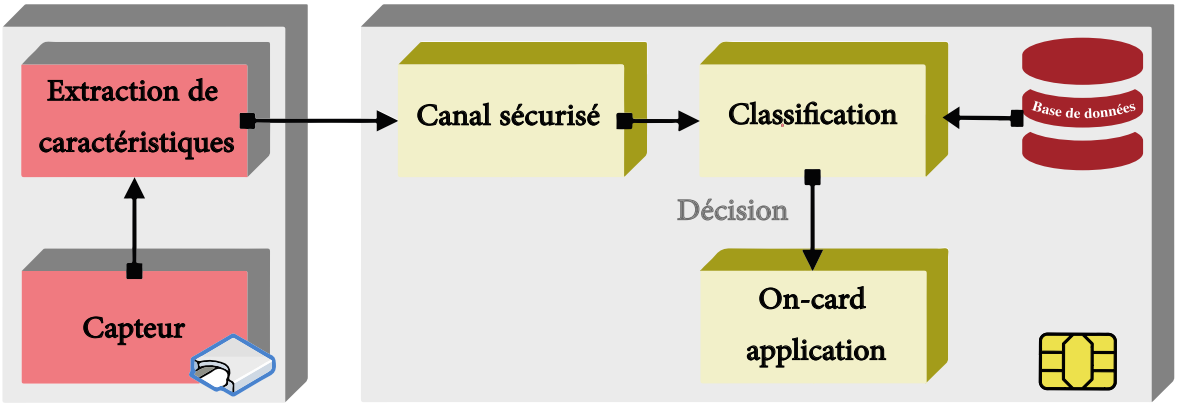


FIGURE 3.2 – Architecture de l'approche *Match-on-Card*

3.3.1.2 System-on-device / system-on-a-chip

Dans les systèmes MoC, même si le modèle est protégé, le scanner/capteur est relié à un système hôte, dont la sécurité peut être faible et indigne de confiance. L'extraction des caractéristiques est également effectuée sur l'hôte et reste donc vulnérable aux intrusions et aux attaques par déni de service. Ces vulnérabilités peuvent être corrigées en déplaçant les modules restants, c'est-à-dire l'extracteur de caractéristiques et peut-être même le scanner/capteur, vers une plate-forme matérielle sécurisée (Figure 3.3). Lorsque la plate-forme matérielle cible est le scanner/capteur, la solution est appelée *System-on-Device* (SoD) ; lorsque la cible est une carte à puce spéciale ou une puce sécurisée, l'architecture est appelée *System-on-Card* ou *System-on-a-Chip* (SoC). Par conséquent, aucune donnée biométrique brute ne transite en dehors de l'espace sécurisé.

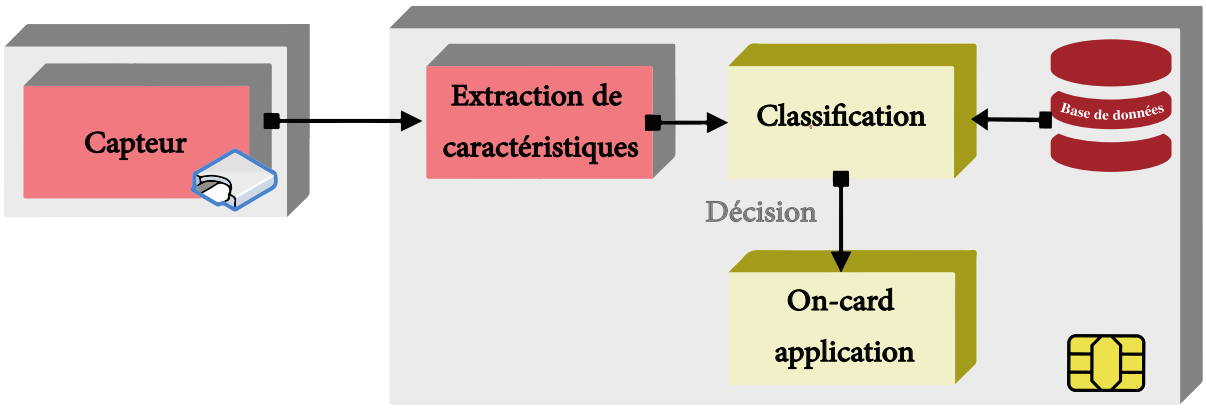


FIGURE 3.3 – Architecture de l'approche *System-on-a-chip*

3.3.2 Transformations de caractéristiques

Cette approche consiste à appliquer une fonction de transformation $f(.)$ au modèle biométrique x^{Enr} et à ne stocker dans la base de données que le modèle transformé $f(x^{Enr}, K)$. Les paramètres de la fonction de transformation sont généralement dérivés d'une clé aléatoire, K , ou d'un mot de passe. La même fonction de transformation est appliquée aux caractéristiques de la requête, x^{Aut}, et la requête transformée, $f(x^{Aut}, K)$, est directement comparée au modèle transformé, $f(x^{Enr}, K)$. La figure 3.4 montre clairement que l'approche de transformation des caractéristiques est analogue au cryptage ou au hachage des mots de passe.

Les fonctions de transformation sont une solution particulièrement intéressante pour compenser la variabilité des données biométriques. Le point commun à ces techniques est d'exécuter la comparaison entre la donnée capturée et la donnée stockée directement dans le domaine de transformation. La révocabilité est garantie car, lorsqu'une donnée transformée est compromise, il suffit de changer la fonction de transformation. La diversité est également assurée par le choix de fonctions différentes pour des applications distinctes. Cependant, trouver de telles fonctions n'est pas simple. En effet, outre la non-inversibilité, ces fonctions doivent exhiber deux propriétés essentielles : une robustesse intra-classe (c'est-à-dire une robustesse vis-à-vis les variations d'une donnée biométrique d'un individu) et une sensibilité inter-classe (le fait de pouvoir distinguer deux individus différents). En fonction des caractéristiques de la fonction de transformation $f(.)$, les schémas de transformation de caractéristiques peuvent être classés en deux catégories : *les transformations inversibles* (connue sous la dénomination anglaise *Salting* ou aussi *BioHashing*) et *les transformations non inversibles*.

3.3.2.1 Transformations inversibles

Lorsque la fonction de transformation $f(.)$ est inversible, la sécurité du modèle transformé se base principalement sur le secret de la clé. En d'autres termes, si un adversaire a accès à la clé et au modèle transformé, il peut récupérer le modèle biométrique origi-

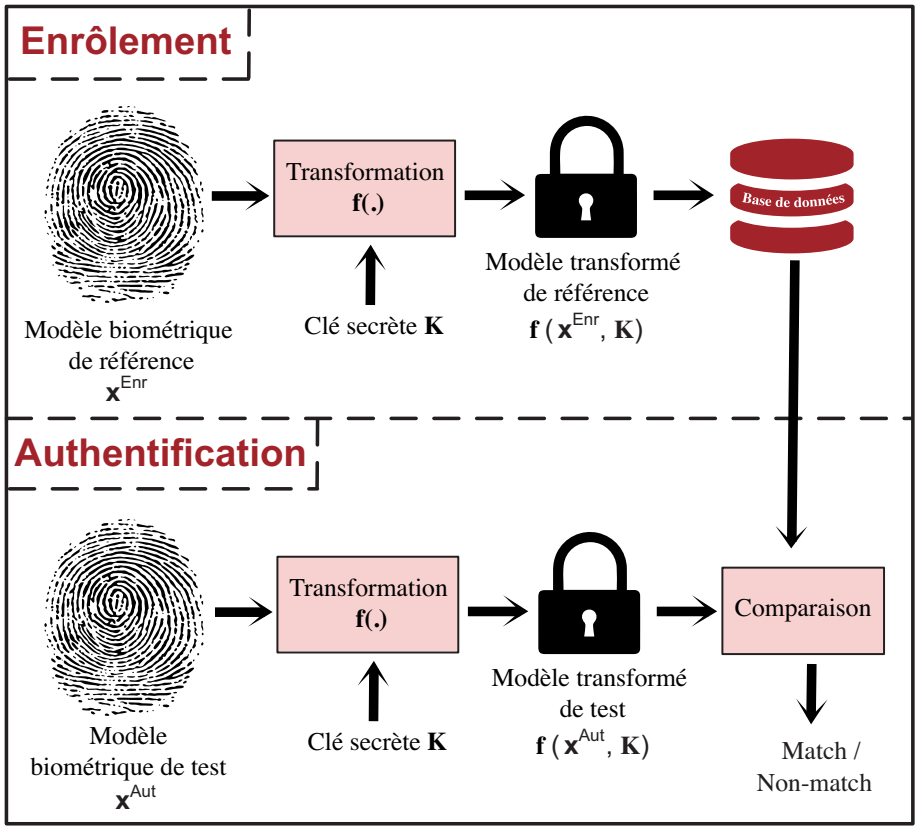


FIGURE 3.4 – Protection des modèles biométriques par la transformation de caractéristiques.

nal ou une approximation proche de celui-ci. Ainsi, un modèle protégé par l'approche de la transformation de caractéristiques inversibles est similaire à un mot de passe crypté. Les difficultés pratiques de gestion des clés limitent la sécurité de ce genre de transformations. De plus, les algorithmes de mise en correspondance doivent être repensés pour permettre la mise en correspondance dans le domaine transformé. Cependant, si les clés sont faites de manière à être spécifiques à l'utilisateur et qui sont censées être des secrets connus uniquement par les utilisateurs légitimes, il existe deux avantages potentiels. Premièrement, l'utilisation des informations aléatoires supplémentaires sous forme d'une clé spécifique augmente généralement la séparabilité entre les utilisateurs dans l'espace des caractéristiques. La capacité de discrimination dans le domaine transformé devient donc plus importante que dans le domaine d'origine, ce qui entraîne un taux de fausse acceptation plus faible. Deuxièmement, les clés spécifiques à l'utilisateur facilitent la révocabilité des modèles transformés.

Le régime *BioHashing* [Jin *et al.*, 2004, Teoh *et al.*, 2006] est par nature une approche à deux facteurs dans laquelle le modèle biométrique est transformé à l'aide d'une fonction dont les paramètres sont définis par une clé externe. C'est une transformation, applicable à différentes modalités biométriques, qui génère un vecteur binaire appelé *BioCode*. Le

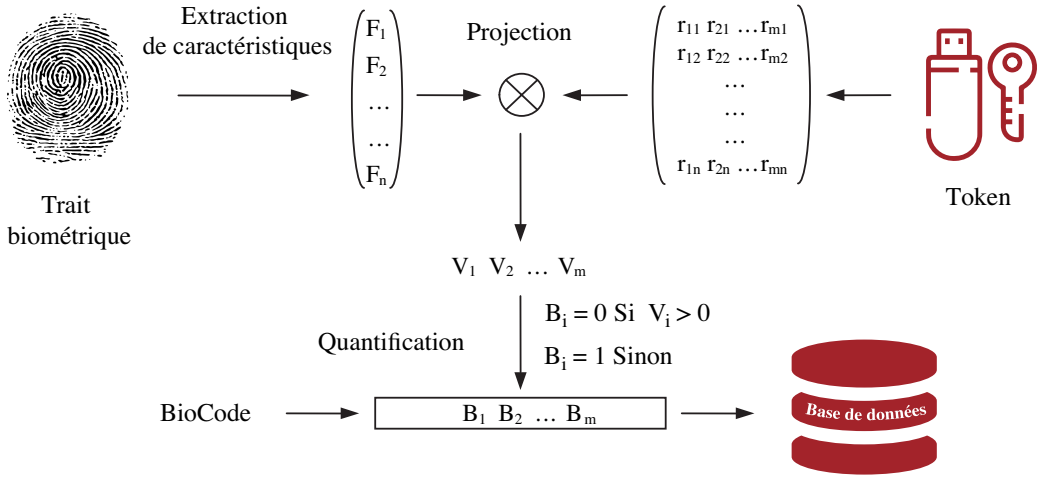


FIGURE 3.5 – Protection des modèles biométriques en utilisant le régime *BioHashing*.

principe du *BioHashing* (Figure 3.5) consiste à projeter la donnée biométrique (normalisée) sur une base orthonormée (matrices orthogonale aléatoire) générée à partir de la clé (la projection aléatoire [Goel *et al.*, 2005, Rathgeb et Uhl, 2011]). L'utilisation d'une base orthonormée permet de garantir la conservation des relations de similarité entre deux données biométriques projetées (comme cela a été démontré dans [Dasgupta et Gupta, 1999]). L'étape suivante consiste à quantifier ce résultat à l'aide d'un simple seuillage. Cette étape permet de garantir la non-inversibilité du procédé (obtention de la donnée biométrique initiale à partir du *BioCode*) et de rendre robuste le procédé (en autorisant des différences mineures dans le vecteur projeté inhérant à l'acquisition de la donnée biométrique). Les étapes du régime *BioHashing* peuvent être résumé comme suit :

1. Générer n vecteurs aléatoires en utilisant une clé utilisateur.
2. Appliquer l'algorithme de *Gram-Schmidt* (Annexe A.2) sur les n vecteurs aléatoires pour calculer une matrice orthogonale Δ.
3. Transformer le modèle d'origine z en utilisant la matrice Δ comme suit :

$$y = \Delta z \tag{3.1}$$

   y est le modèle transformé.
4. Quantifier le modèle transformé comme suit :

$$b_i = \begin{cases} 0 & \text{si} \quad y_i \leq \tau \\ 1 & \text{si} \quad y_i > \tau \end{cases} \quad \text{avec } i = 1, ..., n \tag{3.2}$$

   τ est le seuil de comparaison.
   t est le modèle protégé de taille n.

Il faut noter que plusieurs travaux [Jin *et al.*, 2004, Wang et Plataniotis, 2007] ont proposé d'appliquer l'étape de quantification (equation 3.2) sur le modèle transformé y

pour rendre la projection non inversible. Cependant, il est toujours possible de récupérer le modèle transformé (ou une approximation de celui-ci) à partir du modèle protégé comme proposé dans [Nagar *et al.*, 2010] :

- Premièrement, nous résolvons le problème des moindres carrés suivant :

$$argmin \parallel z - r \parallel_2 \ \ subject\ to \begin{cases} y_i \leq \tau \text{ si } t_i = 0 \\ y_i > \tau \text{ si } t_i = 1 \end{cases} \tag{3.3}$$

  — $i = 1, ..., n$ et $y_i = \sum_{j=1}^{n} \Delta_{ij} z_i$

  — r est un vecteur aléatoire de taille n avec $r \times z \leq \tau$.

  — n est la taille du modèle original et du modèle transformé.

- Deuxièmement, afin d'optimiser l'approximation de z, nous résolvons le problème k fois en utilisant k valeurs différentes de r. L'approximation finale $\tilde{z}$ est calculée comme suit :

$$\tilde{z} = \frac{\sum_{i=1}^{k} \frac{z_i}{d_i^2}}{\sum_{i=1}^{k} \frac{1}{d_i^2}} \tag{3.4}$$

  — z_i est le vecteur original estimé en utilisant r_i.

  — $d_i = \frac{1}{d_i'}$ avec d_i' est la distance de *Hamming* entre z_i et t.

Dans le *BioHashing*, les modèles protégés ne sont pas réversibles, sauf si le modèle et la clé sont connus simultanément. Cependant, si la clé est connue (ou si elle est suffisamment faible pour être craquée par une simple attaque par dictionnaire), le caractère "aléatoire" s'estompe [Kong *et al.*, 2006] et la quantification ne préserve pas suffisamment le modèle protégé. Même si une petite quantité d'information est perdue à cause de la quantification, une assez bonne approximation du modèle peut être récupérée [Jain *et al.*, 2008]. Pour améliorer la sécurité des approches *BioHashing*, il est recommandé que la clé ne soit pas stockée mais plutôt mémorisée par l'utilisateur, mais cela réintroduit la faiblesse des systèmes basés sur des mots de passe que nous essayons de contourner. Une propriété intéressante du *BioHashing* est que la clé externe fournit non seulement la propriété de non-inversibilité, mais améliore également la précision de la correspondance, ce qui n'est pas surprenant étant donné la nature à deux facteurs de l'approche. En changeant la clé, la diversité (et donc la révocabilité) peut être facilement atteinte.

Le régime *BioHashing* a été utilisée sur plusieurs modalités biométriques [Campisi, 2013] (essentiellement les empreintes digitales [Jin *et al.*, 2004, Sakata *et al.*, 2006], le visage [Savvides *et al.*, 2004, Teoh *et al.*, 2006, Lumini et Nanni, 2007], l'iris [Chin *et al.*, 2006], la texture de la paume de la main [Connie *et al.*, 2004, Connie *et al.*, 2005], etc.). L'objet de ces travaux est essentiellement d'augmenter la taille du *BioCode* (plus il est grand, moins une attaque par force brute sera possible) et d'améliorer les performances. La problématique de protection de données

biométriques a été souvent abordée de façon étonnante par le biais de la performance (minimisation du taux d'erreur et maximisation de la taille du *BioCode*).

3.3.2.2 *Transformations non inversibles*

Les schémas de transformation non inversibles appliquent généralement une fonction à sens unique sur le modèle biométrique [Rathgeb et Uhl, 2011], et il est difficile sur le plan informatique d'inverser le modèle transformé, même si la clé et/ou le modèle transformé sont connus. Généralement, on peut dire que la transformation non-inversible est équivalente à un schéma de hachage de mot de passe. Comme il est difficile de récupérer le modèle biométrique d'origine même lorsque les paramètres de la transformation sont compromis, ce schéma offre une meilleure sécurité par rapport à l'approche de transformation inversible. En littérature, le principe de transformation non inversible (bien que le principe de la biométrie révocable) a été proposée pour la première fois par [Ratha *et al.*, 2001b] (il a été nommé *Cancelable biometrics* dans [Bolle *et al.*, 2002]).

Les données biométriques révocables sont transformées avec un schéma de distorsion qui varie pour chaque application. Le concept a été développé dans [Ratha *et al.*, 2001b] (et clarifié dans [Ratha *et al.*, 2006, Ross *et al.*, 2005]), pour répondre aux préoccupations de confidentialité et de sécurité, à savoir que les données biométriques ne sont pas secrètes et ne peuvent pas être révoquées. Pendant l'enrôlement, l'image biométrique d'entrée est soumise à une distorsion connue (Figure 3.6) contrôlée par un ensemble de paramètres de distorsion. L'échantillon biométrique déformé peut ensuite être traité par des algorithmes biométriques standard, qui ignorent que les caractéristiques qui leur sont présentées sont déformées. Pendant la comparaison, l'échantillon biométrique de test doit être déformé exactement de la même manière, sinon il ne peut pas correspondre à l'échantillon enregistré. Cette distorsion doit également satisfaire la contrainte selon laquelle différents types de distorsion ne peuvent pas correspondre. Ainsi, la nature révocable de ce schéma est fournie par la distorsion, dans la mesure où ce n'est pas la biométrie réelle de l'utilisateur qui est stockée, mais simplement une parmi un nombre arbitrairement élevé de permutations possibles. L'un des principaux avantages de ce système est qu'il est indépendant de l'algorithme de comparaison biométrique.

Le régime de protection, appelé BioPhasor [Teoh *et al.*, 2007], est un exemple de cette catégorie. C'est une approche basée sur la projection aléatoire qui adresse les exigences de sécurité et d'inversibilité manquantes dans la projection aléatoire traditionnelle et le régime *BioHashing*. Les étapes de *BioPhasor* sont les suivantes :

1. Générer m vecteurs aléatoires et les stocker dans une carte à puce infalsifiable.
2. Appliquer l'algorithme de *Gram-Schmidt* (Annexe A.2) sur les m vecteurs aléatoires pour calculer une matrice orthogonale Δ.
3. Transformer le modèle original z (de taille n) en utilisant la formule suivante :

$$y_i = \frac{1}{n}\sum_{i=1}^{n} \arctan(\frac{z_i}{\Delta_{ij}}) \tag{3.5}$$

$j = 1, ..., m \mid m \leq n$ et $\Delta_{ij} \neq 0$
y est le modèle transformé.

4. Quantifier le modèle transformé y comme suit :

$$t_i = \begin{cases} 0 & \text{si} \quad 0 < y_i \leq \pi \\ 1 & \text{si} \quad \pi < y_i \leq 0 \end{cases} \quad \text{avec } i = 1, ..., n \tag{3.6}$$

t est le modèle protégé de taille m.

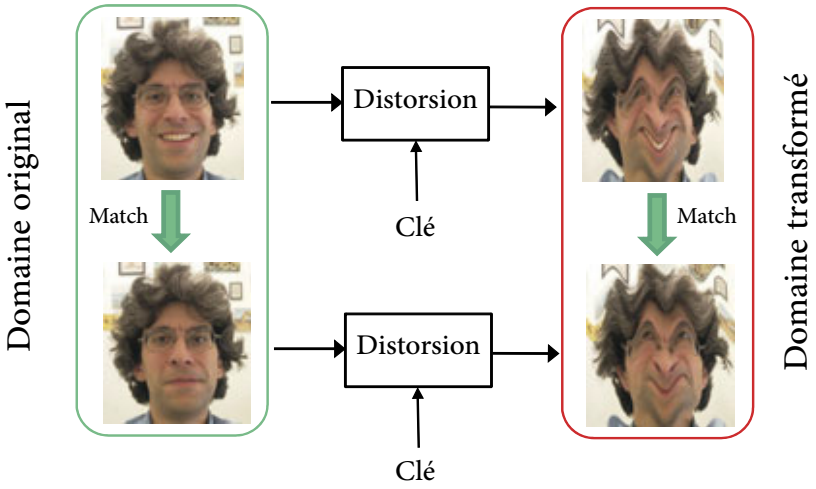
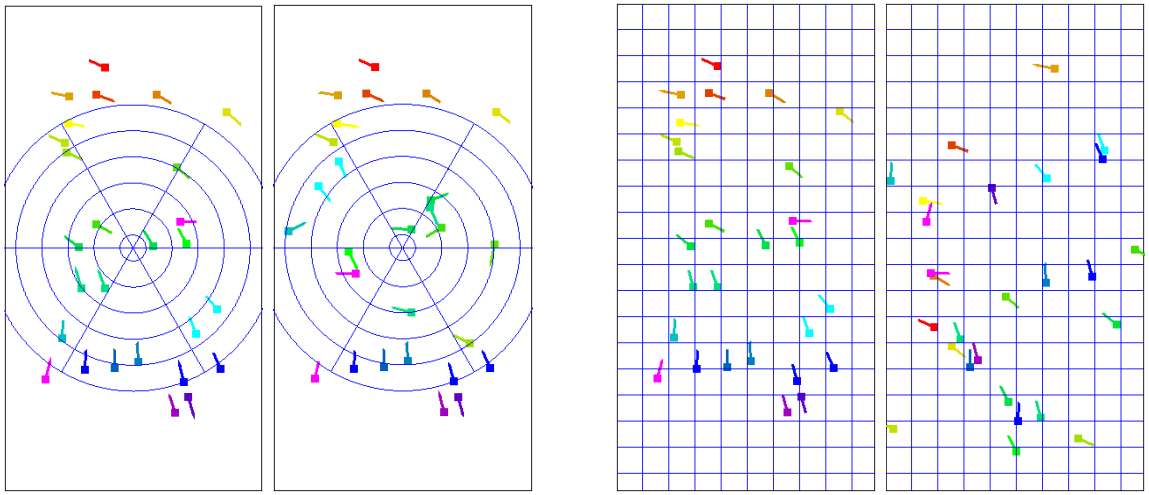


FIGURE 3.6 – Exemples de transformations géométriques.

L'avantage de *BioPhasor* par rapport au *Biohashing* est que la transformation *Arc tangente* améliore l'inversibilité et la sécurité, et elle rend la récupération du modèle original encore plus compliquée. Cependant, il est toujours possible d'avoir une approximation acceptable à partir d'un modèle protégé par *BioPhasor*.

L'une des formes de la fonction de transformation les plus significatives dans le contexte des approches de transformation non inversible, est l'utilisation des distorsions ou des transformations géométriques pour protéger les modèles biométriques [Ratha *et al.*, 2006, Ratha *et al.*, 2007]. Parmi les exemples de fonctions non inversibles qui ont été proposées dans le but de transformer les minuties des empreintes digitales, on trouve *les transformations Cartésiennes, Polaires et Fonctionnelles*[Ratha *et al.*, 2007]. Ces fonctions ont été utilisées pour transformer les points caractéristiques (Les minuties) des empreintes digitales de telle sorte qu'un classifieur puisse encore être appliqué aux minuties transformées. Dans *la transformation Cartésienne*, l'espace des minuties est mis en mosaïque dans une grille rectangulaire et chaque cellule (contenant éventuellement des minuties) est déplacée vers une nouvelle position dans la grille correspondant aux translations définies par la clé spécifiée par l'utilisateur. *La transformation Polaire* est similaire à la transformation cartésienne, à la différence que l'image est maintenant organisée en un certain nombre de coquilles concentriques et que chaque coquille est divisée en secteurs.

Comme la taille des secteurs peut être différente (les secteurs proches du centre sont plus petits que ceux qui en sont éloignés), des restrictions sont imposées au vecteur de translation généré à partir de la clé, de sorte que la distance radiale du secteur transformé ne soit pas très différente de la distance radiale de la position d'origine. La figure 3.7 présente des exemples de points caractéristiques avant et après des transformations polaires et cartésiennes.

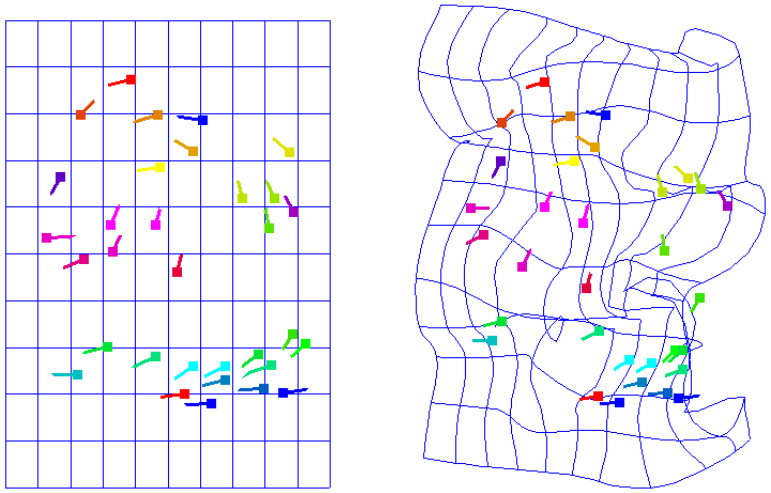


(a) (b) (c) (d)

FIGURE 3.7 – Illustration des fonctions de *transformation Cartésienne* et *Polaire* utilisées pour générer des données biométriques révocables. (a) Points de repère originaux sur une grille radiale, (b) points de repère transformés après *transformation polaire*, (c) points de repère originaux sur une grille rectangulaire, et (d) points de repère transformés après *transformation Cartésienne*.

Pour *la transformation fonctionnelle*, les auteurs ont proposé un mélange de gaussiennes 2D et de champ de potentiel électrique dans une distribution 2D de charge aléatoire comme moyen de translation pour les minuties. Dans les trois transformations, deux ou plusieurs points caractéristiques peuvent correspondre au même point dans le domaine transformé. Par exemple, dans la transformation cartésienne, deux cellules ou plus peuvent être mises en correspondance avec une seule cellule, de sorte que même si un adversaire connaît la clé et donc la transformation entre les cellules, il ne peut pas déterminer la cellule d'origine à laquelle appartient un point caractéristique, car chaque point caractéristique peut appartenir indépendamment à l'une des cellules possibles. Cela confère un degré limité de non-inversibilité à la transformation. De plus, comme les transformations utilisées sont localement lisses, les taux d'erreur ne sont pas affectés de manière significative et la discriminabilité des points caractéristiques est préservée dans une large mesure. Notez que la clé pour obtenir de bonnes performances de reconnaissance est la disponibilité d'un algorithme d'alignement qui peut pré-aligner (enregistrer) avec précision les images d'empreintes digitales ou les caractéristiques de minuties avant la transformation.

Les transformations cartésiennes et *polaires* ont pour inconvénient de convertir de petites différences dans les emplacements relatifs de deux points caractéristiques dans l'espace original en grandes différences dans l'espace transformé, ce qui entraîne un grand nombre de faux rejets. En conséquence, les auteurs recommandent l'utilisation d'une transformation localement lisse. Cependant, pour être cryptographiquement sûre, la transformation ne doit pas être globalement lisse, sinon il serait facile de l'inverser. *La transformation fonctionnelle de pliage de surface* que les auteurs proposent est localement lisse mais pas globalement (Figure 3.8) ;



(a) (b)

FIGURE 3.8 – Dans une transformation par pliage de surface, la position et l'orientation des points caractéristiques sont modifiées par une fonction de correspondance (en anglais *Mapping function*).

En littérature, plusieurs approches de transformation non-inversible ont été appliquées et testées sur plusieurs modalités biométriques. Les méthodes [Zuo *et al.*, 2008, Hämmerle-Uhl *et al.*, 2009, Rathgeb et Uhl, 2010] sont des exemples d'approche de transformation non-inversible destinées aux systèmes de la reconnaissance d'iris. [Sutcu *et al.*, 2005, Teoh *et al.*, 2007] sont des régimes de protection non-inversible pour les systèmes de reconnaissance faciale. Les travaux [Maiorana *et al.*, 2008c, Maiorana *et al.*, 2008b, Maiorana *et al.*, 2010] sont des schémas de protections des systèmes de signature en ligne. Pour les systèmes de reconnaissance des empreintes digitales, nous pouvons distinguer les travaux significatifs suivants [Lee *et al.*, 2007a, Chikkerur *et al.*, 2008, Ferrara *et al.*, 2012] .

3.3.3 Crypto-systèmes biométriques

Les crypto-systèmes biométriques [Uludag *et al.*, 2004] permettent d'adapter les protocoles cryptographiques aux données biométriques, qui sont par nature des données brui-

tées. Les cryptosystèmes biométriques sont en quelque sorte similaires aux systèmes de génération de clés basés sur des mots de passe car ils ont été développés à l'origine dans le but de sécuriser une clé cryptographique à l'aide de caractéristiques biométriques ou de générer directement une clé cryptographique à partir de caractéristiques biométriques. Étant donné que les caractéristiques biométriques acquises lors de l'inscription et de l'authentification sont différentes, ces caractéristiques ne peuvent pas être utilisées directement pour la génération de clés cryptographiques. Afin de faciliter la génération de clés, certaines informations publiques sur les caractéristiques biométriques sont stockées dans la base de données lors de l'inscription. Ces informations publiques sont généralement appelées données d'aide ou données auxiliaires (*Helper Data* en anglais). Le *Helper Data* est utilisé lors de l'authentification pour extraire une clé cryptographique des caractéristiques biométriques de test par un processus connu sous le nom de mécanisme de récupération. La correspondance est effectuée indirectement en vérifiant la validité de la clé extraite ou en utilisant directement la clé dans une autre application.

Les crypto-systèmes biométriques se répartissent généralement en deux catégories [Jain *et al.*, 2008, Rathgeb et Uhl, 2011] : les crypto-systèmes de type *key-binding* et les crypto-systèmes de type *key-generation*, en fonction de la manière dont le *Helper Data* est obtenu. Lorsque le *Helper Data* est obtenu en liant une clé cryptographique (indépendante des caractéristiques biométriques) au modèle biométrique, on parle de crypto-système biométrique de type *key-binding*. Si les données auxiliaires sont dérivées uniquement du modèle biométrique et que la clé cryptographique est directement générée à partir des données auxiliaires et des caractéristiques biométriques de la requête, il s'agit alors d'un crypto-système biométrique de type *key-generation*. Il est essentiel de mettre l'accent sur le fait qu'il n'est pas nécessaire que le modèle sécurisé soit secret dans ce genre de système. Par conséquent, il ne doit pas révéler des informations pertinentes sur le modèle biométrique original ou la clé cryptographique. Ainsi, les crypto-systèmes biométriques résolvent simultanément les problèmes difficiles de gestion des clés cryptographiques et de protection des modèles biométriques. Pour cette raison, ce sujet fait l'objet de recherches actives dans les communautés biométriques et cryptographiques.

Le mécanisme de récupération d'un crypto-système biométrique est capable de compenser les variations intra-classe des données biométriques, généralement par l'utilisation de techniques de codage de correction d'erreurs. Ce dernier est couramment utilisé dans les systèmes de télécommunication pour permettre la transmission fiable de données numériques sur des canaux de communication non fiables. Ces systèmes assurent une tolérance aux erreurs en ajoutant des informations supplémentaires (redondantes) au message avant sa transmission. Dans le contexte des crypto-systèmes biométriques, les caractéristiques biométriques disponibles lors de l'inscription sont analogues au message transmis. Les données auxiliaires sont à peu près équivalents aux informations redondantes qui sont ajoutées au message. La capacité d'un crypto-système biométrique à gérer les variations intra-classe dépend directement de la quantité d'informations redondantes utilisées pour la correction des erreurs. Une plus grande redondance entraîne généralement une plus

grande tolérance aux erreurs et, par conséquent, une plus grande stabilité de la clé. La stabilité de la clé fait référence à la probabilité de récupérer la bonne clé secrète ou de générer la même clé cryptographique lors de chaque tentative d'authentification.

3.3.3.1 *Crypto-systèmes biométriques de type key-binding*

Dans un crypto-système biométrique de type *key-binding*, le modèle biométrique est sécurisé en le liant de manière monolithique à une clé secrète dans un cadre cryptographique, comme le montre la figure 3.9, une entité unique qui intègre à la fois la clé et le modèle est stockée dans la base de données en tant que *Helper data*. Cette derniére ne révèle pas beaucoup d'informations sur la clé ou le modèle biométrique, c'est-à-dire qu'il est difficile de décoder la clé ou le modèle sans connaître les données biométriques de l'utilisateur. La correspondance dans un système de type *key-binding* implique la récupération de la clé à partir du *Helper data* en utilisant les caractéristiques biométriques de la requête et en vérifiant la validité de la clé. En général, les données d'aide sont une association d'un code correcteur d'erreurs (qui est indexé par la clé secrète) et du modèle biométrique. Lorsqu'une requête biométrique diffère du modèle dans une certaine tolérance d'erreur, un *codeword* associé avec une quantité d'erreur similaire peut être récupéré. Ce *codeword* avec des erreurs peut être décodé pour obtenir le *codeword* exact et donc, récupérer la clé intégrée. La récupération de la clé correcte implique une correspondance réussie. La tolérance aux variations intra-classe des données biométriques est déterminée par la capacité de correction des erreurs du *codeword* associé.

Parmi les protocoles cryptographiques les plus couramment utilisés dans un scénario *key-binding*, nous pouvons mentionner le *fuzzy commitment* [Juels et Wattenberg, 1999] où une clé secrète est choisie par l'utilisateur, codée et le résultat est soumis à un test XOR avec le modèle biométrique pour assurer la sécurité et la confidentialité du modèle. Dans le même contexte, et afin de faire face à un ensemble de données non ordonnées, il existe une autre approche connue sous le nom de *fuzzy vault* [Juels et Sudan, 2006], qui se base sur le partage de secret par polynômes. Le *fuzzy commitment* et le *fuzzy vault* ont été largement utilisés pour les systèmes biométriques reposant sur différents identifiants. Le concept *Fuzzy commitment* a été appliqué, entre autres, à la biométrie auriculaire [Tuyls *et al.*, 2004], aux empreintes digitales [Tuyls *et al.*, 2004, Tuyls *et al.*, 2005], au visage 3D [Kelkboom *et al.*, 2007], à l'iris [Hao *et al.*, 2006, Rathgeb et Uhl, 2009] et aux signatures en ligne [Maiorana *et al.*, 2008a, Maiorana et Campisi, 2009]. Concernant l'approche *Fuzzy vault*, elle a été appliquée aux empreintes digitales [Uludag et Jain, 2004b, Yang et Verbauwhede, 2005, Nandakumar *et al.*, 2007a], aux signatures [Freire *et al.*, 2007], aux visages [Nyang et Lee, 2007], aux iris [Lee *et al.*, 2007b] et aux empreintes palmaires [Kumar et Kumar, 2009].

L'approche Fuzzy commitment peut être appliqué aux systèmes biométriques où le vecteur de caractéristiques est une chaîne binaire de longueur fixe. Supposons que le modèle d'enrôlement X^{Enr} soit une chaîne binaire de longueur d bits. Pendant l'enrôlement, un codeword C de la même longueur est sélectionné. Ce codeword est indexé de manière

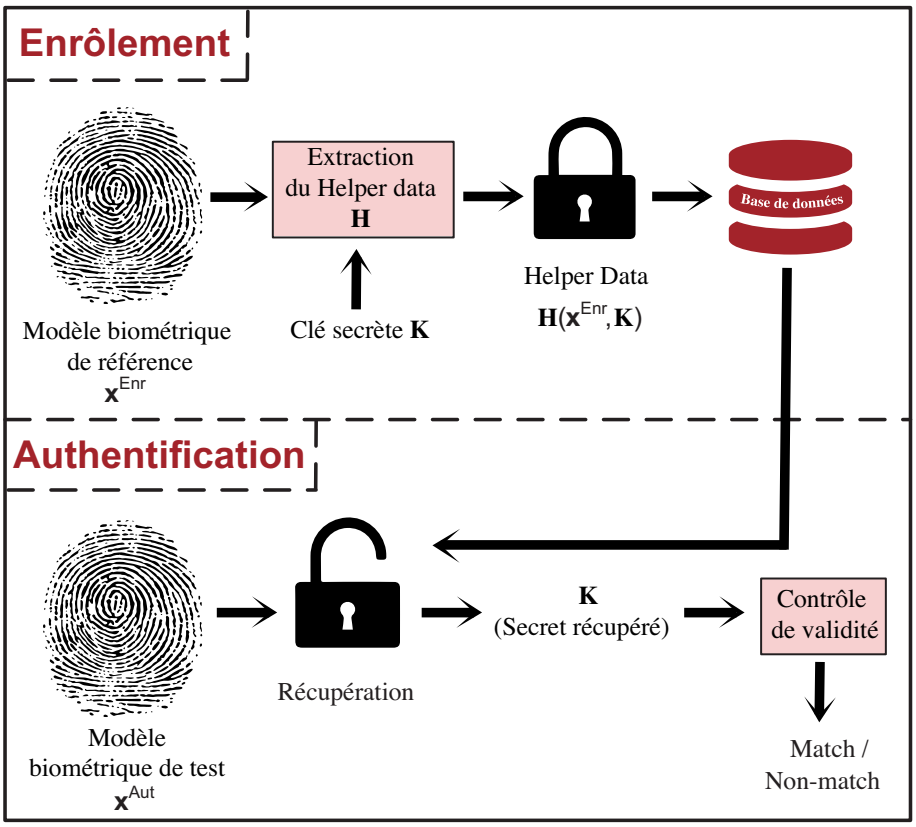


FIGURE 3.9 – Processus d'authentification lorsque le modèle biométrique est sécurisé à l'aide d'un crypto-système biométrique de type *key-binding*.

unique par une clé secrète K de longueur m bits (m est inférieur à d) et le paramètre $(d-m)$ est une mesure de la redondance dans le code correcteur d'erreurs. Le codeword C est ensuite lié au vecteur de caractéristiques biométriques X^{Enr} pour générer le *Helper data*. Les données auxiliaires se composent du *fuzzy commitment* $(X^{Enr} \oplus C)$ et de $g(K)$, où $g(.)$ est une fonction cryptographique et $\oplus$ représente l'opération OU exclusif (XOR). Lors de l'authentification, l'utilisateur présente un vecteur biométrique X^{Aut}. Désormais, le calcul du codeword C' avec erreurs peut se faire, $C' = X^{Aut} \oplus (X^{Enr} \oplus C)$. Si X^{Aut} est proche de X^{Enr}, C' sera proche de C puisque $X^{Aut} \oplus X^{Enr} = C' \oplus C$. Par conséquent, C' peut être décodé pour obtenir le codeword le plus proche C^* , qui serait égal à C à condition que la distance entre C et C^* soit inférieure à la capacité de correction d'erreurs du code. A partir de C^*, on peut calculer K^*. La correspondance est réussi si $g(K^*) = g(K)$.

L'approche *fuzzy vault* est une certaine forme d'amélioration de *fuzzy commitment*. Le principe du fonctionnement général de *fuzzy vault* se traduit comme suit : pendant l'enrôlement, une clé utilisateur K est mise à profit pour construire un polynôme P. La projection polynomiale $P(T)$ du modèle biométrique de référence T est par la suite calculée. Enfin, un bruit est ajouté à $P(T)$ pour générer le *Helper data* de *fuzzy vault*. Durant l'authentification, un code correcteur d'erreur est appliqué sur le modèle de test

et le *helper data* pour reconstruire le polynôme P et récupérer ainsi la clé K.

3.3.3.2 Crypto-systèmes biométriques de type key-generation

La génération directe de clés cryptographiques à partir de la biométrie est une proposition attrayante [Monrose *et al.*, 2000, Chang *et al.*, 2004, Blanton et Aliasgari, 2013], mais c'est un problème difficile pour deux raisons : (a) la variabilité intra-classe des caractéristiques biométriques, et (b) la nature non uniforme de la distribution de probabilité des caractéristiques biométriques. Le concept du *Helper data* peut être utilisé pour résoudre le premier problème. Dans ce scénario, le *Helper data* est dérivé en utilisant uniquement le modèle biométrique et le mécanisme de récupération facilite la reconstruction exacte du modèle lorsqu'on lui présente une requête proche du modèle, comme l'illustre la figure 3.10. Les premiers schémas de génération de clés biométriques utilisaient des schémas de quantification spécifiés par l'utilisateur. Les informations sur les limites de quantification sont stockées en tant que données auxiliaires et sont utilisées pendant l'authentification pour tenir compte des variations intra-classe. Il est également possible de faire appel à des schémas de codage de correction d'erreurs pour générer le *Helper data* à partir des caractéristiques biométriques. Les méthodes [Arakala *et al.*, 2007, Li *et al.*, 2008, Yang *et al.*, 2012] sont des exemples des crypto-systèmes biométriques de type key-generation pour les systèmes de la reconnaissance des empreintes digitales. [Zhou, 2007, Sutcu *et al.*, 2007] sont des crypto-systèmes key-generation destinés pour les systèmes de reconnaissance faciale.

Les approches les plus populaires de cette catégorie sont les systèmes connus sous les nominations : *Secure Sketch* [Bringer *et al.*, 2008] et *Fuzzy Extractor* [Dodis *et al.*, 2008]. Le *Fuzzy Extractor* extrait une chaîne uniformément aléatoire d'une entrée d'une manière tolérante aux erreurs, c'est-à-dire d'une manière telle que même si l'entrée réelle diffère de l'entrée originale, tout en restant proche, la chaîne peut être exactement récupérée. Tandis que le *Secure Sketch* permet une reconstruction exacte de l'entrée en utilisant certaines informations publiques extraites de celle-ci, à savoir le *Sketch*, qui ne révèle pas d'informations significatives sur l'entrée elle-même, et une réplique bruyante de l'entrée suffisamment proche de l'originale. Dans [Sutcu *et al.*, 2007], les questions pratiques liées à la conception d'un système *Secure Sketch* ont été analysées avec une application spécifique à la biométrie du visage.

Il faut rappeler que la cryptographie traditionnelle exige que les clés cryptographiques aient une distribution aléatoire uniforme. Or, il est bien connu que les caractéristiques biométriques ne sont pas uniformément distribuées. Le *Fuzzy Extractor* a été proposé comme primitif cryptographique qui génère une clé cryptographique uniformément aléatoire à partir des caractéristiques biométriques. Le *Helper* data fait partie intégrante d'un *Fuzzy Extractor*, ce qui permet de résoudre le problème de la stabilité de la clé. Le problème de non-uniformité peut être traité en appliquant des fonctions de hachage cryptographiques au modèle biométrique. Rappelons qu'une fonction de hachage cryptographique possède des propriétés souhaitables telles que la résistance à la pré-image et la résistance aux collisions. Ces propriétés facilitent l'extraction de chaînes binaires uniformément aléatoires

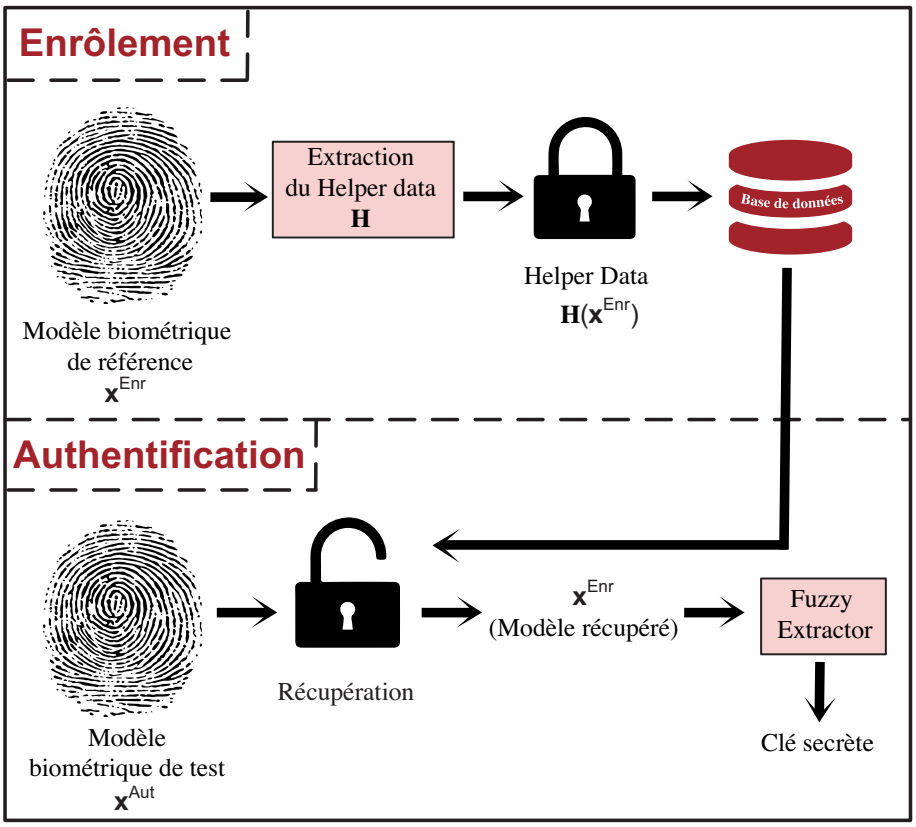


FIGURE 3.10 – Processus d'authentification lorsque le modèle biométrique est sécurisé à l'aide d'un crypto-système biométrique de type *key-generation*.

à partir des caractéristiques biométriques.

3.3.4 Méthodes hybrides

Les méthodes hybrides ont tendance à combiner deux ou plusieurs approches de base (biohashing, transformation non-inversible, crypto-système de type *key-binding* et crypto-système de type *key-generation*) pour générer un modèle biométrique révocable. Les systèmes de protection des modèles présentés ci-dessus ont leurs propres avantages et limites en termes de sécurité des modèles, de coût de calcul, d'exigences de stockage, d'applicabilité à différents types de représentations biométriques et de capacité à gérer les variations intra-classes des données biométriques. Les méthodes hybrides essaient de combiner les avantages de ces régimes de protection sans souffrir de leurs inconvénients respectifs. Les phases de révocabilité, de discrimination et de sécurité constituent la partie principale de ce genre de méthode.

[Boult, 2006] a proposé de telles méthodes pour la biométrie du visage. Dans ses approches, les bio-tokens ont été utilisés, qui sont des jetons d'identité révocables produits en appliquant une transformation révocable aux données biométriques, de sorte que la correspondance d'identité est effectuée sous la forme codée/révocable. Cette approche

combine les concepts de transformation des données, de mesures d'apprentissage fortes et de cryptage des données biométriques. [Ghany *et al.*, 2012, Zhu *et al.*, 2012] ont employé la projection aléatoire avec le *Fuzzy Vault* pour générer des modèles d'empreintes vocales. [Wong *et al.*, 2014] ont proposé une méthode hybride qui est une combinaison de *Multi-Line Code* (MLC) et de *Secure Sketch* (SS) et l'ont appelée *Cancelable Secure Sketch* (CaSS). Le MLC est généré en cinq phases, à savoir : (i) extraction des points caractéristiques (ii) code multiligne (MLC) (iii) projection aléatoire (iv) analyse en composantes principales du noyau (KPCA) (v) binarisation. Le *Secure Sketch* est une combinaison de la chaîne de bits générée par le MLC (c'est-à-dire la chaîne de bits de l'étape de binarisation) et d'un *Codeword* aléatoire, choisi dans le *Codebook*, par exemple les codes *Reed Solomon* (RS) et BCH (Bose-Chaudhuri-Hocquenghem).

Les travaux [Voderhobli *et al.*, 2006, Nandakumar *et al.*, 2007b, Boult *et al.*, 2007, Ong *et al.*, 2008] sont également de bons exemples qui illustrent les méthodes hybrides adoptées pour protéger les gabarits biométriques.

3.4 État de l'art des schémas de protection des modèles d'empreinte digitale

Dans la littérature, plusieurs méthodes abordent la question de la protection des gabarits d'empreintes digitales, chaque méthode a ses avantages et ses inconvénients face aux défis précédemment mentionnés. Dans cette section, nous nous intéressons plus particulièrement aux approches de protection des modèles des empreintes digitales proposées dans la littérature et qui font usage des transformations non inversibles. En général, elles peuvent être distinguées par le mode d'alignement qu'elles utilisent. À noter que l'objectif de l'alignement de minuties de deux empreintes est de trouver la configuration géométrique (rotation, translation) qui produit un maximum de paires de minuties semblables.

3.4.1 Approches avec alignement

Il existe d'abord des schémas qui utilisent des pré-alignements et reposent généralement sur la détection et l'enregistrement des singularités (Core, Delta) pour surmonter les problèmes de translation, de rotation et de changement d'échelle. À cet égard, Ratha et al. [Ratha *et al.*, 2007] ont proposé trois types de transformations : Cartésienne, Polaire et Fonctionnelle, qui sont extrêmement difficiles à inverser mais souffrent d'une dégradation des performances due aux variations intra-classes. De plus, certains travaux ont réussi à dégénérer la troisième transformation en cas de compromission du modèle protégé et des paramètres de la fonction [Quan *et al.*, 2008, Shin *et al.*, 2009] .

Dans la même catégorie, Ang et al [Ang *et al.*, 2005] ont proposé une méthode de génération de modèles d'empreintes digitales révocable. Le principe est d'appliquer une transformation géométrique dépendante de la clé qui implique le point Core et une ligne dont l'orientation est comprise entre 0° et 180° (définie par la clé délivrée). Par la suite, chaque minutie située sous la ligne se déplace symétriquement vers la région supérieure

pour créer une nouvelle représentation protégée. Le problème de cette méthode est qu'elle préserve tout de même certaines informations de l'empreinte digitale originale, même après la transformation.

Yang et al. [Yang *et al.*, 2009] ont exploité les caractéristiques globales et locales extraites pour développer un nouveau système de protection. Il consiste à projeter perpendiculairement la distance entre chaque paire de minuties dans le cercle formé autour du point Core de l'empreinte digitale. Ensuite, les propriétés triangulaires formées sont extraites pour effectuer la correspondance. La sécurité de ce système est effectivement garantie mais les performances ne sont pas assez prometteuses.

Moujahdi et al. [Moujahdi *et al.*, 2014] ont produit une nouvelle représentation sécurisée et révocable basée sur les distances entre les points singuliers et toutes les autres minuties des empreintes digitales. L'idée principale est de construire des courbes spéciales en spirale, qui font référence aux modèles d'empreintes digitales originaux au lieu des représentations basées sur les minuties. Les performances de reconnaissance de ce système semblent être préservées, mais le problème est que si une courbe d'empreinte digitale est compromise, l'attaquant peut révéler les distances utilisées pour générer le modèle protégé. Cependant, de nombreuses versions améliorées sont apparues par la suite [Ali et Prakash, 2015, Ali et Prakash, 2017, Ali et Prakash, 2019] .

Bien que la plupart des schémas basés sur l'enregistrement soient robustes et offrent un haut degré de sécurité, de révocabilité, de diversité et de non-inversibilité, les performances de reconnaissance restent proportionnelles à la précision de la détection des points singuliers, car de légères déformations de leur position peuvent entraîner des erreurs de correspondance. Leur utilisation comme caractéristiques globales, dans ce cas, est donc très risquée et demande une attention particulière. D'autre part, il existe de nombreuses approches qui ne nécessitent aucun point d'enregistrement ou alignement et qui sont connues sous le nom de *Registration-free / Alignement-free methods*. Dans ce type d'approche, la comparaison de deux empreintes se base sur la structure locale des minuties. Les structures locales sont caractérisés par des attributs invariants ce qui élimine l'étape de pré-alignement. Ces approches sont donc plus rapides, plus robustes aux distorsions que les approches globales mais généralement moins distinctives (car elles relâchent les relations spatiales entre les minuties sur le plan global qui sont extrêmement discriminantes). Nous citons dans ce qui suit quelques travaux de cette catégorie.

3.4.2 Approches sans alignement / alignement implicite

Farooq et al. [Farooq *et al.*, 2007] ont introduit une représentation sous forme de chaîne binaire des empreintes digitales dans un contexte sans alignement. Le principe de l'approche qu'ils proposent est d'extraire les caractéristiques invariantes d'un ensemble de triplets de minuties, qui sont ensuite quantifiées et hachées dans un modèle de chaîne binaire aléatoire (224 bits) via une clé spécifique à l'utilisateur. Les caractéristiques sont sélectionnées de manière à être peu sensibles aux transformations rigides telles que la

rotation et la translation. Les auteurs ont utilisé les distances locales entre les minuties, les angles formés entre l'orientation des minuties et les côtés et enfin la hauteur du triangle comme caractéristiques invariantes pour représenter chaque triplet de minuties. Bien que cette méthode se révèle très robuste contre les attaques par force brute, elle est très coûteuse, car elle traite les caractéristiques invariantes de chaque triplet de minuties possible.

Lee et al. [Lee *et al.*, 2007a] ont proposé un nouveau schéma sans alignement, qui peut fournir des modèles révocables à partir de minuties des empreintes digitales. Dans cette méthode, certaines valeurs invariantes sont principalement dérivées de l'orientation des régions locales voisines autour de chaque minutie, puis sont utilisées pour générer un déplacement pour chaque minutie sous forme de translation et de rotation en utilisant deux fonctions différentes. La méthode répond parfaitement aux propriétés de révocabilité et de non-inversibilité en cas d'attaque par force brute ; en revanche, ses performances diminuent considérablement lorsque la clé utilisateur utilisée est volée.

Ahmad et al. [Ahmad *et al.*, 2011] ont développé une autre approche sans alignement, qui consiste en un schéma de protection des modèles d'empreintes digitales basé sur les coordonnées polaires par paire. L'idée principale est d'étudier la relation relative des minuties dans un cadre bipolaire sans rotation ni décalage. Le schéma aborde les questions de révocabilité, de diversité et de non-inversibilité, mais il s'est avéré que les performances dépendent de la qualité des images des empreintes digitales.

Toujours dans le même contexte, on trouve également le schéma proposé par Wang et Hu [Wang et Hu, 2012], qui exploite les vecteurs de paires-minutiae pour générer des modèles des empreintes digitales révocables. La technique proposée est basée sur une transformation non inversible qui permet d'obtenir une correspondance de type *Infinite-to-one mapping*. Il semble que les performances du système soient satisfaisantes, mais la cohérence de la matrice de clés utilisateur pose des problèmes de stockage.

Dans le travail de Jin et al [Jin *et al.*, 2014], une autre technique de protection des modèles d'empreintes digitales a été proposée, qui est basée sur la décomposition du voisinage des minuties. Le schéma est également en mesure de générer des modèles de chaînes binaires révocables, mais cette fois par le biais de la technique *the graph-based Hamming embedding* (RGHE). La haute résistance des modèles résultants a été démontrée contre l'inversion, mais les performances restent proportionnelles à la qualité des images des empreintes digitales acquises.

Parmi les travaux récents dans ce contexte, on peut citer le travail dans [Wang et Hu, 2016], où un traitement aveugle est proposé pour produire des modèles révocables. L'idée derrière cette méthode est qu'au lieu de protéger directement les chaînes binaires dérivées des vecteurs quantifiés des paires des minuties, les échantillons de fréquence sont protégés. Nous citons ensuite le travail proposé dans [Wang *et al.*, 2017], qui présente une autre forme de génération de modèles des empreintes digitales révocables sans tenir compte de l'alignement. Le schéma est basé sur l'utilisation des structures de

minuties locales formées par les paires de minuties délimitées, car elles sont plus discriminantes et donnent de bonnes performances. Cette technique fait appel à une transformation non-inversible basée sur la transformée de Fourier discrète partielle. Bien que cette proposition semble réduire les risques d'attaque par *Attack via Record Multiplicity* (ARM), elle démontre une grande supériorité par rapport à de nombreux systèmes d'empreintes digitales révocables de l'état de l'art. Dans cette approche, les auteurs se sont également basés sur l'utilisation du transformée de Fourier discrète pour convertir les modèles de chaînes binaires en vecteurs complexes, étant donné que de nombreux modèles biométriques révocables basés sur des chaînes binaires souffrent de problèmes de récupération.

Kho et al. [Kho *et al.*, 2019] ont proposé un système de protection des empreintes digitales basé sur un descripteur de minuties sans alignement appelé *Partial Local Structure and Permutated Randomized Non-Negative Least Square* (PR-NNLS). Nous citons également un autre système récemment proposé par Ali et al. [Ali *et al.*, 2020] , qui présente une nouvelle méthode d'authentification basée sur la modification des attributs des minuties. Plusieurs autres propositions ont été suggérées dans la même thématique [Ali *et al.*, 2018, Tran *et al.*, 2018, Das *et al.*, 2012, Ahn *et al.*, 2008, Derman et Keskinöz, 2016, Kumar *et al.*, 2010, Ferrara *et al.*, 2014].

En général, chaque minutie est décrite par sa position dans un espace bidimensionnel et son orientation. Ces caractéristiques peuvent varier radicalement en présence de variations de l'empreinte digitale telles que le déplacement, la rotation et les distorsions non linéaires survenant lors de l'acquisition de l'image d'empreinte digitale. Tous ces facteurs conduisent certainement à une diminution de la précision de la correspondance (la protection des modèles d'empreintes digitales devient plus compliquée dans ce cas). Les systèmes de reconnaissance d'empreintes digitales traitent normalement ce problème en appliquant l'alignement des minuties, ce qui n'est pas toujours évident dans le contexte de la protection des modèles d'empreintes digitales.

3.5 Bilan du chapitre

Après avoir exposé les exigences d'un schéma idéal de protection des modèles biométriques, nous avons pu voir deux grandes familles de solutions. Principalement, des solutions basées sur l'architecture du système et des solutions logicielles orientées sur le traitement du modèle biométrique. Dans ce chapitre, nous avons accordé une attention particulière aux approches de transformation non-inversible dédiées pour la protection des modèles des empreintes digitales, puisqu'elles représentent la même catégorie des approches que nous proposons dans le chapitre 5.

Chapitre

4 Évaluation de performance / robustesse des schémas de protection

Sommaire

4.1 Introduction

Avec l'utilisation de plus en plus répondue de la biométrie, le besoin d'évaluer et de comparer les différents systèmes entre eux devient impératif. Malheureusement, jusqu'à présent, il n'existe pas encore de méthode systématique et standardisée pour certifier et garantir la fiabilité des systèmes biométriques. Les points à évaluer se présentent généralement sous trois critères :

1. L'évaluation de la performance du système, qui mesure les taux d'erreur du système ainsi que son efficacité.

2. L'évaluation de la sécurité et du degré de préservation de la vie privée, qui mesure la robustesse du système aux différentes attaques.

3. L'évaluation de l'usage, qui mesure l'acceptabilité et le taux de satisfaction des utilisateurs.

Dans ce chapitre, nous ferons état des différents facteurs de performance et de sécurité qui seront utilisés dans le chapitre 5 pour analyser les deux méthodes proposées dans cette thèse. La performance sera analysée en terme des taux d'erreur [Adler et Schuckers, 2007]. La sécurité sera analysée en terme de robustesse contre les attaques qui visent les modèles biométriques. L'analyse de la sécurité des systèmes existants est principalement basée sur la complexité des attaques à force brute qui suppose que les données biométriques sont uniformes. Cependant, en pratique, un adversaire peut exploiter la nature non uniforme des données biométriques pour lancer une attaque qui peut exiger beaucoup moins de tentatives pour atteindre la sécurité du système. Alors une analyse de sécurité rigoureuse, comme celle de [Nagar *et al.*, 2010], est très nécessaire pour analyser correctement la sécurité des approches de protection des modèles biométriques.

Dans ce chapitre, dans un premier temps, nous allons présenter les critères d'évaluation de la convivialité des systèmes biométriques (sous-section 4.2.1), en mettant l'accent sur des métriques de divers natures que sont les mesures des taux d'erreur fondamentaux (sous-section 4.2.1.1) et ceux des systèmes d'authentification (sous-section 4.2.1.2), ainsi que les courbes de performances (sous-section 4.2.1.3).

Ensuite, nous allons détailler le protocole FVC [Maio *et al.*, 2002, Maltoni *et al.*, 2009, Cappelli *et al.*, 2005] (*Fingerprint Verification Competition protocol*) et le protocole *1-vs-1* dans la sous-section 4.2.2, qui seront très utile pour évaluer la convivialité de nos deux approches proposées et pour les comparer correctement et d'une manière équitable avec les approches les plus significatives dans l'état de l'art. Dans la section 4.3, nous allons détailler plusieurs équations analytiques, basées sur [Nagar *et al.*, 2010], pour mesurer la résistance des approches de transformation de caractéristiques face à divers types d'attaques qui menacent les modèles biométriques.

4.2 Évaluation de performance

La discussion suivante se concentre sur les empreintes digitales puisqu'elles sont l'objet de cette thèse, bien qu'elle soit valable pour tout autre identifiant biométrique. Dans un système de reconnaissance d'empreintes digitales, la réponse d'un classificateur (comparateur) est généralement un score de correspondance (compris dans l'intervalle $[0, 1]$) qui quantifie la similarité entre les représentations du modèle enrôlé et du modèle de test. Plus le score est proche de 1, plus le système est certain que les deux empreintes digitales proviennent du même doigt ; plus le score est proche de 0, moins le système est sûr que les deux empreintes digitales proviennent du même doigt. La décision du système est régulée par un seuil, les paires d'empreintes digitales générant des scores supérieurs ou égaux au seuil sont déduites comme paires correspondantes (c'est-à-dire appartenant au même doigt), tandis que les paires d'empreintes digitales générant des scores inférieurs au seuil sont déduites comme paires non correspondantes (c'est-à-dire appartenant à des doigts différents).

Un système de vérification biométrique typique commet deux types d'erreurs : il confond les mesures biométriques de deux doigts différents avec celles du même doigt (on parle de fausse correspondance) et il confond deux mesures biométriques du même doigt avec celles de deux doigts différents (on parle de fausse non-correspondance). Il convient de noter que ces deux types d'erreurs sont également souvent désignés par les termes *fausse acceptation* et *faux rejet*.

4.2.1 Convivialité des systèmes biométriques

Selon l'organisation Internationale de Normalisation ISO/IEC 19795-1 [19795-1, 2006], les mesures des taux d'erreur peuvent être des taux d'erreur fondamentale ou de taux d'erreur du système d'authentification.

4.2.1.1 Taux d'erreur fondamentale

- Taux d'echec à l'acquisition (*failure-to-acquire rate*, FTA) : proportion des tentatives de vérification 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 digitale, 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.

4.2.1.2 Taux d'erreur des systèmes d'authentification

- Taux de faux rejets (*false rejection rate*, FRR) : la 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 fixe, le taux de faux rejets est calculé par :

$$FRR(\tau) = FTA + FNMR(\tau) \times (1 - FTA) \tag{4.1}$$

  Le FRR est un critère de confort, car un faux rejet réduit la crédibilité des systèmes biométriques. Cette valeur ne dépend pas seulement de la conception du système mais des utilisateurs également. La probabilité de défaillance d'une personne légitime U est :

$$FAR(U) = \frac{\text{Nombre de vérifications rejetées}}{\text{Nombre total de vérifications}} \tag{4.2}$$

  Une vérification est rejetée si la distance entre le modèle de la personne U et le modèle de référence est strictement supérieure à un seuil τ.

  Le FRR final de N utilisateurs est la moyenne de tous les FRR(U) :

$$FRR = \frac{1}{N} \sum_{i=1}^{N} FRR(U) \tag{4.3}$$

- 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 fixe, le taux de fausses acceptations est calculé par :

$$FAR(\tau) = FMR(\tau) \times (1 - FTA) \tag{4.4}$$

  Le FAR est une mesure de sécurité pertinente, car une fausse acceptation peut souvent conduire à des dégâts critiques. Dans une attaque de fraude, le système est attaqué par un modèle d'une personne non autorisée qui essaie de contourner le système. La probabilité de succès contre une personne légitime U est :

$$FAR(U) = \frac{\text{Nombre d'attaques de fraude réussies}}{\text{Nombre total d'attaques de fraude}} \tag{4.5}$$

L'attaque de fraude est réussie si la distance entre le modèle de la personne U et le modèle de l'adversaire est inférieure ou égale à un seuil τ.

Le FAR final pour N utilisateurs est la moyenne de tous les FAR(U) :

$$FAR = \frac{1}{N}\sum_{i=1}^{N} FAR(U) \tag{4.6}$$

Par ailleurs, il faut noter que la mesure FAR n'est pas suffisante pour valoriser la sécurité d'un système (seulement pour mesurer relativement la résistance contre les attaques à force brute), car il y a beaucoup d'autres possibilités pour lancer des attaques prometteuses comme nous l'avons mentionné précédemment.

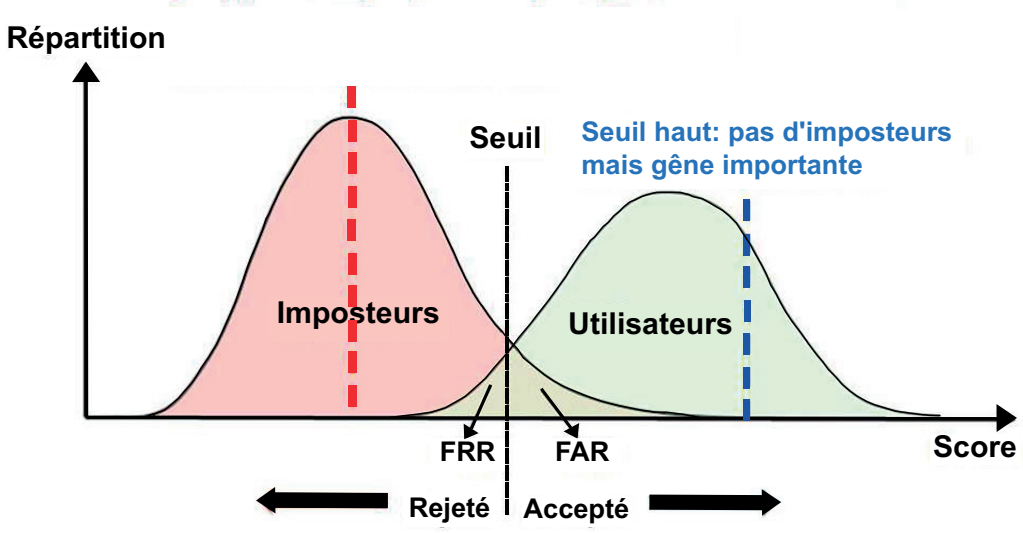
4.2.1.3 Les courbes de performance

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 systémes biométriques qui ont des performances similaires.

Il faut noter qu'un score de correspondance de similarité est dit légitime ou véritable s'il résulte de la comparaison de deux échantillons du même trait biométrique d'un utilisateur. Il est dit score imposteur s'il résulte de la comparaison de deux échantillons biométriques provenant de différents utilisateurs. Un score imposteur qui dépasse le seuil entraîne une fausse acceptation, tandis qu'un score légitime qui est inférieur au seuil entraîne un faux rejet.

- La distribution intra-classe/inter-classe ou la distribution légitime/imposteur : pour évaluer la performance d'un système de vérification, on doit calculer les scores à partir d'un large nombre de comparaisons entre des gabarits d'un même sujet. On obtient alors la distribution intra-classe (*Genuine distribution*). Il faut aussi collecter les scores des comparaisons entre des gabarits appartenant à des sujets différents pour obtenir la distribution inter-classe (*Impostor distribution*). La distribution intra-classe/inter-classe se présente comme sur la figure 4.2. Le seuil de décision du système est ensuite choisi parmi les scores possibles suivant le niveau de sécurité souhaité.
- La courbe réceptrice des caractéristiques (*Receiver operating characteristic curve*, ROC) [Egan et Egan, 1975] qui est la plus couramment utilisée pour représenter les performances du système. Elle représente l'évolution du FAR en fonction du FRR suivant les différents seuils de décision possibles.

  Au lieu de ROC, parfois le terme DET *Detection Error Tradeoff* est utilisé (Figure 4.3). Dans ce cas, le terme ROC est réservé pour représenter les taux de vrais rejets ($1 - FRR$) (*appelé Genuine Accept Rate* (GAR)) contre les taux de fausses acceptations (FAR). La principale différence entre les courbes DET et ROC est



(a)

FIGURE 4.1 – La distribution inter-classe/intra-classe.

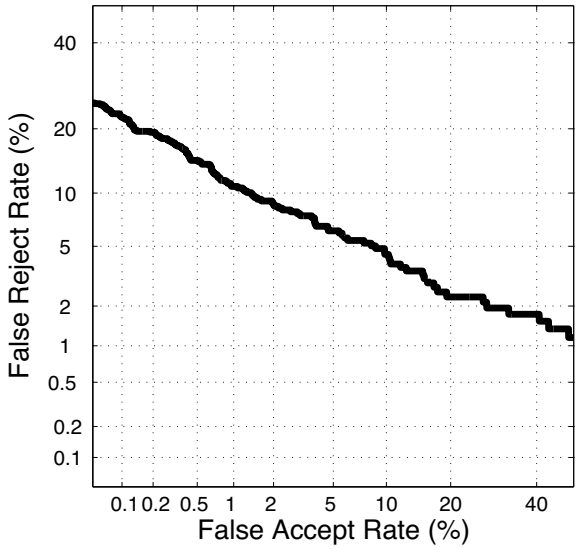
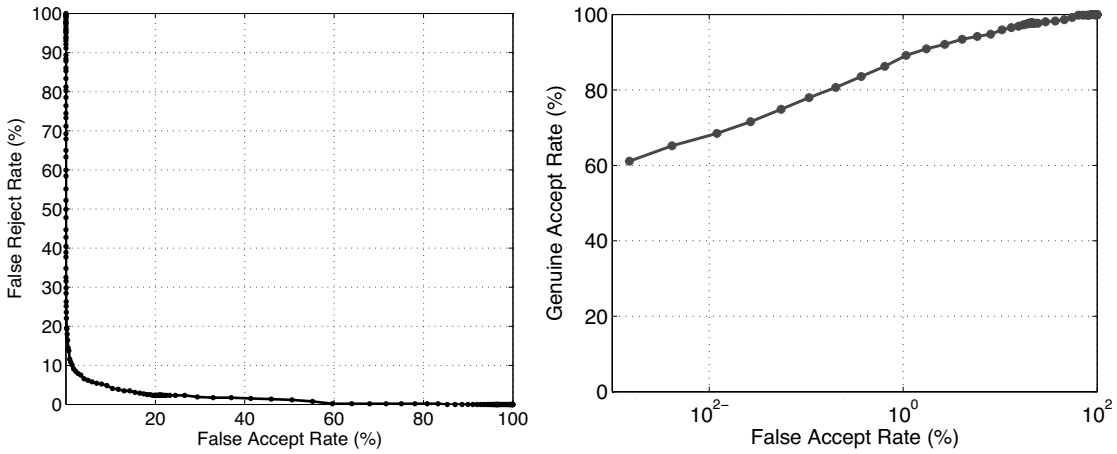


FIGURE 4.2 – Courbe DET qui trace le FRR contre le FAR dans l'échelle normale de déviation

l'utilisation de l'échelle de déviation normale dans la première. Cette courbe est utile pour donner une représentation globale sur le comportement du système. Un exemple de cette courbe est donné sur la figure 4.3.

La valeur EER (*Equal Error Rate*) est calculée comme le point où FRR(τ)=FAR(τ). Il convient de noter que les valeurs de EER dépendent également des définitions de FAR / FRR. Alors une comparaison de deux valeurs EER qui appartiennent à deux différents systèmes est raisonnable seulement si ses définitions se coïncident. L'EER constitue un



(a) (b)

FIGURE 4.3 – (a) Courbe ROC (FRR contre FAR dans l'échelle linéaire) ; (b) Courbe ROC (GAR contre FAR dans une échelle semi-logarithmique).

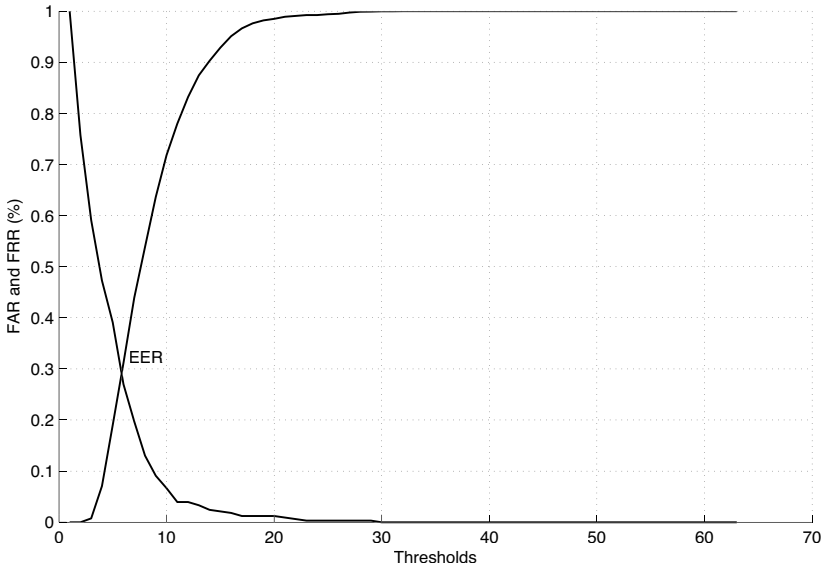
indicateur de la précision du dispositif biométrique. En d'autres termes, plus l'EER est faible, plus le système est performant. À noter que ce taux d'erreur est le plus couramment utilisé dans la littérature pour illustrer la performance des systèmes biométriques. La figure 4.4 est un exemple de courbes FAR et FRR et son intersection EER.

4.2.2 Protocoles Expérimentaux

4.2.2.1 Protocole FVC

FVC est l'abréviation de *Fingerprint Verification Competition*. Il s'agit d'une compétition internationale qui a été organisée par des laboratoires académiques pour évaluer des algorithmes de vérification d'empreintes digitales. Plusieurs bases de données (FVC2000, FVC2002 4.5, FVC2004 et FVC2006), qui sont acquises avec divers types de capteurs, ont été fournies aux participants pour leur permettre de tester leurs algorithmes en respectant un protocole de test prédéfini. Le but de cette sous-section est de détailler ce protocole FVC, qui sera utilisé pour tester notre premier algorithme proposé dans le chapitre suivant.

Dans ce protocole, pour chaque base de données FVC (chaque personne est représentée par 8 impressions), la première impression de chaque doigt est comparée avec la première impression des autres doigts pour obtenir la distribution de scores des imposteurs (*Impostor score distribution*). Pour obtenir la distribution de scores des légitimes (*Genuine score distribution*), chaque empreinte est comparée aux impressions restantes du même doigt (même identité). Il convient de noter que dans le protocole FVC, les scores calculés doivent être dans l'intervalle [0, 1], et les comparaisons symétriques ne sont pas lancées pour éviter la répétition des scores (par exemple, si un modèle biométrique T1 est déjà comparé avec un modèle T2, alors le score de T2 contre T1 ne sera pas calculé).



(a)

FIGURE 4.4 – Exemples des courbes FAR, FRR et le point EER.

La distribution des scores légitime / imposteur peut être illustrée graphiquement pour montrer comment un algorithme peut séparer ou distinguer entre les imposteurs et les utilisateurs légitimes 4.6.

Dans notre évaluation, nous utiliserons aussi deux autres facteurs pour mesurer la séparabilité des scores : premièrement, le test de *Kolmogorov-Smirnov* (K-S test) [Holland et Komogortsev, 2013]. Plus ce test est proche de 1, meilleure est la séparation des scores, ce qui signifie que la diversité est grande. Deuxièmement, le critère de séparabilité proposé par [Lee *et al.*, 2007a] :

$$Separability = \frac{\mid \mu_g - \mu_i \mid}{\sqrt{(\sigma_g^2 + \sigma_i^2)/2}} \tag{4.7}$$

μ_g et μ_i sont la moyenne de distributions des légitimes et imposteurs respectivement. σ_g^2 et σ_i^2 sont la variance de distributions des légitimes et imposteurs respectivement.

Selon le protocole FVC, les *Genuine matching scores* (gms) et les *Impostor matching scores* (ims) sont utilisées pour calculer le *False Match Rate* (FMR) et le *False Non Match Rate* (FNMR). Pour un seuil τ allant de 0 à 1 [Maio *et al.*, 2002] :

$$FMR(\tau) = \frac{\text{Cardinalité}\{ims \mid ims \geq \tau\}}{\text{Nombre de tentatives de reconnaissance imposteur}} \tag{4.8}$$

$$FNMR(\tau) = \frac{\text{Cardinalité}\{gms \mid gms < \tau\} + REJ}{\text{Nombre de tentatives de reconnaissance légitime}} \tag{4.9}$$

REJ est le nombre de rejets. Si une image ne peut pas être inscrite avec succès dans le système, le score de comparaison sera 0 pour toutes les tentatives de reconnaissance

FIGURE 4.5 – Exemple d'empreintes digitales de la base FVC2002 DB2

possibles en utilisant un des modèles rejetés. On trace FNMR en fonction de FMR pour obtenir la courbe ROC du protocole FVC en utilisant une échelle logarithmique dans les deux axes (Figure).

4.2.2.2 Protocole 1-vs-1

Concernant le protocole 1-vs-1, le modèle généré à partir du premier échantillon biométrique d'un utilisateur est comparé au modèle généré à partir du second échantillon biométrique du même utilisateur pour calculer le FRR. Pour calculer le FAR, le modèle généré à partir du premier échantillon biométrique d'un utilisateur est comparé à tous les modèles générés à partir du premier échantillon biométrique du reste des utilisateurs.

4.3 Évaluation de sécurité

L'évaluation de la sécurité d'un système de vérification biométrique concerne dans un premier lieu la réaction du système en termes de précision de reconnaissance dans les deux scénarios les plus fréquents : la vérification dans les cas de clés différentes *Different-key* et de clés volées *Stolen-key*. Le premier scénario consiste à attribuer à chaque utilisateur une clé spécifique pour pouvoir générer son modèle biométrique. Alors que pour le deuxième scénario, on suppose qu'un utilisateur légitime a perdu sa clé, qui finira par tomber entre les mains d'un intrus, qui tentera alors d'effectuer une authentification en tant qu'utilisateur légitime. Cette attaque est connue sous le nom de *Lost Token Attack*. La simulation de ce scénario consiste à attribuer la même clé utilisateur à tous les utilisateurs de la base de données et à étudier par la suite le comportement du système.

Par ailleurs, Nous nous concentrons sur la vulnérabilité d'un schéma de transformation de gabarit aux attaques par intrusion et par liaison de bases de données (*linkage* ou *cross-matching* en anglais) qui peuvent être organisées en utilisant la connaissance d'un gabarit stocké. L'intrusion consiste à accéder à un système de reconnaissance biométrique en

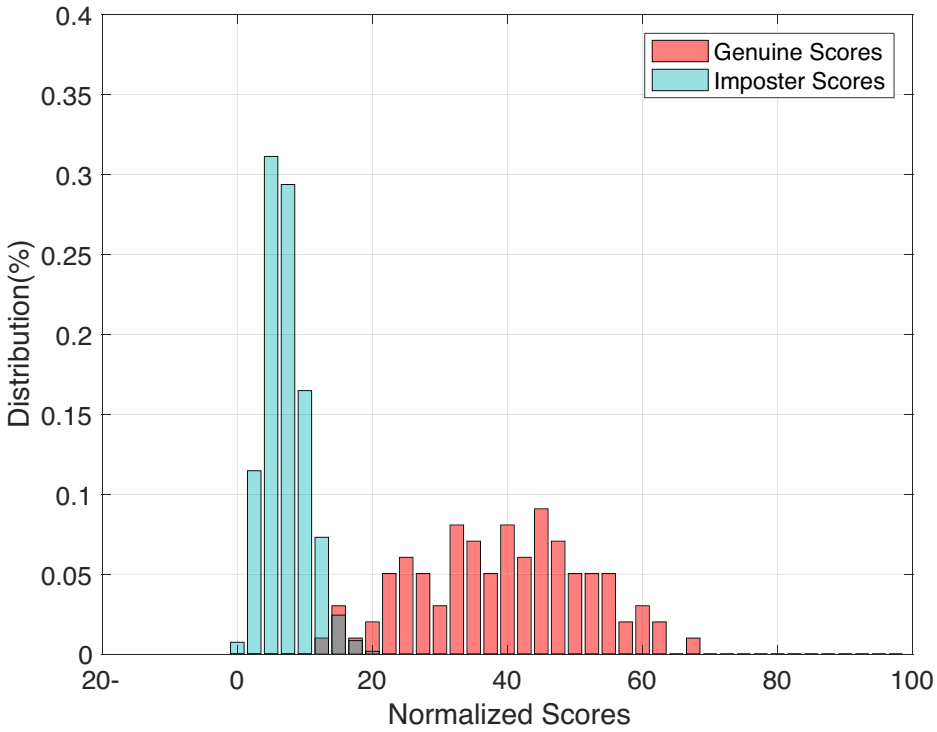


FIGURE 4.6 – Exemple d'un histogramme de distribution des scores légitime/imposteur.

présentant au système des données d'authentification falsifiées. L'intrusion compromet l'un des avantages fondamentaux de l'utilisation d'un système biométrique, à savoir la non-répudiation. D'autre part, les attaques par liaison impliquent une correspondance croisée entre les systèmes biométriques afin de suivre les utilisateurs en secret, ce qui compromet la vie privée de l'utilisateur. Il est donc important d'analyser la probabilité de succès de ces deux attaques dans un système biométrique.

Dans les sous-sections 4.3.1 et 4.3.2, nous allons détailler plusieurs équations analytiques, basées sur [Nagar *et al.*, 2010], pour mesurer la résistance des approches de transformation de caractéristiques aux attaques d'intrusion et de liaison de bases de données.

4.3.1 Mesures d'évaluation de sécurité pour les menaces d'intrusion

Pour mesurer la vulnérabilité des approches de transformation de caractéristiques à des attaques d'intrusion, nous considérons le scénario où un modèle protégé est volé à partir de la base de données et que les paramètres de transformation sont connus par l'attaquant.

Dans ce scénario, l'adversaire va essayer de récupérer le modèle original et rejouer le modèle récupéré (dans le même système) en utilisant les paramètres de transformation. Ce critère a été nommé IRIS (*Intrusion Rate due to Inversion for the Same biometric system*) :

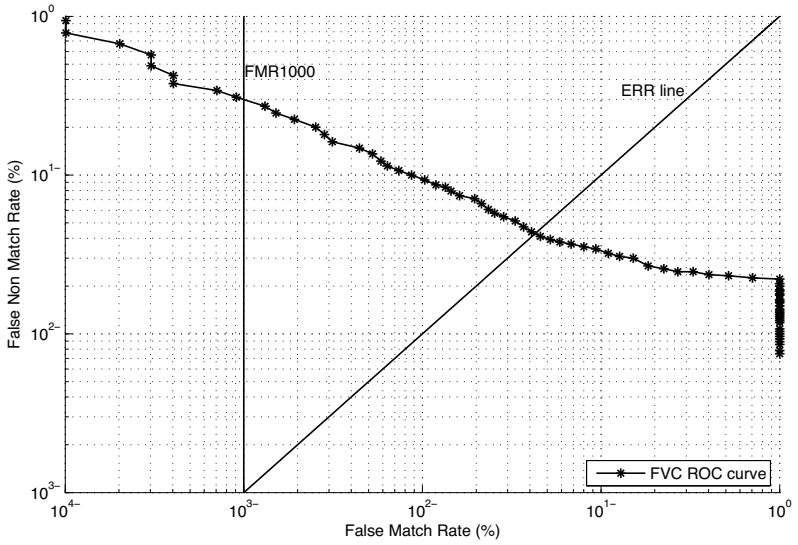


FIGURE 4.7 – Courbe ROC selon le protocole FVC.

$$IRIS(\tau) = P[D(f(f^{-1}(T_i, k_i), k_i), T_i) < \tau] \tag{4.10}$$

— D est la fonction de distance.

— f est la fonction de transformation.

— f^{-1} est la fonction de transformation inverse.

— $T_i = f(t_i, k_i)$ est le modèle transformé volé.

— t_i est le modèle original de l'identité i.

— k_i sont les paramètres de transformation de l'identité i.

Considérant le scénario précédent, mais cette fois nous allons supposer que l'adversaire lance une attaque contre un autre système biométrique qui contient la même identité volée ou compromise, en supposant que l'attaquant connaît aussi les paramètres de transformation du deuxième système. Ce critère a été nommé IRID (*Intrusion Rate due to Inversion for a Different biometric system*) :

$$IRID(\tau) = P[D(f(f^{-1}(T_i, k_i), k'_i), T'_i) < \tau] \tag{4.11}$$

— $T'_i = f(t'_i, k'_i)$ est le modèle transformé de l'identité i dans le deuxième système.

— t'_i est le modèle original de l'identité i dans le deuxième système.

— k'_i sont les paramètres de transformation de l'identité i dans le deuxième système.

IRIS et IRID dépendent du seuil τ. Il est facile d'inverser un modèle biométrique, protégé par une approche de transformation de caractéristiques, si IRIS et IRID sont à proximité de ou égal à 1.

4.3.2 Mesures d'évaluation de sécurité pour les menaces de liaison

La vulnérabilité des approches de transformation de caractéristiques aux attaques de liaison de bases de données peut être mesurée en considérant le scénario où l'adversaire vole deux modèles biométriques d'une même identité à partir de deux systèmes différents qui utilisent le même schéma de protection ; en supposant que l'attaquant connaît les paramètres de transformation de ces deux systèmes. L'adversaire peut lancer son attaque soit dans le domaine original ou dans le domaine de la transformation. Ce critère a été nommé CMR (*Cross-Matching Rate*). Il peut être définis comme suit dans les domaines des caractéristiques transformées (CMR_t) et originales (CMR_o).

$$CNR_o(\tau) = P[D(f^{-1}(T_i, k_i), f^{-1}(T'_i, k'_i)) < \tau] \tag{4.12}$$

$$CNR_t(\tau) = P[D(f(T_i, k_i), f(T'_i, k'_i)) < \tau] \tag{4.13}$$

Bien que le test de *Kolmogorov-Smirnov* (K-S test) (qui est pour mesurer la diversité), une étude des attaques de de liaison de bases de données est très utile aussi pour mesurer la diversité des approches de transformation de caractéristiques.

4.4 Bilan du chapitre

Dans ce chapitre, nous avons présenté les critères d'évaluation de performance et de sécurité qu'on va utiliser dans le chapitre 5 pour tester et comparer nos deux approches proposées avec les approches de l'état de l'art. Nous avons commencé par la description des critères généraux pour évaluer la convivialité des systèmes biométriques. Ensuite, nous avons présenté le protocole d'évaluation *FVC* et *1-vs-1*. Enfin, nous avons présenté les équations analytiques pour mesurer la vulnérabilité des approches de transformation de caractéristiques aux attaques d'intrusion et de liaison de bases de données.

Chapitre

5 Contributions à la protection des modèles d'empreinte digitale.

Sommaire

5.1 Introduction

Ce chapitre s'intéresse principalement à la présentation de deux nouvelles approches de protection des modèles pour les empreintes digitales, qui ont été proposées lors de l'élaboration de cette thèse. Les méthodes proposées sont :

1. *Fingerprint Template Protection Using Irreversible Minutiae Tetrahedrons* [Lahmidi *et al.*, 2021] : Un nouveau schéma de protection des modèles biométriques dédié pour les systèmes de reconnaissance par empreintes digitales. L'approche proposée est une technique basée principalement sur les minuties, qui permet de mettre en correspondance les empreintes digitales dans un espace transformé en utilisant des tétraèdres de minuties irréversibles.
2. *On the methodology of fingerprint template protection schemes conception : Meditations on the reliability* [Lahmidi *et al.*, 2022] : nous avons adopté une nouvelle méthodologie pour la conception de schémas de protection de modèles pour les empreintes digitales. En effet, nous avons pris en considération la spécification d'un système de vérification d'empreintes digitales non protégé pour construire un schéma de protection spécifique et adapté qui fournit le meilleur compromis entre performance et sécurité comparé à toute solution de protection générique.

Dans ce chapitre, une propre section est consacrée à chacune des nouvelles méthodes proposées. La premiére approche sera présentée dans la section 5.2 et la deuxième sera abordé dans la section 5.3. Ces deux sections partagent une structure commune : Une brève introduction du problème visé, ensuite une description de l'approche, et enfin une présentation et une discussion des résultats expérimentaux. Le bilan du chapitre sera l'objet de la section 5.4.

5.2 Contribution 1 : Protection à l'aide de tétraèdres de minuties irréversibles

5.2.1 Introduction

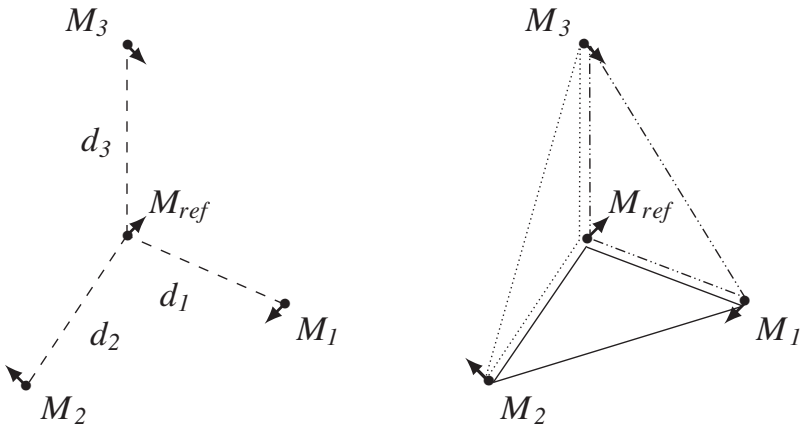
La proposition est une approche à deux facteurs d'authentification [Jain *et al.*, 2008, Maltoni *et al.*, 2009, Rathgeb et Uhl, 2011], où chaque utilisateur est censé présenter son échantillon d'empreinte digitale ainsi qu'une clé utilisateur spécifique à chaque vérification. Les points forts de ce type d'approche résident dans sa capacité à maintenir la précision de la correspondance en évitant à la fois les faux positifs (c'est-à-dire les fausses acceptations) et les faux négatifs (c'est-à-dire les faux rejets) et également dans son potentiel à fournir une grande robustesse. Dans le contexte du système proposé, pour chaque image d'empreinte digitale acquise lors de l'inscription, les propriétés des minuties originales sont extraites puis soumises à des transformations géométriques irréversibles, qui sont effectuées par le biais d'une clé utilisateur exprimé en tant que vecteur de longueur fixe généré aléatoirement. Le concept adopté consiste à transformer les tétraèdres de minuties formés à partir de la structure locale de chaque point caractéristique (minutie) sélectionné, puis à

déduire des propriétés significatives à partir des transformations résultantes et à les combiner pour produire le modèle d'empreinte digitale protégé. La phase de vérification passe évidemment par les mêmes séquences que pour l'inscription, ainsi que par un processus de mise en correspondance pour vérifier les images d'empreintes digitales de la requête. Le schéma proposé peut être résumé durant la phase de vérification par les étapes suivantes :

1. Sélection des minuties.
2. Transformation des tétraèdres de minuties.
3. Génération du modèle protégé.
4. Processus de correspondance.

5.2.2 Régime de protection proposé

5.2.2.1 Sélection des minuties



(a) (b)

FIGURE 5.1 – Structure extraite d'une zone de minuties : (a) les minuties voisins, (b) le tétraèdre de minuties formé.

En ne conservant que les terminaisons et les bifurcations des crêtes après la phase d'extraction, les caractéristiques dérivées de chaque image d'empreinte digitale sont présentées comme un ensemble de n points de minuties $\{M_i \mid i = 1, ..., n\}$, dans lequel le $i$ème point de minutie est exprimé par $(x_i\ ,\ y_i\ ,\ \theta_i)$, représentant les coordonnées dans l'espace cartésien et l'angle d'orientation qui est compris entre $[0, 2\pi)$. Il ne faut pas nier que des distorsions locales et des erreurs d'extraction de minuties peuvent souvent se produire et sont donc inévitables. Cependant, pour minimiser leur impact sur la performance de reconnaissance, la transformation par laquelle nous procédons concernera les zones locales plutôt que les minuties individuelles, car leur comportement permet de préserver la stabilité des caractéristiques même en cas de déformation de l'empreinte digitale.

Considérons par exemple les minuties illustrées dans la figure 5.1(a). À partir d'une minutie de référence M_{ref} (choisie au hasard dans l'ensemble des minuties extrait), nous déterminons d'abord les trois minuties les plus proches (M_1 ; M_2 ; M_3) de telle sorte que ces voisines soient ordonnées en fonction de leurs distances euclidiennes (respectivement ; d_1, d_2, d_3) par rapport à M_{ref}, où $d_1 < d_2 < d_3$. Sur la base de cet ordre, trois faces triangulaires différentes peuvent être formées à partir des minuties de référence, donnant ainsi naissance à une forme de tétraèdre de minuties (la figure 5.1(b) illustre cette situation). En fonction des faces triangulaires qui le compose, nous pouvons définir un tel tétraèdre de minuties comme suit :

$$Tetrahedron(M_{ref}) = \begin{cases} (M_{ref}; M_1; M_2); \\ (M_{ref}; M_1; M_3); \\ (M_{ref}; M_2; M_3). \end{cases} \tag{5.1}$$

De la même manière, à partir de chaque minutie sélectionnée, nous sommes censés identifier le tétraèdre de minuties correspondant. La figure 5.2 illustre quelques tétraèdres de minuties sélectionnés à partir d'une empreinte digitale.

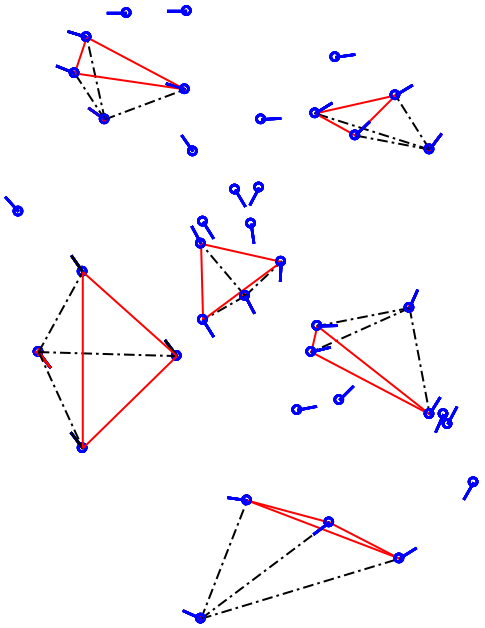


FIGURE 5.2 – Formes de tétraèdres sélectionnées à partir d'un modèle original d'empreinte digitale.

Pour réduire le risque de considérer une fausse minutie ou le problème d'une minutie authentique manquante, nous avons fait en sorte que chaque face triangulaire considérée concerne exactement deux minuties voisines différentes du tétraèdre sélectionné. Chaque face triangulaire subira indépendamment une transformation utilisant certains paramètres

spécifiques, et devra se référer au tétraèdre formé à partir de M_{ref}. De cette façon, en cas de minutie manquante ou de minutie parasite, nous pouvons être sûrs qu'une face triangulaire ne la concerne pas.

5.2.2.2 *Transformation des tétraèdres de minuties*

La transformation en question consiste à changer les faces définies d'un tétraèdre (sans considérer la base du tétraèdre) par le biais d'un ensemble de paramètres, en formant des nouvelles faces triangulaires indépendantes avec des caractéristiques différentes. Soit, par exemple, la face initiale du tétraèdre $(M_i\,;\,M_j\,;\,M_k)$. La première étape concerne le couple de minuties $(M_i,\ M_j)$ où M_i sera d'abord soumis à une rotation autour de la minutie M_j selon un angle défini α_i dans l'intervalle $[0, 2\pi)$. Ensuite, le point résultant devra être décalé vers un autre qui est aligné avec M_j selon une longueur définie L_i. Ces deux étapes principales du processus de transformation (c'est-à-dire la rotation et la translation) seront détaillées dans les deux sous-sections suivantes.

Sens de retation : Pour effectuer la rotation, il est tout d'abord nécessaire de définir une direction. Pour cela, nous faisons intervenir à la fois l'angle d'orientation θ_j du point M_j et l'orientation du côté β_{ji} comme illustré sur la figure 5.4. Le sens de rotation considéré est donc proportionnel à ces deux angles d'orientation et peut être déterminé selon la condition suivante :

$$Sens = \begin{cases} Positive & \text{Si } 0 \leqslant \theta_j - \beta_{ji} < \pi, \\ Negative & \text{Sinon.} \end{cases} \tag{5.2}$$

Où le sens positif fait référence ici au sens trigonométrique.

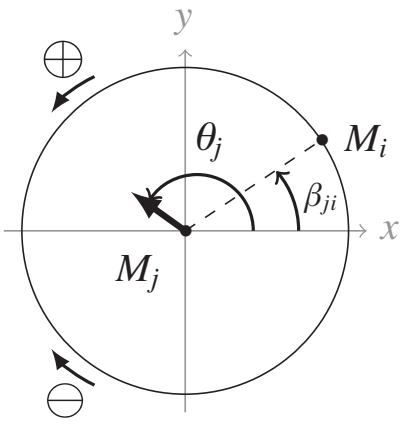


FIGURE 5.3 – Illustration des angles impliqués à la rotation.

Repositionnement des minuties : Supposons que le sens de rotation est positif, les nouvelles coordonnées (x_{R_i}, y_{R_i}) après l'application d'une rotation de α_i autour de M_j peuvent

simplement être obtenues comme suit :

$$\begin{bmatrix} x_{R_i} \\ y_{R_i} \end{bmatrix} = \begin{bmatrix} \cos(\alpha_i) & -\sin(\alpha_i) \\ \sin(\alpha_i) & \cos(\alpha_i) \end{bmatrix} \begin{bmatrix} x_{M_i} - x_{M_j} \\ y_{M_i} - y_{M_j} \end{bmatrix} + \begin{bmatrix} x_{M_j} \\ y_{M_j} \end{bmatrix} \tag{5.3}$$

Alors que pour le cas du sens négatif, il suffit d'appliquer la formule avec $-\alpha_i$. L'étape suivante consiste à décaler le point résultant R_i d'une distance L_i sur la ligne comprenant M_j et R_i. Le point transformé qui en résulte T_i (lié à M_i), peut être schématisée dans la figure 5.4 (cas où L_i est positif).

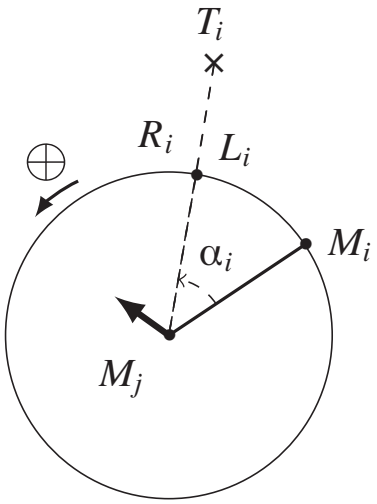


FIGURE 5.4 – Schématisation du point final T_i.

Le processus est réitéré pour les couples (M_j, M_k) et (M_k, M_i) en utilisant respectivement (α_j, L_j) et (α_k, L_k) afin de générer les deux autres points transformées T_j et T_k. De cette maniére, les minuties transformées peuvent occuper n'importe quelle position dans le plan 2D selon α et de L (les paramètres de la clé utilisateur), et permettant ainsi de former différentes formes de face triangulaire. Un exemple illustratif de la transformation est donné dans la figure 5.5.

Il convient de mentionner que même si la face triangulaire $(M_i\,;\, M_j\,;\, M_k)$ est transformée sur le plan des positions en $(T_i\,;\, T_j\,;\, T_k)$, les angles d'orientation initiaux restent les mêmes après transformation. Au cours de l'étape suivante, nous allons les combiner avec les nouvelles positions des minuties pour produire une représentation compacte.

5.2.2.3 *Génération du modèle protégé*

Pour représenter la face transformée du tétraèdre, nous avons utilisé les caractéristiques qui peuvent être dériver entre les minuties transformées car elles ne révèlent aucune indication sur les positions réelles des minuties originales. Ces caractéristiques ne doivent pas être sensibles aux variations de rotation et de translation en cas de déformation élastique de l'empreinte digitale. Elles doivent cependant être stables pour aboutir

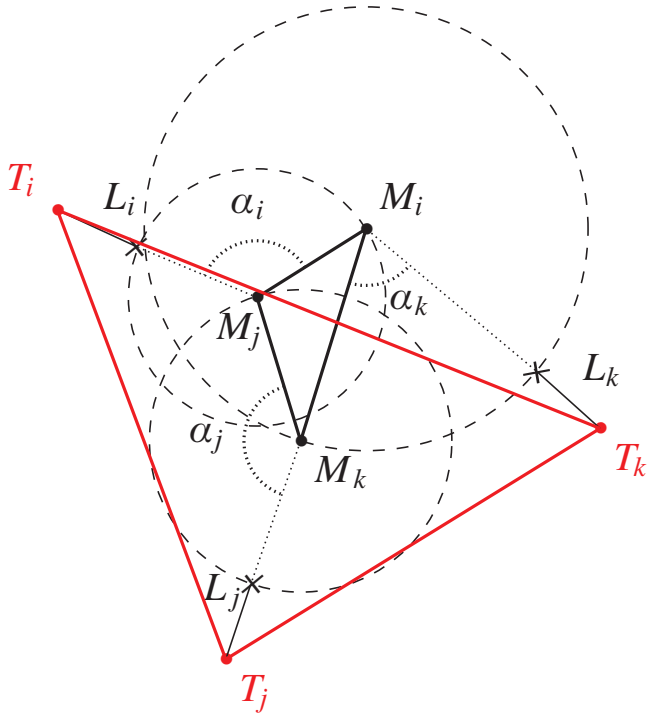


FIGURE 5.5 – Génération de la face de tétraèdre transformée $(T_i\,;\,T_j\,;\,T_k)$ à partir de $(M_i\,;\,M_j\,;\,M_k)$ selon, respectivement, $(\alpha_i,\,L_i)$, $(\alpha_j,\,L_j)$, $(\alpha_k,\,L_k)$.

à un vecteur fiable et significatif. Il existe plusieurs caractéristiques invariantes que l'on peut dériver d'une face triangulaire, telles que les longueurs des côtés, les angles internes, les angles entre les côtés et les minuties, ainsi qu'un certain nombre d'autres invariants [Farooq *et al.*, 2007]. En ce qui nous concerne, nous avons opté pour la hauteur du côté le plus large par rapport au minutie opposé, les trois côtés et les angles entre l'orientation de chaque minutie et le côté qui les relie, car ils sont pertinents et offrent un certain degré de discrimination. La figure 5.6 les schématise comme des caractéristiques de la face transformée du tétraèdre $(T_i\,;\,T_j\,;\,T_k)$.

— Les trois côtés S_{ij}, S_{jk}, S_{ki} ;

$$S_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{5.4}$$

— La hauteur depuis T_i (h_{ijk}) ; (Heron's Theorem(Annexe A.3))

$$h_{ijk} = \frac{2\sqrt{S(S - S_{ij})(S - S_{jk})(S - S_{ki})}}{S_{jk}} \tag{5.5}$$

où $S = \frac{1}{2}(S_{ij} + S_{jk} + S_{ki})$

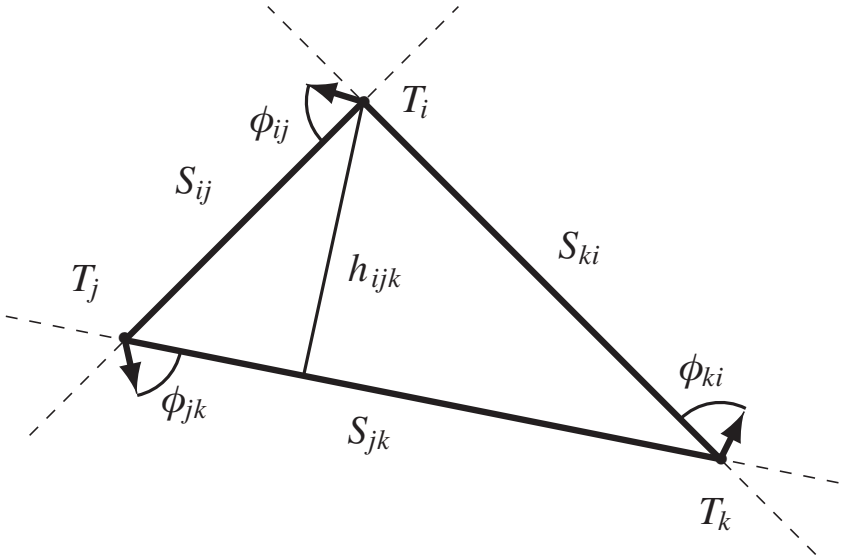


FIGURE 5.6 – Les caractéristiques extraites de la face d'un tétraèdre.

— Les angles entre les côtés et l'orientation des minuties ϕ_{ij}, ϕ_{jk}, ϕ_{ki} ;

$$\phi_{ij} = \min(\mid \beta_{ij} - \theta_i \mid, 360 - \mid \beta_{ij} - \theta_i \mid) \tag{5.6}$$

où $\beta_{ij} = arctan(\frac{y_i - y_j}{x_i - x_j})$ est l'angle d'orientation du côté S_{ij} par rapport à l'axe des abscisses.

Le vecteur caractéristique final obtenu à partir de la face transformée du tétraèdre $(T_i\,;\,T_j\,;\,T_k)$ peut être représenté dans l'ordre suivant :

$$FV(T_i, T_j, T_k) = (S_{ij}, S_{jk}, S_{ki}, h_{ijk}, \phi_{ij}, \phi_{jk}, \phi_{ki}) \tag{5.7}$$

En résumé, sur la base d'un tétraèdre de minuties défini à partir d'une seule minutie (minutie de référence) dans le domaine original, trois vecteurs caractéristiques différents de taille fixe (les 7 propriétés de la face triangulaire) sont dérivés dans l'espace transformé, qui représentent tous la structure locale de la minutie initiale. Ainsi, le traitement de toutes les minuties selon le processus décrit permet d'obtenir un total de $3 \times n$ vecteurs pour l'ensemble du modèle transformé, où n se réfère au nombre de minuties détectées dans l'empreinte digitale. L'algorithme 5.2 résume toutes les étapes suivies pour produire le modèle transformé.

5.2.2.4 *Processus de correspondance*

Comme nous l'avons déjà mentionné, chaque transformation géométrique dépend évidemment d'un vecteur particulier composé de paramètres générés aléatoirement, qui représente la clé utilisateur du système proposé. Elle peut être définie comme suit :

$$K^{user} = (\alpha_1^{user}, L_1^{user}, \alpha_2^{user}, L_2^{user}, \alpha_3^{user}, L_3^{user}) \tag{5.8}$$

Algorithme 5.1 Génération d'un modèle protégé selon le premier système proposé

Input : *Locations and orientations of all minutiae detected from a fingerprint image* $M = \{(x_u, y_u, \theta_u) \mid u = 1, ..., n\}$; *User key* $\{(\alpha_1, L_1), (\alpha_2, L_2), (\alpha_3, L_3)\}$;

Output : *A protected template* FV_j^i *where* $i = 1, ..., n$ *and* $j = 1, 2, 3$.

While $i \leq n$ **do**
// *Consider the selected minutia point as reference minutia.*
$M_{ref} = M_i$;
// *Definition of the minutiae tetrahedron corresponding to* M_{ref} *(5.2.2.1).*
$Tetrahedron(M_{ref}) =$
$\{\{M_{ref}; M_1; M_2\}, \{M_{ref}; M_1; M_3\}, \{M_{ref}; M_2; M_3\}\}$;
For $j = 1$ to 3
// *For each triangular face of the identified tetrahedron.*
For $k = 1$ to 3
// *For each minutia pair of the triangular face.*
Rotation direction()(5.2.2.2)
Minutiae Re-positioning(α_k, L_k)(5.2.2.2)
End For
// *Extraction of features from the new transformed minutiae*(5.2.2.3)
$FV_j^i(T_1, T_2, T_3) = (S_{12}, S_{23}, S_{31}, h_{123}, \phi_{12}, \phi_{23}, \phi_{31})$
End For
$i = i + 1$
End While
Retourner FV_j^i

Chaque utilisateur doit se voir attribuer un vecteur spécifique afin de préserver l'indépendance entre ses impressions et celles des autres utilisateurs (discrimination inter-classe). De plus, il sera nécessaire que ce vecteur soit stocké normalement dans une base de données, or ceci n'est pas très recommandé en raison de certaines mesures de protection de la vie privée. Pour remédier à ce problème, nous choisissons de les présenter dans un jeton externe lors de chaque vérification en plus de l'échantillon biométrique, tout comme la plupart des approches révocables de la littérature.

Lors de la phase de vérification, le même processus de transformation doit être appliqué, conduisant également à trois vecteurs de caractéristiques pour chaque minutie de l'empreinte digitale, exactement comme lors de la phase d'enrôlement, à moins que la différence entre le modèle enrôlé et le modèle de test résultant soit en termes du nombre de minuties acquises, donc le nombre de vecteurs diffère souvent même si les modèles proviennent de la même empreinte digitale.

Cependant, pour effectuer la correspondance entre un modèle enrôlé E^n et un modèle de test Q^m où leurs tailles sont, respectivement, $3 \times n$ et $3 \times m$, il est nécessaire d'identifier, pour chaque zone de minutie, qui est représentée par trois vecteurs dans le modèle enrôlé,

le correspondant le plus similaire dans le modèle de test. Ainsi, nous adoptons une stratégie de recherche exhaustive où les $3 \times n$ vecteurs de E^n sont comparés aux $3 \times m$ vecteurs de Q^m comme illustré dans la figure 5.7.

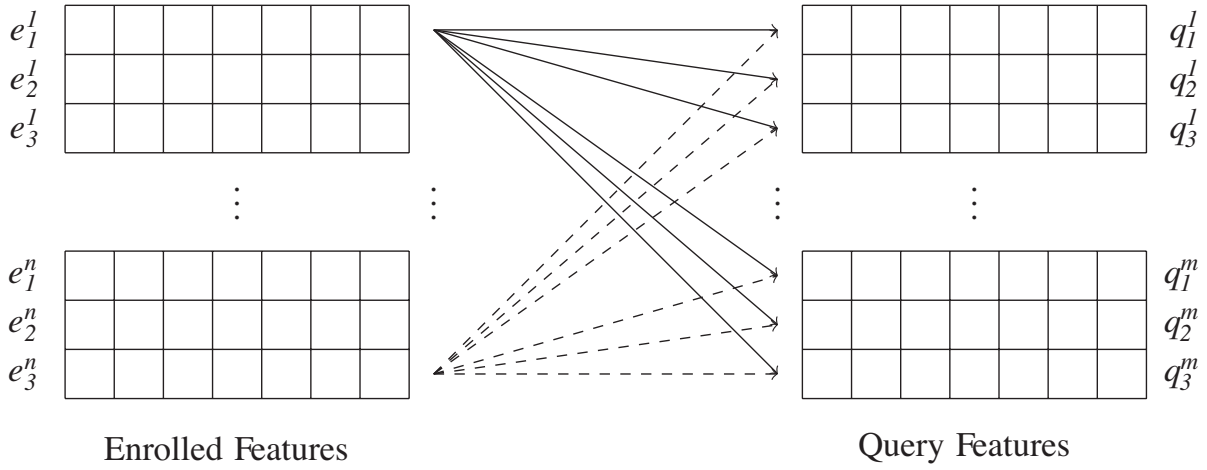


FIGURE 5.7 – Comparaison entre le modèle inscrit et le modèle de test.

Étant donné $E^i = [e_1^i, e_2^i, e_3^i]$ le vecteur contenant les représentations associées à la zone du $i$ème minutie du modèle enrôlé, et $Q^j = [q_1^j, q_2^j, q_3^j]$ le vecteur qui concerne la zone du $j$ème minutie du modèle de test. Le score du couple apparié c_{ij} lors de la comparaison de E^i et de Q^j est calculé à travers l'expression suivante :

$$c_{ij} = \min_{a,b=1,2,3}(D(e_a^i, q_b^j)), i = 1, ..., n \mid j = 1, ..., m \tag{5.9}$$

où D indique la distance Euclidienne entre e_a^i et q_b^j. Après avoir calculé tous les scores, ils sont rangés dans une matrice de scores $C = [c_{ij}]$ dont la taille est de $n \times m$. Ensuite, on enregistre dans le vecteur V_i la valeur minimale sur chaque ligne de C comme indiqué ci-dessous :

$$V_i = \min_j(c_{ij}), i = 1, ..., n \mid j = 1, ..., m \tag{5.10}$$

A ce stade, nous utilisons une autre mesure pour filtrer correctement les correspondances, qui consiste à faire la différence entre les vecteurs dont les distances Euclidiennes sont dans V_i. Cette mesure permet de fixer des seuils à la fois pour les distances et les angles qui constituent les vecteurs caractéristiques.

Étant donné deux vecteurs e_a et q_b dont la distance est incluse dans V_i, leur différence est la suivante :

$$\Delta FV_i = \mid e_a - q_b \mid \tag{5.11}$$

Le score final de correspondance est obtenu en quantifiant le nombre de V_i qui sont inférieurs à un seuil prédéfini t, et dont le vecteur de différence ΔFV_i est également

inférieur à un vecteur seuil prédéfini TV (chaque élément de ΔFV_i doit être inférieur à sa valeur seuil correspondante dans TV). Notons que TV est composé de deux valeurs, à savoir Δd et $\Delta\phi$, la première concerne les quatre premières distances de ΔFV et la seconde pour les trois autres angles. Le score final peut être normalisé comme suit :

$$Score = \frac{\sum_{i=1}^{n}((V_i \leqslant t), (\Delta FV_i \leqslant TV))}{\sqrt{n \times m}} \tag{5.12}$$

Où le score est compris entre [0, 1] et 1 est la correspondance parfaite.

5.2.3 Évaluation du régime proposé

Nous avons mené nos expérimentations sur trois bases de données d'empreintes digitales, à savoir, FVC2002 DB1, FVC2002 DB2 et FVC2002 DB3, qui sont fournies publiquement dans le cadre de la compétition internationale pour la vérification d'empreintes digitales [Maio *et al.*, 2002]. Les images sont acquises avec un capteur optique d'une résolution de 569 dpi, générant des images de 560 × 296 pixels. Pour chaque base de donnée, elle est composée de 800 images pour 100 individus et 8 échantillons par individu. En ce qui concerne l'extraction des coordonnées des minuties et leurs angles d'orientation, nous avons utilisé la version d'essai du logiciel commercial de reconnaissance d'empreintes digitales Verifinger 6.0 SDK de [Neurotechnology, 2010].

Pour obtenir une vision claire sur le comportement du système proposé en matière de performance, nous avons étudié sa réaction en termes de la précision de reconnaissance dans les deux scénarios les plus fréquents : la vérification dans les cas de *Different-key* et de *Stolen-key*. Le premier scénario consiste à attribuer à chaque utilisateur un vecteur aléatoire spécifique représentant la clé de l'utilisateur. Alors que pour le deuxième scénario, on suppose qu'un utilisateur légitime a perdu la clé, qui finira par tomber dans les mains d'un intrus, qui tentera alors d'effectuer une authentification en tant qu'un utilisateur légitime. La simulation de ce scénario consiste à attribuer la même clé utilisateur à tous les utilisateurs de la base de données. Pour réaliser toutes ces simulations sur FVC 2002, nous nous sommes basés sur le protocole original FVC (Fingerprint Verification Competition).

5.2.3.1 *Ajustement des paramètres*

Avant d'effectuer le processus de mise en correspondance, nous devions ajuster les valeurs des trois paramètres dont nous avons fait état dans la sous section 5.2.2.4 : Δd, $\Delta\phi$ et t. Pour cela, nous avons étudié l'impact de leurs différentes valeurs sur le système (déterminées empiriquement), puis nous avons défini la configuration à laquelle le système répond le mieux.

Pour cela, nous avons adopté la même stratégie que [Nandakumar *et al.*, 2007a, Ahmad *et al.*, 2011], à savoir le protocole *1 Vs 1* [Ferrara *et al.*, 2012]. En effet, nous avons utilisé dans cette expérience uniquement les deux premières images pour chaque empreinte digitale, en considérant la première comme le modèle orignal enrôlé et la deuxième comme

le modèle original de test. Il en résulte un total de 10000 scores de test, qui comprennent 9900($= 99 \times 100$) de scores imposteurs et 100 de scores authentiques.

En utilisant un vecteur aléatoire unique généré dans tous les tests, nous pouvons remarquer dans le tableau 5.1 que le paramétrage ($\Delta d = 10$, $\Delta\phi = 20$, $t = 20$) offre la meilleure performance parmi toutes les autres, que ce soit sur FVC2002 DB1 (1.93% d'EER) ou sur FVC2002 DB2 (1.25% d'EER). De ce fait, toutes les expériences suivantes seront réalisées en utilisant ce paramétrage.

Dataset/Parameter	Δd	$\Delta\phi$	t	$EER(\%)$
FVC2002 DB1	10	20	10	4.10
	10	**20**	**20**	**1.93**
	10	20	25	3.58
	20	20	30	4.56
	20	30	25	4.69
	20	30	30	3.88
FVC2002 DB2	10	20	10	3.00
	10	**20**	**20**	**1.25**
	10	20	25	2.40
	20	20	30	4.21
	20	30	25	3.57
	20	30	30	4.81

TABLE 5.1 – Valeurs EER selon différentes valeurs de seuil.

5.2.4 Précision de la vérification

Nous avons pu avoir un aperçu du comportement du système proposé dans les deux scénarios susmentionnés (*Different-key* et *Stolen-key*). Il s'est avéré que les résultats expérimentaux de la proposition dans le scénario *Différent-key* ont été parfaits pour toutes les empreintes digitales enregistrées sur FVC2002 DB1 et FVC2002 DB2, atteignant 0% d'EER. Nous pouvons le remarquer dans le tableau 5.6, où les valeurs résultantes du test de Kolmogorov Smirnov indiquent 1, qui signifie qu'il y a une séparation totale entre les distributions des scores légitime/imposteur. Alors que pour FVC2002 DB3, l'EER dans ce scénario a atteint 0.13% avec 0.9972 comme résultat du test de Kolmogorov Smirnov. Ce résultat est tout à fait normal car la qualité des images est moins bonne par rapport à celles de FVC2002 DB1 et DB2. d'après ces expériences, on peut déduire que l'utilisation de différents jetons rend les modèles d'empreintes digitales plus discriminants et plus précis malgré les distorsions infligées et les erreurs d'extraction des minuties.

Alors que pour le scénario *Stolen-key*, nous pouvons observer sur la figure 5.8 qui illustre les distributions légitimes/imposteurs quelques petits chevauchements pour

Database	Separability	K-S test
FVC2002 DB1	5.7187	1
FVC2002 DB2	5.6939	1
FVC2002 DB3	3.0356	0.9972

TABLE 5.2 – Valeurs obtenues à partir des tests de séparabilité et de *Kolmogorov-Smirnov* (K-S) sous le scénario *Différent-key*.

FVC2002 DB1, DB2 et DB3. Ceci est clairement visible sur la figure 5.9 qui illustre les graphiques FAR / FRR en fonction du seuil de notre proposition. Pour plus de clarté, nous affichons le comportement du système sur la figure 5.10, qui expose *the receiver operating characteristic*(ROC). D'après les graphiques, il est remarquable que les performances du système pour FVC2002 DB3 soient inférieures à celles de FVC2002 DB1 et DB2. Cette dégradation est certainement due aux taux de minuties erronées et manquantes. Comme la performance du schéma dépend fortement des résultats de l'extraction des minuties, il est évident que les résultats pour FVC2002 DB3 soient moins bons.

On peut affirmer que dans le cas du scénario *Stolen-key*, la performance est légèrement réduite. On peut le constater en comparant le tableau 5.6 et le tableau 5.7 en termes de séparabilité et de test de *Kolmogorov-Smirnov*. Cependant, nous sommes bien conscients que ces déficiences sont normales, sauf qu'elles ne doivent pas réduire considérablement les performances et restent donc tolérables.

Database	Separability	K-S test
FVC2002 DB1	3.4936	0.9221
FVC2002 DB2	3.5225	0.9364
FVC2002 DB3	2.1081	0.7690

TABLE 5.3 – Valeurs obtenues à partir des tests de séparabilité et de *Kolmogorov-Smirnov* (K-S) sous le scénario *Stolen-key*.

D'après les résultats obtenus en utilisant le protocole FVC original, tel que nous pouvons le voir dans le tableau 5.8, l'EER du schéma proposé dans ce scénario peut atteindre jusqu'à $3,90\%$, $3,17\%$ et $11,54\%$ respectivement pour, FVC2002 DB1, DB2 et DB3. Nous pensons que, dans le cadre du scénario d'attaque *Stolen-key*, les performances demeurent tout à fait acceptables, surtout lorsqu'elles sont comparées à certaines méthodes de la littérature qui utilisent à peu près le même protocole et la même stratégie d'évaluation (voir tableau 5.8).

Par ailleurs, il convient de noter que les performances varient en fonction du vecteur aléatoire utilisé. Ceci est en fait dû aux paramètres de rotation et de translation appliqués,

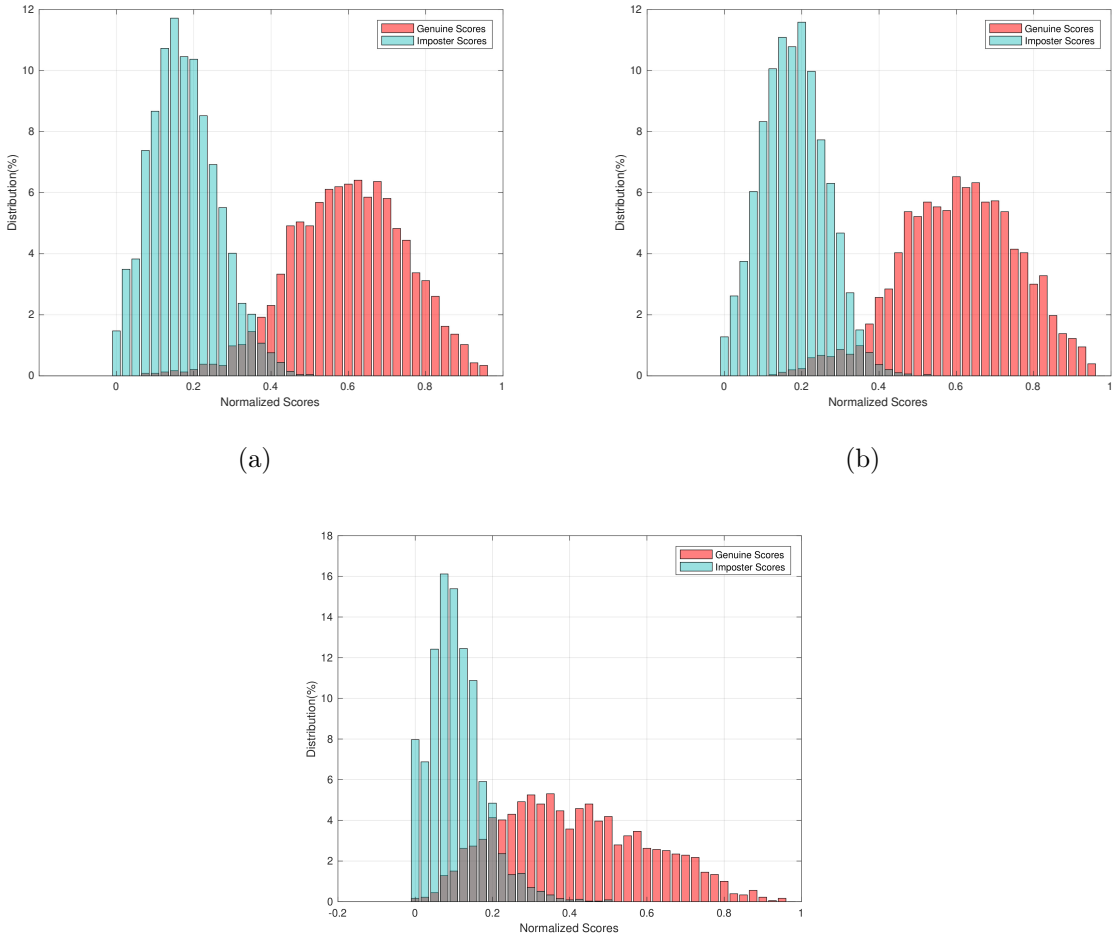


FIGURE 5.8 – Les distributions des scores légitimes/imposteurs selon le scénario *Stolen-key* pour **(a)** FVC2002 DB1, **(b)** FVC2002 DB2, **(c)** FVC2002 DB3.

ainsi qu'à la façon dont les erreurs et les imprécisions générées pendant la phase d'acquisition sont traitées. C'est tout à fait naturel puisque les transformations sont basées sur la géométrie.

5.2.5 Conformité aux exigences de révocabilité, diversité et non-inversibilité

5.2.5.1 Révocabilité

La révocabilité est un aspect indispensable pour tout système biométrique révocable. Elle permet de révoquer un modèle utilisateur qui a été intercepté ou compromis suite à une attaque, de telle sorte que les modèles compromis et révoqués soient différents et non corrélés, même s'ils proviennent des mêmes données biométriques. Dans le contexte de la biométrie révocable, la construction du nouveau modèle protégé n'est effectuée qu'en

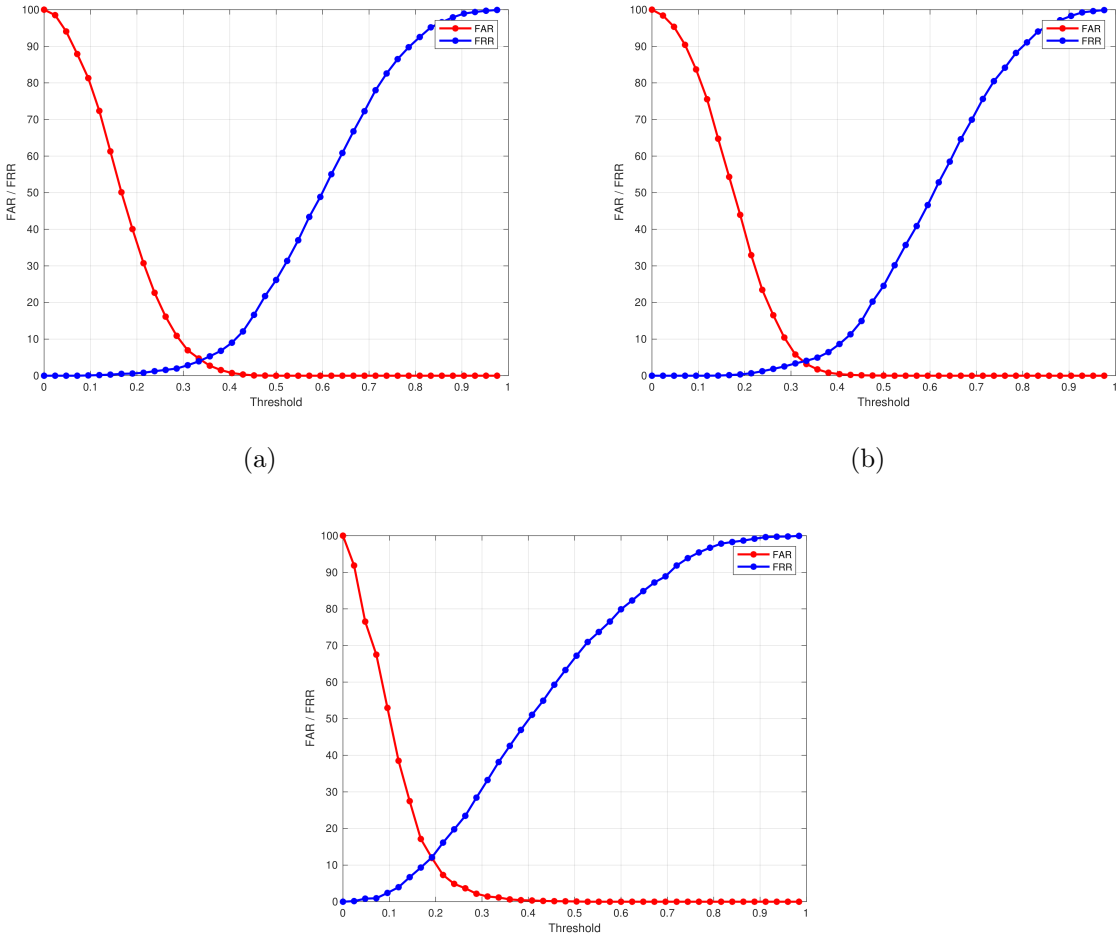


FIGURE 5.9 – Courbes FAR/FRR en fonction du seuil sous le scénario *Stolen-key* pour **(a)** FVC2002 DB1, **(b)** FVC2002 DB2, **(c)** FVC2002 DB3.

utilisant une nouvelle clé utilisateur. Pour notre approche, il suffit de définir différentes valeurs de rotation et de translation pour produire un nouveau modèle. Il convient de noter que chaque ensemble spécifique de paramètres donne lieu à une déformation différente en termes de position. La figure 5.11 illustre les multiples déformations dans l'espace transformé à partir d'une seule face de tétraèdre en utilisant différents vecteurs générés aléatoirement.

Dans le cas d'une attaque qui vise à compromettre un modèle, notre système doit être capable de générer un nouveau modèle protégé qui ne correspond pas au modèle compromis en utilisant la même empreinte digitale originale. Pour évaluer notre système en termes de révocabilité, nous avons étudié la réaction du système en cas d'attaque par un modèle révoqué, un scénario dans lequel un attaquant tente d'accéder au système en utilisant un modèle compromis. Pour cela, nous avons considéré deux scénarios d'attaque :

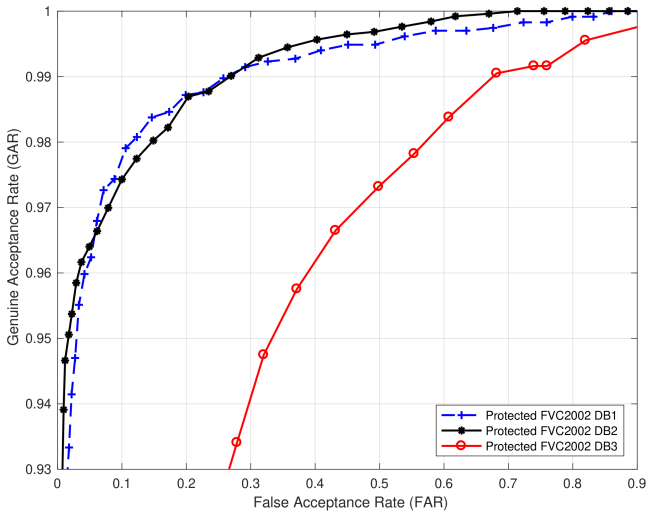


FIGURE 5.10 – Courbes ROC selon le scénario *Stolen-key* pour FVC2002 DB1, DB2 et DB3.

Method/Dataset	DB1	DB2	DB3
Ahn et al. [Ahn *et al.*, 2008]	7.18	3.61	11.80
Ferrara et al. [Ferrara *et al.*, 2012]	3.33	1.76	7.78
P. Das et al. [Das *et al.*, 2012]	4	-	-
Wang and Hu [Wang et Hu, 2016]	4	3	8.50
Lahmidi and al. [Lahmidi *et al.*, 2021]	3.90	3.17	11.54

TABLE 5.4 – Comparaison entre le schéma proposé et certaines méthodes de l'état de l'art en termes de EER(%) sous le scénario *Stolen-key* et selon le protocole FVC sur FVC2002 DB1, DB2 et DB3.

Note : "-" dénote la non-disponibilité des données.

— **Scénario - I** : Suite à une attaque visant la base de données, un modèle protégé est compromis, et un nouveau modèle est reproduit à partir de la même image d'empreinte digitale en utilisant un nouvel ensemble de paramètres. Dans ce scénario, l'attaquant tente de s'authentifier avec le modèle compromis ;

— **Scénario - II** : Suite à une attaque visant la base de données, un modèle protégé est compromis, et un nouveau modèle est reproduit à partir d'une image d'empreinte digitale différente du même doigt en utilisant un nouvel ensemble de paramètres. Dans ce scénario, l'attaquant tente de s'authentifier avec le modèle compromis.

Sur la base des résultats obtenus suite à la simulation de l'attaque *the revoked template*

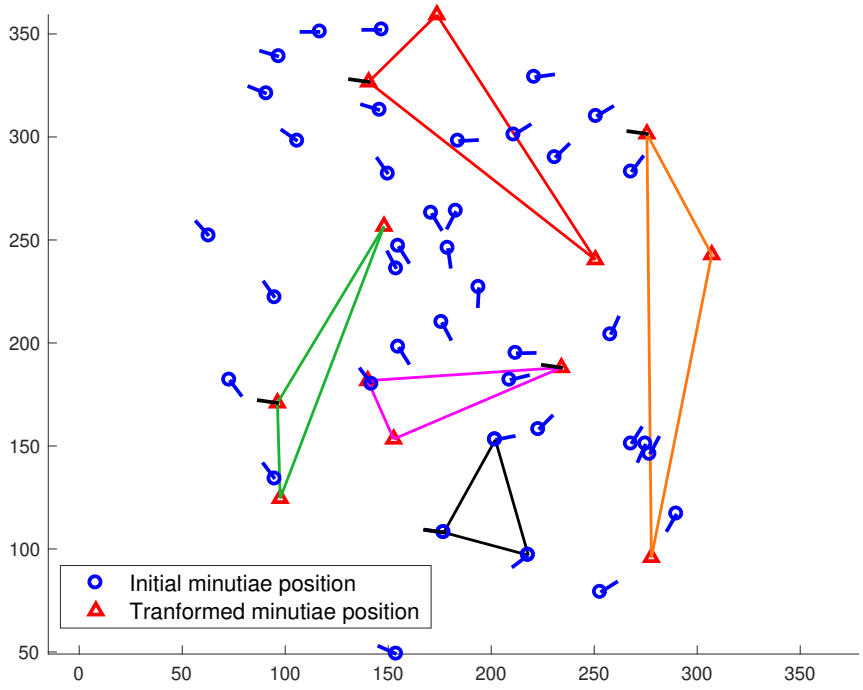


FIGURE 5.11 – Multiples déformations d'une face de tétraèdre en utilisant différents vecteurs générés de façon aléatoire.

attack (Tableau 5.5), qui a été réalisée sur FVC 2002 DB1, DB2 et DB3. Nous pouvons dire que l'approche proposée garantit une grande séparabilité entre les anciens modèles compromis et les nouveaux modèles, même s'ils proviennent de la même image d'empreinte digitale. Cela signifie que notre proposition est parfaitement conforme à la propriété de révocabilité.

Database	Scenario - I	Scenario - II
FVC2002 DB1	0.0%	0.0%
FVC2002 DB2	0.0%	0.0%
FVC2002 DB3	0.0%	0.0%

TABLE 5.5 – Vérification sous *the revoked template attack* sur FVC2002 DB1, DB2 et DB3.

5.2.5.2 Diversité

La diversité, est le fait de produire plusieurs modèles protégés pour différentes applications à partir des mêmes données biométriques, de manière à ce que les modèles générés

ne soient pas corrélés entre eux. Pour répondre à cette exigence, chaque système biométrique doit être capable de garantir une grande dissimilarité entre les modèles compromis et révoqués.

Pour évaluer l'approche que nous proposons en termes de diversité, plusieurs modèles protégés ont été dérivés de chaque image d'empreinte digitale, de sorte que chaque génération a impliqué un ensemble différent de clés utilisateur. Ces modèles ont ensuite été comparés en termes de similarité. Pour cette simulation de scénario, nous avons considéré deux systèmes, chacun d'entre eux traitant une gamme spécifique de valeurs de paramètres. Pour le premier système, nous avons choisi au hasard des valeurs dans les plages suivantes : $\alpha_i \in [0, 30]$, $\alpha_j \in [0, 30]$, $\alpha_k \in [0, 30]$, $Li \in [0, 20]$, $Lj \in [0, 20]$, $Lk \in [0, 20]$. Tandis que pour le deuxième, $\alpha_i \in [60, 90]$, $\alpha_j \in [60, 90]$, $\alpha_k \in [60, 90]$, $Li \in [40, 60]$, $Lj \in [40, 60]$, $Lk \in [40, 60]$. Le principe de la simulation était d'obtenir la distribution pseudo-légitime en comparant pour chaque image d'empreinte digitale les gabarits générés suivant le système 1 avec ceux obtenus par le système 2.

Sur la base des distributions schématiques de la figure 5.12, nous pouvons remarquer pour FVC2002 DB1 et DB2 une grande séparation graphique entre la distribution des scores pseudo-légitimes obtenus à partir de la simulation précédente et la distribution des scores légitimes. Cela prouve que même si plusieurs modèles protégés sont dérivés d'un même doigt, ils peuvent être complètement indépendants en leur sein. Cela nous donne la possibilité de les étendre à diverses applications sans nous soucier du problème de la correspondance croisée (*Linkage attack*) [Ali et Prakash, 2017], puisqu'il suffit d'utiliser un ensemble de clés utilisateur différent. Alors que pour FVC2002 DB3, un léger chevauchement entre les deux distributions peut être observé, cela peut s'expliquer simplement par la mauvaise qualité des images d'empreintes digitales.

5.2.5.3 Non-inversibilité

Le principal critère auquel doit répondre une approche idéale des modèles biométriques est la non-inversibilité. Cette propriété signifie qu'il doit être extrêmement difficile de reconstruire les modèles biométriques originaux à partir des modèles transformés.

En supposant dans notre cas que toutes les étapes du système proposé sont publiques et bien connues, et que les modèles protégés stockés ont été interceptés. Étant donné que les modèles stockés sont simplement constitués de caractéristiques locales dérivées de la transformation des tétraèdres de minuties, et qu'en cas de compromission, un attaquant sera normalement en mesure de retracer facilement leurs formes. Cependant, leurs positions dans l'espace transformé demeurent inconnues. En pratique, il n'y a aucun moyen de révéler l'emplacement des minuties transformées car les caractéristiques considérées représentent les propriétés qui peuvent être formées entre les minuties transformées (par exemple, les distances, les angles) et non les positions. En outre, bien que certains minuties soient impliqués dans plus d'un tétraèdre, il n'est pas du tout évident d'en déduire des données utiles pour découvrir la forme originale du tétraèdre.

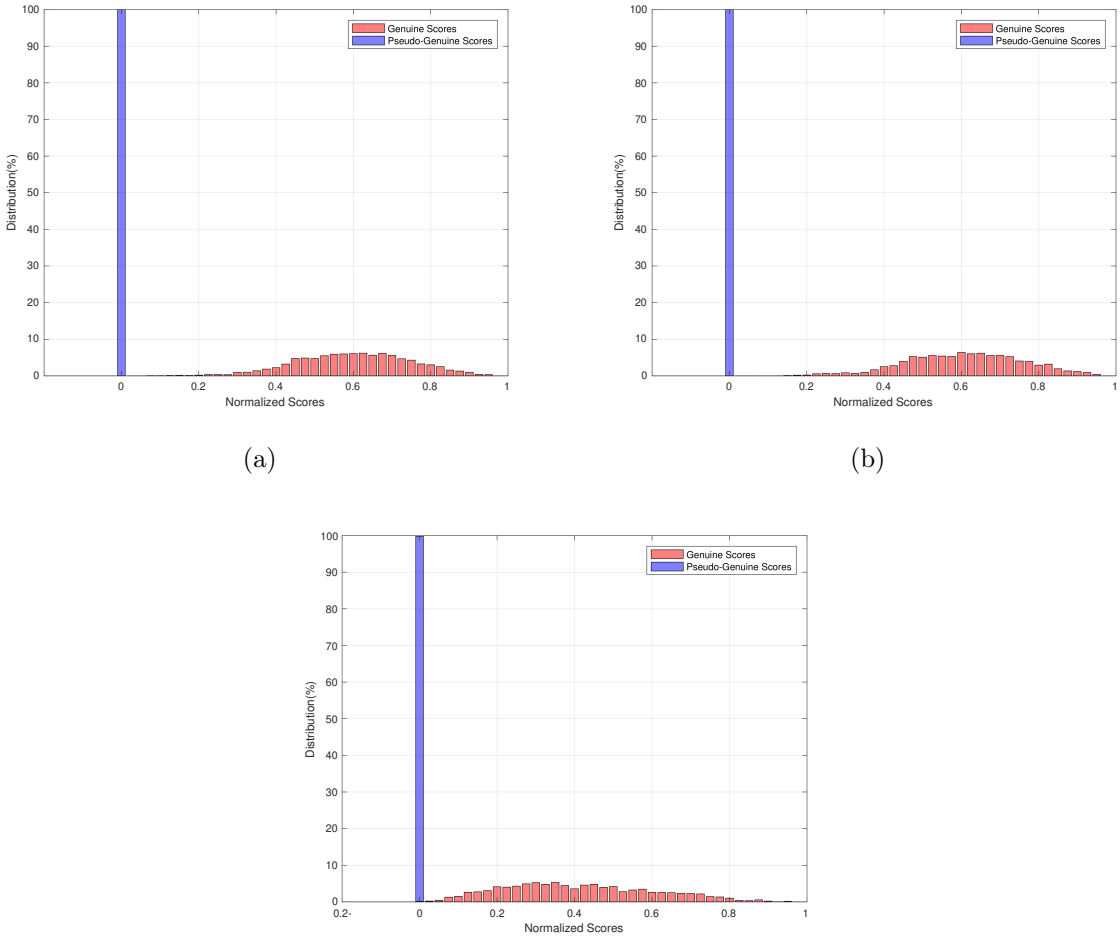


FIGURE 5.12 – Distributions des scores légitimes/pseudo-légitimes pour **(a)** FVC2002 DB1, **(b)** FVC2002 DB2, (c) FVC2002 DB3

Supposons à présent que le jeton (clé utilisateur) soit également en possession de l'attaquant, en utilisant les formes des faces du tétraèdre transformées et les paramètres de rotation et de translation interceptés, une attaque par force brute sur l'orientation de la longueur ajoutée, le centre et la direction de rotation pour chaque minutie est réalisable pour obtenir une représentation approximative des formes originales, sauf que ce résultat ne peut être estimé dans un temps raisonnable, et pour un nombre limité d'essais de vérification possibles, en raison du nombre exponentiel de combinaisons possibles.

À partir d'une face de tétraèdre transformée, chaque minutie a normalement 360 possibilités de revenir au point R sachant les valeurs de translation. Ensuite, il faut déterminer un centre de rotation pour refaire la rotation en sens inverse. Pour cela, il faut tester tous les points possibles le long de la ligne incluant les points caractéristiques transformés et R. Une fois ce point déterminé, il reste deux possibilités pour définir la direction de la ro-

tation. Tous ce processus concerne une seule minutie, le traitement de toutes les minuties transformées est donc très compliqué et prend beaucoup de temps. Par exemple, en supposant que nous allons tester 50 points sur la ligne comprenant les minuties transformées et R, le nombre de combinaisons possibles est d'environ $(360 \times 2 \times 50)^n$. Après tout, les représentations proposées ne concerneront que les formes des tétraèdres initiaux et non leurs positions réelles dans le domaine original.

5.3 Contribution 2 : Nouvelle méthodologie basée sur les spécifications d'un système non protégé

5.3.1 Introduction

Dans cette contribution, la méthodologie de conception que nous avons menée n'est pas celle que l'on retrouve habituellement dans les travaux standards de protection de gabarits d'empreintes digitales. En effet, nous avons adopté une réflexion à partir d'un processus bien connu de correspondance d'empreintes digitales, à savoir celui proposé dans [Jiang et Yau, 2000], dont le but est de trouver la correspondance entre deux ensembles de minuties sans l'implication des caractéristiques globales. Le concept que nous avons proposé consiste à convertir le modèle d'empreinte digitale original à l'aide d'une transformation non inversible, de sorte que le processus de comparaison des empreintes digitales reste fonctionnel même dans le domaine transformé. C'est pour cette raison que nous avons veillé à préserver la même qualité des minuties. Nous proposons donc une technique qui provoque une certaine perturbation en termes d'emplacement et d'orientation des minuties originales, ce qui donne lieu à un nouveau modèle d'empreinte digitale. Pour être cohérent avec le processus de comparaison appliqué, le modèle d'empreinte digitale comparé ne doit pas contenir des points singuliers. Nous avons donc pris soin de respecter cette exigence en exploitant les points singuliers uniquement pour effectuer la transformation et en les excluant de l'ensemble des minuties transformées. Dans les deux sous-sections suivantes (5.3.2,5.3.3), nous présentons d'abord le système de vérification d'empreintes digitales non protégé sur lequel on s'est basé, puis nous décrivons le schéma de protection proposé.

5.3.2 Système de vérification non protégé

Au cours des dernières décennies, le processus de mise en correspondance a toujours été considéré comme une tâche délicate dans les systèmes d'authentification par empreintes digitales, en particulier lorsqu'on est confronté au problème critique des variations intra-classes (les acquisitions d'un même doigt subissent un degré élevé de variabilité). Selon [Maltoni *et al.*, 2009], la comparaison d'empreintes digitales peut être généralement classée en trois grandes familles :

- Correspondance basée sur la corrélation : C'est un processus qui consiste à calculer

la corrélation entre les valeurs des pixels de l'image d'empreintes digitales de la requête et celles de la référence pour différents alignements.

- Correspondance basée sur les minuties : C'est la technique la plus utilisée dans la reconnaissance des empreintes digitales. Elle est principalement basée sur les caractéristiques des points caractéristiques (c'est-à-dire leur emplacement et leur orientation). L'objectif est d'atteindre l'alignement entre deux modèles d'empreintes digitales qui donne lieu au nombre maximal de paires de points caractéristiques.

- Correspondance basée sur des caractéristiques non minutieuses : Cette technique repose sur la comparaison d'autres caractéristiques du motif de la crête de l'empreinte digitale qui peuvent être mieux extraites que les caractéristiques minutieuses, telles que l'orientation locale, la fréquence, la forme de la crête et les informations de texture.

Dans ce travail, nous avons abordé la catégorie des correspondances basées sur les minuties, qui peuvent à leur tour être divisées en correspondances locales et globales. La correspondance locale des minuties permet de comparer deux empreintes digitales sur la base des structures locales des minuties. Celles-ci sont définies par rapport au voisinage des minuties (souvent en termes de distances Euclidiennes). Le principe est de se servir des propriétés qui peuvent être extraites dans cette zone et qui sont invariantes aux transformations globales (par exemple, la translation, la rotation). Alors que la correspondance globale des minuties reflète l'unicité des empreintes digitales comparées. L'idée est de réaliser l'alignement des minuties grâce à des caractéristiques globales telles que les points singuliers ou les champs d'orientation.

Une approche particulière de correspondance par minuties a été introduite par Jiang et Yau [Jiang et Yau, 2000], où les deux types de correspondance par minuties sont impliqués (structures locales et globales). L'idée est de rechercher d'abord la meilleure paire de minuties similaires entre deux empreintes digitales à l'aide d'un descripteur de minuties local, basé sur des caractéristiques invariantes (distances et angles) extraites du voisinage des minuties. Le but est d'identifier les structures locales les plus similaires. Ensuite, une consolidation globale est élaborée qui consiste à faire un alignement en fonction des minuties sélectionnées dans chaque empreinte digitale.

Étant donné que chaque minutie acquise $\{M_i \mid i = 1, ..., n\}$ est présenté en tant que vecteur caractéristique $F_i = (x_i, y_j, \theta_i, t_i)$, où (x, y) sont les coordonnées sur le plan cartésien, θ est l'angle d'orientation de la minutie et t fait référence au type de minutie (terminaison ou bifurcation). La première étape de cette comparaison de minuties consiste à définir la structure locale de chaque minutie en impliquant ses deux plus proches voisins, puis à extraire les caractéristiques invariantes en rotation et en translation qui peuvent être formées dans le voisinage. En considérant M_j et M_k comme les deux plus proches voisins de M_i (où M_j est le premier plus proche de M_i et M_k le deuxième). Le vecteur de caractéristiques locales résultant FV_i lié à M_i peut être écrit comme suit :

$$FV_i = (d_{ij}, d_{ik}, \alpha_{ij}, \alpha_{ik}, \phi_{ij}, \phi_{ik}, n_{ij}, n_{ik}, t_i, t_j, t_k) \tag{5.13}$$

où d_{ij} est la distance euclidienne entre M_i et M_j, α_{ij} représente la différence d'orientation entre θ_i et θ_j, ϕ_{ij} correspond à la différence d'orientation entre θ_i et l'orientation de l'arête reliant M_i et M_j. n_{ij} fait référence au nombre d'arêtes entre M_i et M_j, et t_i est le type de minutie de M_i (La figure 5.13 illustre certaines de ces caractéristiques).

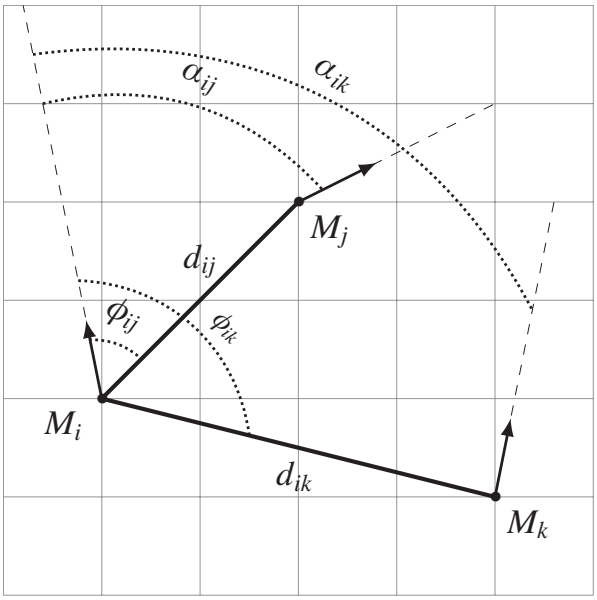


FIGURE 5.13 – Propriétés de la structure locale impliquée dans [Jiang et Yau, 2000].

L'étape suivante du processus vise à trouver la paire de minuties la plus similaire entre les modèles d'empreintes digitales de référence et de test. Cela se fait par le biais d'une stratégie de recherche exhaustive où tous les vecteurs caractéristiques locaux extraits du modèle de référence sont comparés à ceux du modèle de test. Les paires de structures les mieux appariées obtenues sont ensuite utilisées pour effectuer un alignement entre les deux empreintes digitales. Tous les minuties restantes seront alignées sur la base de la paire de minuties en les convertissant dans le système de coordonnées polaires. Enfin, un score est calculé en tenant compte des apports des deux étapes de comparaison. Dans ce travail, nous avons étudié le cas d'un système de vérification non protégé en utilisant notre propre configuration de cet algorithme de comparaison par minuties (décrit dans la sous-section 5.3.4.1).

5.3.3 Régime de protection proposé

La nature protectrice des modèles d'empreintes digitales proposés dans ce travail est une forme de désordre provoqué à l'emplacement et à l'orientation des minuties, principalement par le biais d'une clé spécifique que l'utilisateur est censé fournir lors de l'authentification. Le modèle généré est un ensemble de minuties avec différentes caractéristiques

qui est en fait conçu pour être révocable et non inversible, c'est-à-dire que le système est capable de générer plusieurs modèles protégés à partir de la même empreinte digitale en utilisant différentes clés de sorte qu'il n'y a pas de corrélation entre les modèles. En outre, lorsque le modèle généré est intercepté, il est presque impossible de révéler des données significatives. L'idée principale de la technique de protection proposée est de construire quatre groupes différents de minuties, chaque groupe subissant une transformation particulière en fonction des paramètres de la clé utilisateur. Dans ce qui suit, nous décrivons le processus de génération d'un modèle sécurisé à l'aide de notre proposition qui comprend trois étapes essentielles :

(i) Extraction des caractéristiques invariantes.

(ii) Désignation du groupe d'appartenance.

(iii) transformation des minutiae.

Toutes les étapes nécessaires à la construction du modèle protégé sont présentées dans l'algorithme 5.2.

5.3.3.1 *Extraction des caractéristiques invariantes*

La première étape consiste à dériver deux caractéristiques invariantes pour chaque minutie acquise, formées entre la minutie concernée et le point singulier principal (appelé SP). La première propriété à déterminer est la position des minuties par rapport à SP. A cette fin, l'espace bidimensionnel est partitionné en quatre zones selon trois rayons r_1, r_2, r_3 (fixés empiriquement) définis autour de SP. Les zones formées sont étiquetées respectivement 00, 01, 10 et 11 comme le montre la figure 5.14. Pour chaque minutie M_i, les deux bits qui font référence à la zone à laquelle appartient la minutie en question sont affectés à P_i. La deuxième propriété concerne l'angle β_i entre l'orientation de la minutie M_i et l'orientation de l'arête reliant la position de M_i et SP dans le sens inverse des aiguilles d'une montre comme le montre la figure 5.15. En fonction de la valeur de l'angle de β_i, deux bits sont affectés à A_i comme décrit ci-dessous.

$$A_i = \begin{cases} 00 & \text{if} \quad \beta_i < \dfrac{\pi}{2}, \\ 01 & \text{if} \quad \dfrac{\pi}{2} \leqslant \beta_i < \pi, \\ 10 & \text{if} \quad \pi \leqslant \beta_i < \dfrac{3\pi}{2}, \\ 11 & \text{if} \quad \dfrac{3\pi}{2} \leqslant \beta_i \end{cases} \tag{5.14}$$

5.3.3.2 *Désignation du groupe d'appartenance*

Durant cette phase, chaque minutie M_i est attribuée à l'un des quatre groupes définis précédemment en fonction des valeurs de P_i et A_i. Pour établir l'attribution, le système effectue une opération OU exclusif (XOR) entre P_i et A_i conduisant à deux nouveaux bits.

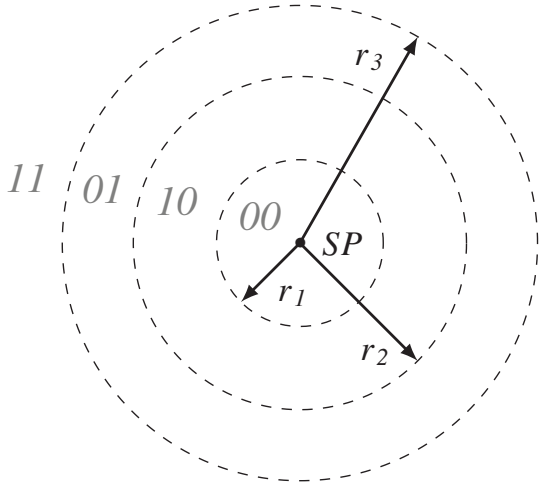


FIGURE 5.14 – Découpage de l'espace bi-dimensionnel en quatre zones selon r_1, r_2, r_3 et SP.

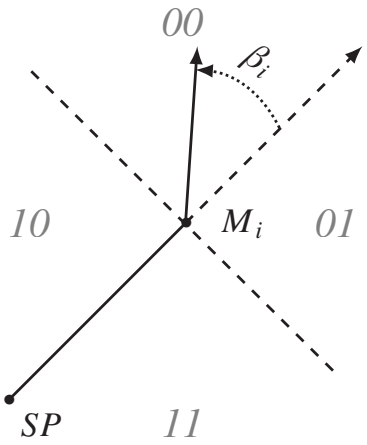


FIGURE 5.15 – Représentation de l'angle β_i entre l'orientation de la minutie M_i et l'orientation de l'arête reliant M_i et SP en rotation anti-horaire.

Les bits résultants sont chargés de décider à quel groupe elle appartiendra et donc quels paramètres de la transformation seront appliqués. Considérant que les quatre groupes prédéfinis sont préalablement référencés respectivement comme 00, 01, 10 et 11. Chaque minutie est assignée au groupe de rattachement dont la référence coïncide avec le résultat du OU exclusif (XOR). Ainsi, il y aura quatre groupes de minuties différentes. Ce mode de répartition empêche en fait d'appliquer la même transformation sur des minuties qui ont des caractéristiques proches, provoquant de ce fait un désordre dont il sera dur à inverser.

5.3.3.3 Transformation des minuties

Après avoir assigné chaque minutie à son groupe d'appartenance, chaque groupe est exposé à une transformation spécifique, c'est-à-dire que toutes les minuties du groupe

seront soumises à la même transformation, qui est effectuée essentiellement à l'aide d'un ensemble de paramètres qui est censé être représenté par une clé utilisateur. La clé utilisateur est constituée de quatre paires de paramètres (Equation 5.15) référencés aussi respectivement comme 00, 01, 10 et 11. Chaque groupe est concerné par la transformation dont l'étiquette du groupe est la même que celle de la paire de paramètres de la clé utilisateur.

$$Key^{user} = \{(\delta_k, \lambda_k)^{user}\}_{k=1}^{4} \tag{5.15}$$

Supposons que le couple (δ_k,λ_k) correspond aux paramètres de transformation du groupe auquel appartient M_i. La transformation de $M_i(x_i, y_i)$ consiste à effectuer une rotation autour du point singulier principal $SP(x_{SP}, y_{SP})$ avec un angle δ_k dans le sens inverse des aiguilles d'une montre (Equation 5.16). Le point résultant (x_i^{Rot}, y_i^{Rot}) sera ensuite utilisé pour construire une image par homothétie en prenant comme centre $SP(x_{SP}, y_{SP})$ et λ_k comme rapport de l'homothétie (Equation 5.17). La position de la nouvelle minutie M_i' est représentée par les coordonnées cartésiennes (x_i', y_i').

Suite à cette transformation, toutes les minuties de l'empreinte digitale acquièrent de nouvelles positions dans l'espace bi-dimensionnel, tandis qu'en termes d'orientation, leurs positions initiales sont ajoutées à l'angle δ. Au final, les quatre groupes sont concaténés pour donner naissance à un nouvel ensemble de minuties représentant le modèle protégé.

$$\begin{bmatrix} x_i^{Rot} \\ y_i^{Rot} \end{bmatrix} = \begin{bmatrix} \cos(\delta_k) & -\sin(\delta_k) \\ \sin(\delta_k) & \cos(\delta_k) \end{bmatrix} \begin{bmatrix} x_{M_i} - x_{SP} \\ y_{M_i} - y_{SP} \end{bmatrix} + \begin{bmatrix} x_{SP} \\ y_{SP} \end{bmatrix} \tag{5.16}$$

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \lambda_k \cdot \begin{bmatrix} x_i^{Rot} - x_{SP} \\ y_i^{Rot} - y_{SP} \end{bmatrix} + \begin{bmatrix} x_{SP} \\ y_{SP} \end{bmatrix} \tag{5.17}$$

5.3.4 Évaluation du régime proposé

5.3.4.1 Stratégie d'évaluation

La stratégie d'évaluation suivie dans la présente étude est similaire à celle utilisée dans [Ahmad *et al.*, 2011, Nandakumar *et al.*, 2007a, Xi et Hu, 2009], qui consiste à comparer le gabarit généré à partir de la première empreinte avec celui construit à partir de la deuxième empreinte du même doigt pour obtenir le FRR, tandis que le gabarit de la première empreinte est comparé avec celui construit à partir de la première empreinte des autres doigts lorsqu'il s'agit du FAR. Selon la stratégie utilisée sur FVC 2002 DB1 et DB2, un total de 100 scores authentiques et 9900 (99×100) imposteurs sont obtenus pour chaque base de données. Comme décrit précédemment, la transformation des minuties effectuée par le système proposé est principalement basée sur les points singuliers. Nous avons donc extrait pour chaque empreinte digitale uniquement le point singulier le plus proche du centre de l'image, considérant que seules les empreintes digitales où des points singuliers sont détectés sont retenues.

Algorithme 5.2 Génération d'un modèle protégé selon le deuxième système proposé

Input : *Minutiae locations and orientations from a fingerprint image* $M_i = \{(x_i, y_i, \theta_i) \mid i = 1, ..., n\}$; *User key* $\{(\delta_1, \lambda_1), (\delta_2, \lambda_2), (\delta_3, \lambda_3), (\delta_4, \lambda_4)\}$; Singular Point $SP = \{(x_{SP}, y_{SP})\}$;
Output : *Modified minutiae locations and orientations* $M'_i = \{(x'_i, y'_i, \theta'_i) \mid i = 1, ..., n\}$;

// *initialize the minutia counter*
$i \leftarrow 1$;
// *Declare four empty structures*
$Grp_1 \leftarrow []$; $Grp_2 \leftarrow []$; $Grp_3 \leftarrow []$; $Grp_4 \leftarrow []$;
Tantque $i \leq n$ **Faire**
    // *Invariant features extraction*(5.3.3.1)
    // *The label corresponding to the zone where* M_i *resides with respect to SP*
    $P_i \leftarrow Label(Position(M_i, SP))$;
    // *The label corresponding to the angle between the orientation of* M_i *and the orientation of the edge linking* M_i *and SP in counter-clockwise rotation*
    $A_i \leftarrow Label(Angle(M_i, SP))$; (Equation 5.14)
    $B_i \leftarrow A_i \oplus P_i$;
    // *Home group designation* (5.3.3.2)
    Si $B_i == 00$ **Alors**
        $Affect(M_i, Grp_1)$; // *Assign* M_i *to* Grp_1
    Sinon Si $B_i == 01$ **Alors**
        $Affect(M_i, Grp_2)$;
    Sinon Si $B_i == 10$ **Alors**
        $Affect(M_i, Grp_3)$;
    Sinon
        $Affect(M_i, Grp_4)$;
    Finsi
    $i \leftarrow i + 1$;
Fin Tantque
// *Minutiae transformation* (5.3.3.3)
// *initialize a counter k to browse groups and pairs of parameters*
$k \leftarrow 1$;
Tantque $k \leq 4$ **Faire**
    Si $IsEmpty(Grp_k) == False$ **Alors**
        // *For each minutia* M_j *of* Grp_k
        Pour $j = 1$ to $size(Grp_k)$ **Faire**
            $\begin{bmatrix} x_j^{Rot} \\ y_j^{Rot} \end{bmatrix} \leftarrow \begin{bmatrix} \cos(\delta_k) & -\sin(\delta_k) \\ \sin(\delta_k) & \cos(\delta_k) \end{bmatrix} \begin{bmatrix} x_j - x_{SP} \\ y_j - y_{SP} \end{bmatrix} + \begin{bmatrix} x_{SP} \\ y_{SP} \end{bmatrix}$;
            $\begin{bmatrix} x'_j \\ y'_j \end{bmatrix} \leftarrow \lambda_k \cdot \begin{bmatrix} x_j^{Rot} - x_{SP} \\ y_j^{Rot} - y_{SP} \end{bmatrix} + \begin{bmatrix} x_{SP} \\ y_{SP} \end{bmatrix}$;
            $\theta'_j \leftarrow \theta_j + \delta_k$;
            // *Replace* M_j *with* $M'_j(x'_j, y'_j, \theta'_j)$ *in* Grp_k
            $Replace(M_j, M'_j)$;
        Fin pour
    Finsi
    $k \leftarrow k + 1$;
Fin Tantque
// *Concatenate all groups to make a single set of transformed minutiae*
$Concatenate(Grp_1, Grp_2, Grp_3, Grp_4)$;
Retourner $M'_i = \{(x'_i, y'_i, \theta'_i) \mid i = 1, ..., n\}$;

Notre propre mise en œuvre du processus de correspondance utilisé a imposé une configuration légèrement modifiée par rapport à celle définie dans [Jiang et Yau, 2000]. En effet, nous n'avons utilisé que les six premiers éléments de l'équation 5.13 avec un ajustement approprié au nombre de propriétés utilisées. D'autre part, il faut noter que la correspondance des minuties a été évaluée dans l'article de référence sur une base de données d'empreintes digitales capturées via le capteur *CMOS* de *Veridicom* d'une taille

de 300×300 pixels, alors que la présente étude a porté comme indiqué précédemment sur les deux bases de données publiques d'empreintes digitales FVC 2002 DB1 et DB2. Les valeurs empiriques des rayons (r_1,r_2,r_3) impliquées dans ce travail et pour lesquelles les expériences ont été réalisées étaient respectivement : 100, 200 et 300.

Les exigences prises en compte dans cette évaluation sont la précision des performances dans deux scénarios différents (*Different-key* et *Stolen-key*), la révocabilité, la diversité et la non-inversibilité.

5.3.5 Précision de la vérification

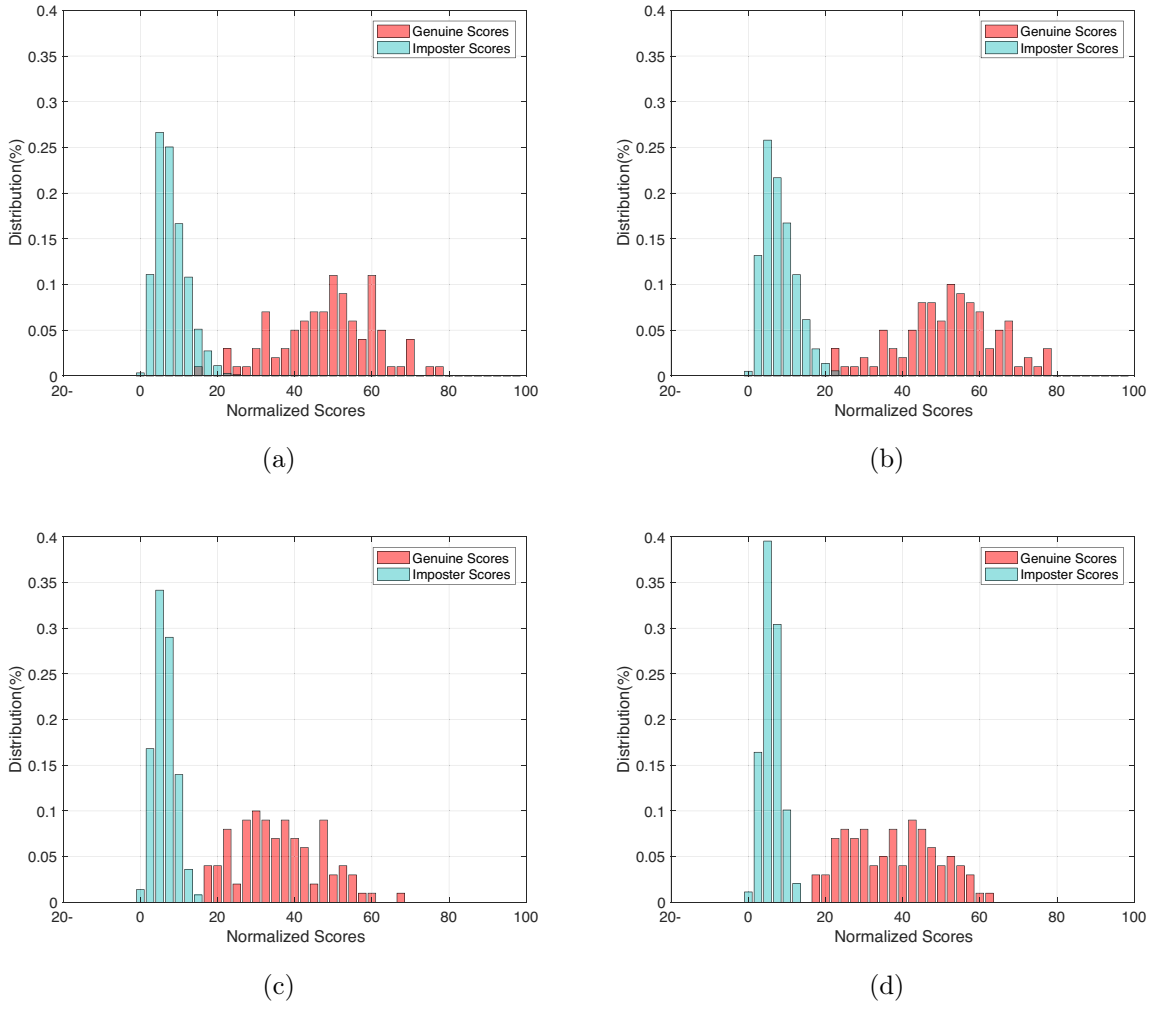


FIGURE 5.16 – Distribution des scores légitimes/imposteurs pour le système non protégé sur FVC 2002 DB1 (a) et DB2 (b), et pour le système proposé sous le scénario Different-key sur FVC 2002 DB1 (c) et DB2 (d).

Pour étudier le comportement du système d'empreintes digitales proposé en termes de

performance de reconnaissance, nous allons comparer la précision du système non protégé (utilisation des modèles d'empreintes digitales originaux) avec la précision du système protégé où chaque utilisateur détient une clé utilisateur spécifique (scénario *Different-key*). L'objectif est d'étudier l'impact des clés utilisateur sur la capacité de discrimination des modèles protégés. D'après les résultats expérimentaux, il s'avère que le taux EER obtenu par le système non protégé sur les deux bases de données FVC 2002 DB1 et DB2, respectivement, est de 1.02% et 0.13% (Tableau 5.6), tandis que le système protégé dans le cadre du scénario *Different-key* a atteint des résultats parfaits avec un taux EER de 0% pour les deux bases de données. Cela signifie que l'utilisation de clés rend les modèles d'empreintes digitales plus significatifs et plus discriminants en termes de précision de la vérification. Cela peut être confirmé dans la figure 5.16 qui représente la distribution des scores légitimes/imposteurs du système non protégé et protégé. Nous pouvons clairement observer que les distributions du système protégé sont très différentes de celles du système non protégé. D'après les résultats du test de Komolorov-Smirnov présentés dans le tableau 5.7, la séparation des distributions du système protégé a atteint une valeur de 1 pour les deux bases de données (une séparation totale), alors que celles du système non protégé correspondent respectivement à 0.9856 et 0.9972 pour FVC 2002 DB1 et DB2, ce qui signifie qu'il existe des chevauchements entre les distributions.

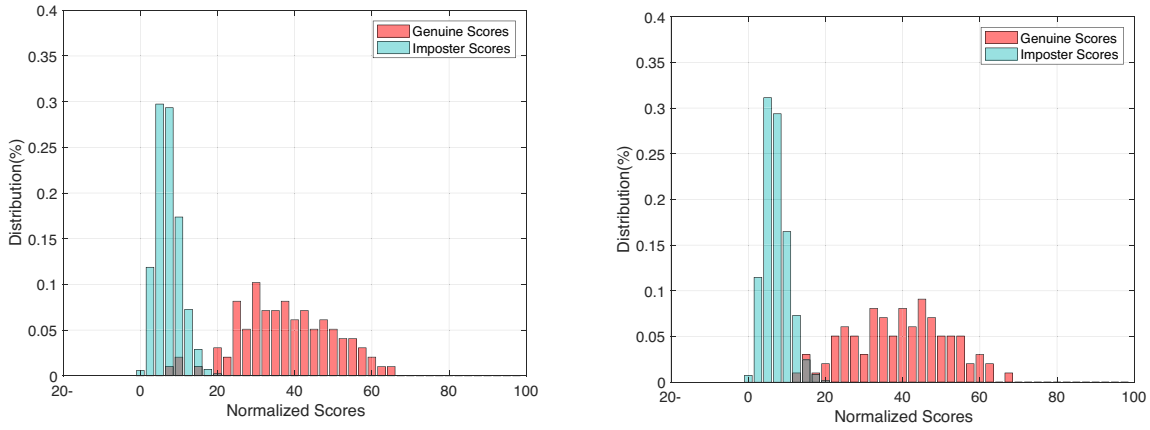
	FVC 2002 DB1	FVC 2002 DB2
Unprotected System	1.02%	0.13%
Different-key scenario	0%	0%
Stolen-key scenario	3.09%	1.83%

TABLE 5.6 – Valeurs EER obtenues à partir du Système non protégé, sous les scénarios *Different-key* et *Stolen-key* sur FVC 2002 DB1 et DB2.

	FVC 2002 DB1	FVC 2002 DB2
Unprotected System	0.9856	0.9972
Different-key scenario	1	1
Stolen-key scenario	0.9582	0.9632

TABLE 5.7 – Tests de *Kolmogorov-Smirnov* concernant le système non protégé, sous les scénarios "Different-key" et "Stolen-key" sur FVC 2002 DB1 et DB2.

Pour le scénario "*Stolen-key*", qui décrit l'événement où un imposteur intercepte la clé d'un utilisateur légitime, et tente ensuite d'accéder au système. Une simulation a été réalisée dans ce contexte, elle consiste à utiliser la même clé pour tous les utilisateurs de la base de données. D'après le graphique de la figure 5.17 qui représente la distribution



(a) (b)

FIGURE 5.17 – Distribution des scores légitimes/imposteurs pour le système proposé dans le scénario "*Stolen-key*" sur FVC 2002 DB1 (a) et DB2 (b).

des scores légitimes/imposteurs dans le scénario "Stolen-key" ainsi que le tableau 5.6, les valeurs EER résultantes étaient, respectivement, de $3,09\%$ et $1,83\%$ pour FVC 2002 DB1 et DB2, avec une séparabilité de $0,9582$ et $0,9632$ (Tableau 5.7). Il est important de noter que c'est tout à fait normal que le système réagisse de cette manière dans un tel scénario. Les performances n'ont pas été dégradées de manière significative, ce qui prouve que le système peut atteindre un certain niveau de robustesse dans le scénario d'attaque par vol de clé. Dans le même contexte, nous pouvons remarquer dans le tableau 5.8 que le système proposé montre sa supériorité par rapport à plusieurs méthodes de la littérature qui ont utilisé le même protocole d'évaluation. Ceci est évidemment dû au fait que la conception de la technique proposée a été réalisée sur la base du système non protégé et donc la méthodologie suivie a été fructueuse en termes de performance. D'autre part, il apparaît que la précision de la performance sur DB1 est supérieure à celle de DB2, comme l'illustre la courbe *ROC* de la figure 5.18, qui est dû à la qualité des empreintes digitales acquises à partir de DB2 qui est bien meilleure que celle de DB1.

5.3.6 Conformité aux exigences de révocabilité, diversité et non-inversibilité

5.3.6.1 Révocabilité

La révocabilité est une exigence essentielle des systèmes de protection des modèles. Cette propriété se manifeste par le fait de remplacer un modèle d'empreinte digitale compromis (en raison d'une attaque sur la base de modèles) par un autre modèle généré à partir du même trait d'empreinte digitale de manière à ce que le modèle compromis et le modèle révoqué soient très dissemblables. Le processus de révocation consiste simplement à utiliser une nouvelle clé utilisateur pour générer un nouveau modèle protégé, comme le

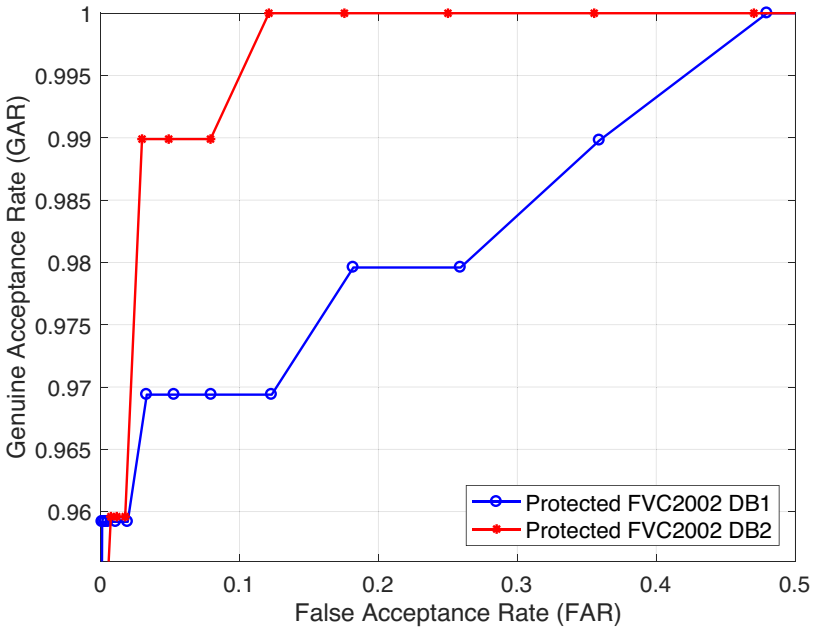


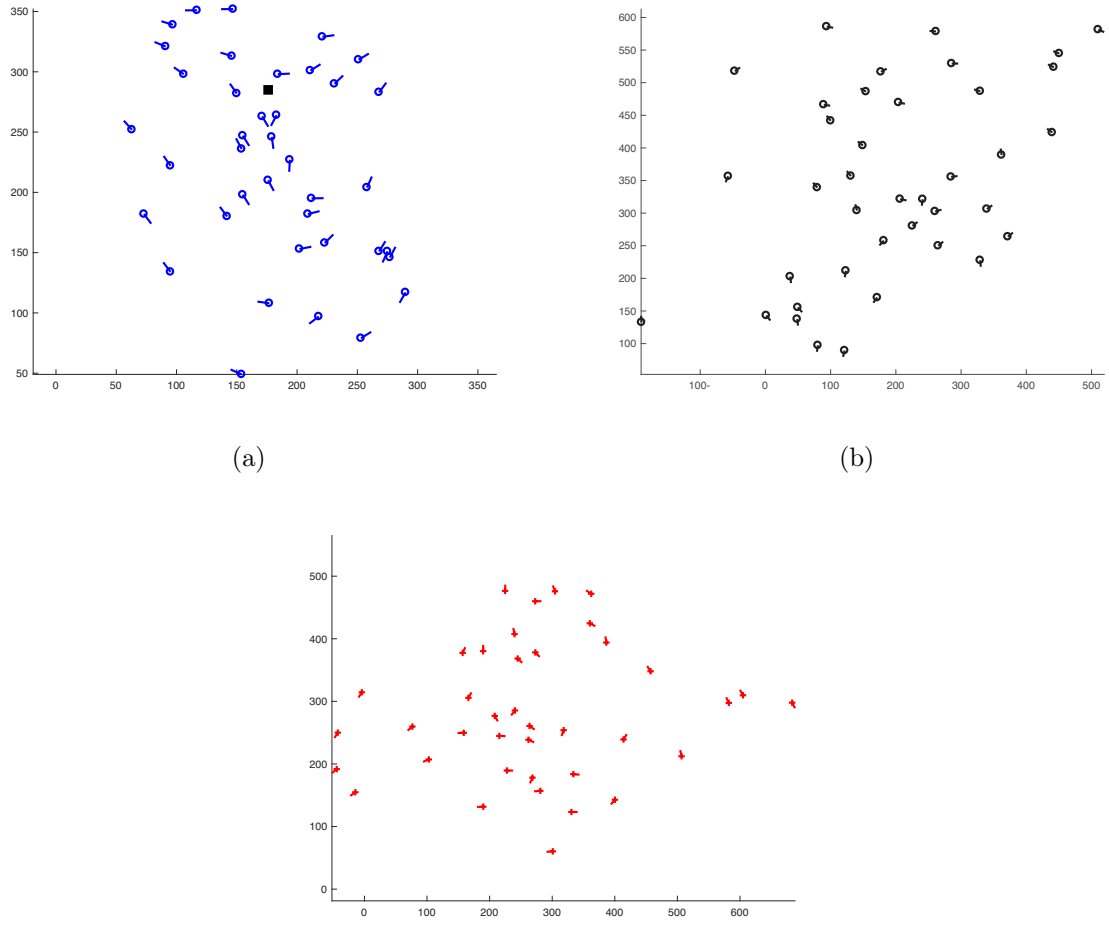
FIGURE 5.18 – Courbe ROC dans le scénario *Stolen-key* pour FVC 2002 DB1 et DB2.

Method/Dataset	FVC 2002 DB1	FVC 2002 DB2
Ahmad et al. [Ahmad *et al.*, 2011]	9	6
Jin et al. [Jin *et al.*, 2014]	4.36	1.77
Yang et al. [Yang *et al.*, 2013]	5.93	4
Yang et al. [Yang *et al.*, 2014]	3.38	0.59
Sandhya et al. [Sandhya *et al.*, 2016]	3.96	2.98
Jin et al. [Jin *et al.*, 2012]	5.19	5.65
Wang and Hu [Wang et Hu, 2012]	3.5	4
Sandhya and Prasad [Sandhya et Prasad, 2015]	4.71	3.44
Wang and Hu [Wang et Hu, 2016]	3	2
Approche Proposée	3.09	1.83

TABLE 5.8 – Comparaison de l'EER(%) avec certaines méthodes de l'état de l'art dans le cadre du scénario *Stolen-key* sur FVC2002 DB1 et DB2.

montre la figure 5.19. Pour évaluer notre proposition en termes de révocabilité, le scénario de l'attaque *the revoked template attack* [Ali *et al.*, 2018] sera étudié, où un adversaire tente de cibler le système avec un modèle d'empreinte digitale compromis. Dans un tel

scénario, il existe deux types d'attaques (sous-section 5.2.5.1). Après avoir expérimenté sur les deux bases de données FVC 2002 DB1 et DB2, le pourcentage de vérification réussie était de 0%, ce qui signifie qu'il y a une dissociation totale entre les gabarits compromis et révoqués. Cela prouve à coup sûr que le système est suffisamment robuste contre l'attaque du modèle révoqué et donc conforme avec la notion de révocabilité.



(a) (b)

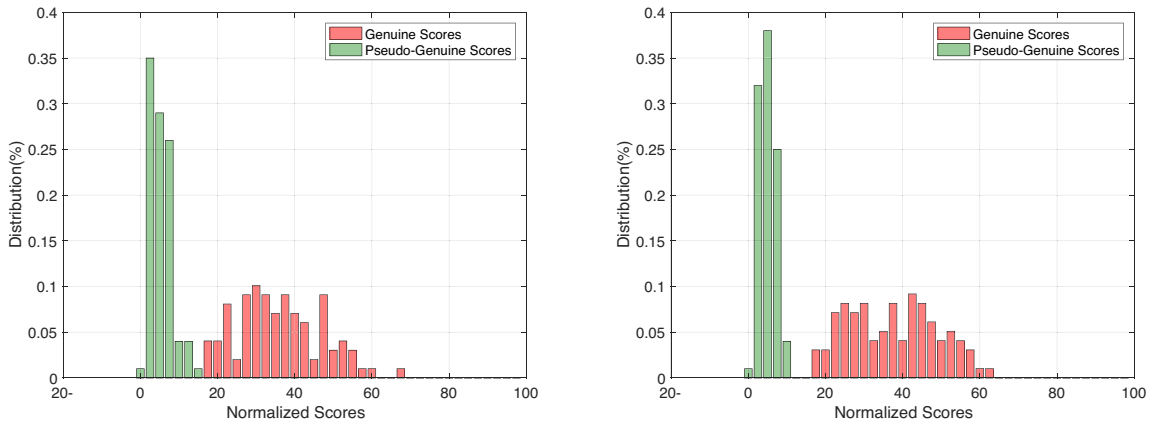
(c)

FIGURE 5.19 – Transformation d'un modèle original d'empreinte digitale en utilisant deux ensembles différents de paramètres de clé utilisateur. (a) Modèle original, (b) Modèle transformé utilisant le premier ensemble, (c) Modèle transformé utilisant le second ensemble.

5.3.6.2 *Diversité*

La diversité fait référence à la possibilité de générer plusieurs modèles distincts à partir d'un même trait biométrique, de telle sorte que tous ces modèles n'ont aucun lien avec le modèle original, mais aussi entre eux-mêmes. Pour évaluer le système en termes de

dissociabilité, nous considérons deux systèmes effectuant des transformations différentes sur les mêmes empreintes digitales. Le premier système prend au hasard des paramètres de clés utilisateur dans les plages suivantes : $\{\delta_i\}_{i=1}^{4} \in [0, 90]$, $\{\lambda_i\}_{i=1}^{4} \in [-1, 0]$. Tandis que le second système fait intervenir les suivantes : $\{\delta_i\}_{i=1}^{4} \in [180, 270]$, $\{\lambda_i\}_{i=1}^{4} \in [2, 3]$. L'objectif du test est de construire la distribution des scores pseudo-authentiques en comparant les modèles transformés du même doigt générés par le système 1 avec ceux produits par le système 2. D'après le graphique de la figure 5.20, nous pouvons remarquer qu'il n'y a pas de chevauchement entre les distributions des scores pseudo-légitimes et légitimes. La séparabilité selon le test de *Komolorov-Smirnov* atteint une valeur de 1 (séparation totale) pour les deux bases de données. La distribution des scores pseudo-légitimes est donc presque identique à celle des imposteurs (Figure 5.16), ceci explique que bien que les modèles soient générés à partir de la même empreinte digitale, ils ne sont pas appariés entre eux. Ainsi, nous pouvons affirmer que le système proposé est conforme à l'exigence de diversité.



(a) (b)

FIGURE 5.20 – La distribution des scores légitime-pseudo-légitime pour FVC 2002 DB1 (a) et DB2 (b).

5.3.6.3 *Non-inversibilité*

La non-inversibilité désigne la possibilité de divulguer une partie ou la totalité du modèle original (modèle sans transformation). Un système de protection robuste doit être capable de rendre les gabarits biométriques originaux difficiles à reconstruire afin de garantir la confidentialité de l'utilisateur. Afin d'étudier les possibilités de révélation en cas d'attaque par inversion de gabarit sur notre système, nous supposons qu'un gabarit protégé est intercepté par un adversaire. Dans cette situation, l'adversaire peut obtenir toutes les positions et les orientations des minuties modifiées à partir du modèle compromis (après la transformation). Cependant, ces informations ne peuvent pas conduire à

celles du domaine original puisque la transformation effectuée dans le schéma proposé est basée principalement sur la position du point singulier utilisé et l'ensemble des valeurs de la clé utilisateur. En fait, en l'absence de ces informations, l'adversaire ne peut en aucun cas calculer les positions et orientations des points caractéristiques originaux. La transformation menée est en fait une sorte de rotation et d'une opération d'homothétie qui prend comme centre le point singulier. L'implication de ce point est cruciale pour effectuer les changements linéaires ainsi que les valeurs de δ et λ fournies par la clé utilisateur, où δ fait référence à l'angle de rotation et λ représente le rapport d'homothétie. Par conséquent, même si l'adversaire est en possession du modèle compromis et des paramètres de la clé utilisateur avec lesquels le modèle a été généré, et qu'il tente d'atteindre les informations de minuties initiales, il ne pourra pas inverser l'homothétie ou la rotation car il ignore la position du point singulier ainsi que les distances entre les minuties et le point singulier. Même si l'adversaire effectue une attaque par force brute sur le point singulier, cela ne mène nulle part.

5.4 Bilan du chapitre

Dans ce chapitre, nous avons présenté nos deux contributions à la protection des modèles d'empreintes digitales : Dans la première contribution, nous avons proposé un nouveau schéma dédié à la protection des modèles d'empreintes digitales. Il s'agit d'une technique basée sur les caractéristiques des minuties et qui ne requière aucun point d'enregistrement ou d'alignement. Le concept de base consiste à ne pas utiliser directement les caractéristiques du modèle original, mais plutôt à exploiter les particularités du tétraèdre formées à partir de chaque minutie et son voisinage avec un vecteur généré de manière aléatoire, afin d'effectuer d'abord des transformations géométriques irréversibles, puis de récupérer uniquement les propriétés discriminantes par lesquelles le modèle protégé final sera représenté. Ce système a été conçu principalement pour réduire l'impact des erreurs d'extraction des minuties et pour surmonter les effets indésirables de la déformation élastique de l'empreinte digitale. Selon les résultats expérimentaux, l'approche proposée fournit de bons résultats par rapport à de nombreuses techniques de l'état de l'art, que ce soit pour FVC2002 DB1 ou FVC2002 DB2, avec une légère dégradation pour FVC2002 DB3 en raison de la mauvaise qualité des images d'empreintes digitales dans cette base de données. Dans le scénario *Stolen-key*, les performances du protocole FVC original sont certes réduites, mais elles restent très acceptables par rapport à certaines techniques existantes. En outre, l'approche proposée garantit un niveau de sécurité élevé et répond aux exigences de révocabilité, de diversité et de non-inversibilité, qui sont nécessaires dans une technique idéale de protection.

La conception des schémas de protection des modèles biométriques a toujours été un problème majeur pour la communauté des chercheurs en sécurité. Le défi consiste à maintenir à la fois un niveau de sécurité élevé et une grande précision de vérification. Dans la deuxième contribution, nous avons adopté une nouvelle méthodologie pour la conception

des schémas de protection de modèles d'empreintes digitales. En effet, nous avons pris en considération la spécification d'un système de vérification d'empreintes digitales non protégé (c'est-à-dire le processus de correspondance des minuties utilisé) pour construire un schéma de protection spécifique et approprié qui fournit le meilleur compromis entre la performance et la sécurité par rapport à toute solution de protection générique. La proposition est une technique basée sur les minuties, entraînant des désordres au niveau des caractéristiques des minuties (à savoir, la position et l'orientation) en se basant sur les informations extraites des minuties et celles du point singulier principal, et en produisant ainsi un nouveau modèle différent qui satisfait parfaitement aux critères de révocabilité, de diversité, de sécurité et de performance, malgré les erreurs d'extraction et les variations dues à la translation et à la rotation du doigt. La précision des performances de la proposition a été évaluée en utilisant les bases de données publiques d'empreintes digitales : FVC 2002 DB1 et DB2. Les résultats expérimentaux montrent que le système proposé, dans le scénario *Different-key*, améliore la diversité des modèles d'empreintes digitales et les rend plus discriminants par rapport au système non protégé. Dans le scénario *Stolen-key*, la performance s'est légèrement dégradée mais reste généralement acceptable par rapport à plusieurs méthodes de l'état de l'art. Par ailleurs, il est à noter que la précision de vérification de notre approche est liée à deux facteurs essentiels : d'une part, la précision de l'extraction des minuties et d'autre part, la présence et la précision de la détection des points singuliers.

Chapitre

6

Conclusion générale et perspectives

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 traditionnels d'authentification des individus, dans la mesure où la relation entre l'authentifiant et l'individu ne peut pas être plus étroite. En revanche, bien que les systèmes biométriques se soient bien positionnés, le problème d'assurer la sécurité et la confidentialité des données biométriques demeure critique. En effet, l'utilisation croissante de la biométrie dans les applications de la sécurité a suscité un regain d'intérêt pour la recherche et l'exploration de nouvelles méthodes pour attaquer les systèmes biométriques. Ces recherches ont prouvé que ces systèmes sont vulnérables à un certain nombre d'attaques.

Dans cette thèse, nous nous sommes intéressés à la protection contre les attaques qui visent les modèles des empreintes digitales. Pourquoi une telle biométrie ? L'empreinte digitale est une technologie biométrique très attrayante, principalement, en raison de ses performances et de son acceptabilité. Malheureusement, de nombreux problèmes lui sont aussi associés. Il existe des préoccupations majeures pour des raisons d'éthique, de sécurité et d'invasion de la vie privée. L'empreinte digitale est une modalité à trace, elle n'est donc pas secrète. En cas de compromission, comme il est impossible de la révoquer, elle devient inutilisable. Les schémas de protection des modèles des empreintes digitales existants sont des solutions prometteuses pour les problèmes de révocabilité et de confidentialité. Pour l'heure, les propositions apportées ne sont pas totalement satisfaisantes en terme de performance. On observe une dégradation des taux de reconnaissance par rapport aux systèmes sans protection. Dans la plupart des cas, le critère de non-inversibilité, souhaité pour garantir le respect de la vie privée n'est que partiellement atteint.

Nos principaux objectifs étaient, premièrement, d'étudier la nature des caractéristiques qui constituent l'empreinte digitale, d'analyser les systèmes de reconnaissance par cette biométrie, puis d'évaluer les menaces liées aux attaques contre les modèles des empreintes digitales, et finalement, de concevoir et de développer des schémas de protection génériques pour ces modèles afin d'améliorer la robustesse générale des systèmes des empreintes digitales sans pour autant dégrader leurs performances.

Le but du chapitre 1 a été d'introduire le problème de la sécurité dans les systèmes biométriques, et de clarifier le contexte de ce travail, la perspective suivie lors de l'éla-

boration de la thèse, les objectifs et les contributions principales. Le chapitre 2 a été consacré aux généralités sur les systèmes biométriques et aux vulnérabilités qui font face à ce genre de système. Une description plus détaillé sur la biométrie des empreintes digitale a été abordé puisqu'elle fait l'objet de notre travail. Dans le chapitre 3, nous avons décrit les exigences d'un schéma idéal de protection des modèles biométriques, puis nous avons présenter un état de l'art des travaux sur la protection des systèmes de sécurité biométriques, et particulièrement sur la protection des modèles des empreintes digitales. Le chapitre 4 a abordé les différentes méthodologies statistiques disponibles pour évaluer la performance/sécurité des systèmes biométriques/schémas de protection. Dans le Chapitre 5, nous avons présenté nos deux contributions à la protection des modèles des empreintes digitales.

Dans la première contribution, nous avons proposé un nouveau schéma dédié à la protection des modèles des empreintes digitales. Il s'agit d'une technique basée sur les points caractéristiques (les minuties) qui ne nécessite aucun point d'enregistrement ou d'alignement. Le concept n'est pas d'utiliser directement les caractéristiques du modèle original, mais plutôt d'exploiter les caractéristiques du tétraèdre formées à partir de chaque minutie avec un vecteur généré aléatoirement pour d'abord effectuer des transformations géométriques irréversibles et ensuite récupérer uniquement les caractéristiques discriminantes par lesquelles le modèle final protégé sera représenté. D'après les résultats expérimentaux, l'approche proposée donne de bons résultats par rapport à de nombreuses techniques de l'état de l'art, que ce soit pour FVC2002 DB1 ou FVC2002 DB2, avec une légère baisse pour FVC2002 DB3 en raison de la mauvaise qualité des images des empreintes digitales. Dans le scénario du *Stolen-key*, les performances du protocole FVC original sont réduites, mais restent acceptables par rapport à certaines techniques existantes. De plus, l'approche proposée garantit un niveau de sécurité élevé et répond aux exigences de révocabilité, de diversité et de non-inversibilité. Dans le cadre de travaux futurs, afin de trouver un moyen d'améliorer encore les performances du système, nous tenterons d'appliquer certains processus pour améliorer la technique d'extraction, comme celui mentionné dans [Jin *et al.*, 2014] (*User-specific Minutia Vicinities*), qui permet de fusionner plusieurs échantillons d'entraînement et d'extraire ensuite les structures locales les plus stables. En outre, nous chercherons des caractéristiques plus stables et plus pertinentes pour représenter les modèles protégés et les transformer en une chaîne binaire.

Dans la deuxiéme contribution, nous avons adopté une nouvelle méthodologie pour la conception de schémas de protection de modèles des empreintes digitales. En effet, nous avons pris en considération la spécification d'un système de vérification des empreintes digitales non protégé (c'est-à-dire le processus de correspondance des minuties utilisé) pour construire un schéma de protection spécifique et approprié qui fournit le meilleur compromis entre la performance et la sécurité par rapport à toute solution de protection générique. La proposition est une technique basée sur les minuties qui provoque des désordres en termes de caractéristiques des minuties (à savoir la position et l'orientation) en se basant sur les informations extraites des minuties et celles du point singulier prin-

cipal, et en produisant ainsi un nouveau modèle différent qui satisfait parfaitement la révocabilité, la diversité, la sécurité et la performance, malgré les erreurs d'extraction et les variations dues à la translation et à la rotation du doigt. La précision des performances de la proposition a été évaluée en utilisant les bases de données publiques des empreintes digitales : FVC 2002 DB1 et DB2. Les résultats expérimentaux montrent que le système proposé, dans le scénario de *Different-key*, améliore la diversité des modèles d'empreintes digitales et les rend plus discriminants par rapport au système non protégé. Dans le scénario de *Stolen-key*, la performance s'est légèrement dégradée, mais reste généralement acceptable par rapport à plusieurs méthodes de l'état de l'art. Dans le cadre de travaux futurs, nous prévoyons d'étendre notre évaluation à d'autres bases de données présentant des empreintes digitales de qualité inférieure (variations intra-classe importantes). Nous avons également l'intention d'améliorer le processus de correspondance des minuties utilisé pour obtenir des résultats parfaits dans le système non protégé, en utilisant par exemple d'autres propriétés dans la structure locale des minuties. Enfin, nous aborderons la question d'éviter l'utilisation de points singuliers, car la précision de leur détection n'est parfois pas évidente. Par ailleurs, nos plans futurs incluent également la conception de nouveaux systèmes de protection pour les systèmes multimodaux et l'inclusion des techniques de tatouage dans nos schémas de protection.

Annexes

ANNEXE

A

NOTIONS

A.1 L'algorithme Hill-Climbing

Pour régénérer les images originales à partir des scores de correspondance, l'algorithme *Hill-Climbing* procède comme suit [Adler, 2003b] :

- On teste plusieurs images dans le système et on collecte les scores de correspondance. L'image I_i qui a le plus grand score S_i est utilisée comme une image d'attaque initiale.
- On va modifier I_i sur plusieurs itératives pour construire l'estimation finale I_f comme suit :
  1. Modifier aléatoirement I_i.
  2. Calculer le score S_f en utilisant l'image modifiée.
  3. Si $S_f > S_i$, Alors $I_i = I_f$ et $S_i = S_f$
  4. Arrêter l'algorithme si S_f ne s'augmente plus.

Le défi majeur dans l'implémentation de cet algorithme d'attaque, est de trouver la façon optimale ou logique pour modifier I_i de telle manière S_f s'augmente.

A.2 Orthogonalisation de Gram-Schmidt

Soit $v_1, v_2, ..., v_n$ un ensemble de vecteurs linéairement indépendants. L'orthogonalisation de *Gram-Schmidt* procède comme suit :

$$u_1 = v_1 \text{ et } u_1 \frac{u_1}{\| u_1 \|} \tag{A.1}$$

$$u_2 = v_2 - proj_{u_1} v_2 \text{ et } u_2 \frac{u_2}{\| u_2 \|} \tag{A.2}$$

$$u_i = v_i - \sum_{k=1}^{i} -1 proj_{u_k} vi \text{ et } u_i \frac{u_i}{\| u_i \|} \tag{A.3}$$

— $proj$ est l'opérateur de projection orthogonale.

— $proj_u v = \langle u, v \rangle u$.

— $\langle u, v \rangle$ est le produit scalaire.

— $\| \, . \, \|$ la norme d'un vecteur.

A.3 Formule de Héron

En géométrie euclidienne, la formule de *Héron* permet de calculer l'aire S d'un triangle quelconque en ne connaissant que les longueurs a, b et c de ses trois côtés :

$$S = \sqrt{p(p-a)(p-b)(p-c)} \tag{A.4}$$

avec

$$p = \frac{a+b+c}{2} \tag{A.5}$$

Annexe

B

Production scientifique

Les travaux réalisés et présentés dans ce mémoire ont été valorisés dans les publications indiquées ci-dessous :

Revues internationales (2)

1. **A. Lahmidi**, K. Minaoui, C. Moujahdi, M. Rziza, "Fingerprint Template Protection Using Irreversible Minutiae Tetrahedrons", The Computer Journal (**IF=1.494**), 2021, DOI : 10.1093/comjnl/bxab111.
2. **A. Lahmidi**, K. Minaoui, C. Moujahdi, M. Rziza, "A New Protection Scheme for Biometric Templates based on Random Projection and CDMA Principle", International Journal of Advanced Computer Science and Applications (**IF=1.092**), vol. 12, no. 10, 2021, DOI : 10.14569/IJACSA.2021.0121088.
3. **A. Lahmidi**, C. Moujahdi, K. Minaoui, M. Rziza, "On the methodology of fingerprint template protection schemes conception : Meditations on the reliability", EURASIP Journal on Information Security (**IF=3.17**), vol. 2022, no 1, p. 1-13, 2022, DOI : 10.1186/s13635-022-00129-6.

Conférences nationales et internationales (2)

4. **A. Lahmidi**, K. Minaoui, M. Rziza, "A robust minutia-based approach for securing fingerprint templates", 9th International Symposium on Signal, Image, Video and Communications (ISIVC), IEEE, 27-30 Novembre 2018.
5. **A. Lahmidi**, K. Minaoui, M. Rziza, "A variant of Biohashing based on the chaotic behavior of the Logistic Map", 4th International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBIoTS), IEEE, 12-13 Decembre 2019.

Bibliographie

[19795-1, 2006] 19795-1, I. (2006). Information technology - biometric performance testing and reporting. part 1 : Principles and framework.

[Adler, 2003a] Adler, A. (2003a). Can images be regenerated from biometric templates ? *In Biometrics Consortium Conference*.

[Adler, 2003b] Adler, A. (2003b). Sample images can be independently restored from face recognition templates. *In CCECE 2003-Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No. 03CH37436)*, volume 2, pages 1163–1166. IEEE.

[Adler, 2004] Adler, A. (2004). Images can be regenerated from quantized biometric match score data. *In Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No. 04CH37513)*, volume 1, pages 469–472. IEEE.

[Adler, 2005] Adler, A. (2005). Vulnerabilities in biometric encryption systems. *In International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 1100–1109. Springer.

[Adler et Schuckers, 2007] Adler, A. et Schuckers, M. E. (2007). Comparing human and automatic face recognition performance. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(5):1248–1255.

[Ahmad *et al.*, 2011] Ahmad, T., Hu, J. et Wang, S. (2011). Pair-polar coordinate-based cancelable fingerprint templates. *Pattern recognition*, 44(10-11):2555–2564.

[Ahn *et al.*, 2008] Ahn, D., Kong, S. G., Chung, Y.-S. et Moon, K. Y. (2008). Matching with secure fingerprint templates using non-invertible transform. *In 2008 Congress on Image and Signal Processing*, volume 2, pages 29–33. IEEE.

[Ali *et al.*, 2018] Ali, S. S., Ganapathi, I. I. et Prakash, S. (2018). Robust technique for fingerprint template protection. *IET Biometrics*, 7(6):536–549.

[Ali *et al.*, 2020] Ali, S. S., Ganapathi, I. I., Prakash, S., Consul, P. et Mahyo, S. (2020). Securing biometric user template using modified minutiae attributes. *Pattern Recognition Letters*, 129:263–270.

[Ali et Prakash, 2015] ALI, S. S. et PRAKASH, S. (2015). Enhanced fingerprint shell. *In 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 801–805. IEEE.

[Ali et Prakash, 2017] ALI, S. S. et PRAKASH, S. (2017). Fingerprint shell construction with prominent minutiae points. *In Proceedings of the 10th Annual ACM India Compute Conference*, pages 91–98.

[Ali et Prakash, 2019] ALI, S. S. et PRAKASH, S. (2019). 3-dimensional secured fingerprint shell. *Pattern Recognition Letters*, 126:68–77.

[Ang *et al.*, 2005] ANG, R., SAFAVI-NAINI, R. et MCAVEN, L. (2005). Cancelable key-based fingerprint templates. *In Australasian conference on information security and privacy*, pages 242–252. Springer.

[Arakala *et al.*, 2007] ARAKALA, A., JEFFERS, J. et HORADAM, K. J. (2007). Fuzzy extractors for minutiae-based fingerprint authentication. *In International conference on biometrics*, pages 760–769. Springer.

[Blanton et Aliasgari, 2013] BLANTON, M. et ALIASGARI, M. (2013). Analysis of reusability of secure sketches and fuzzy extractors. *IEEE transactions on information forensics and security*, 8(9):1433–1445.

[Bleumer, 1999] BLEUMER, G. (1999). Biometric authentication and multilateral security. *Multilaterial security in communications, Addison-Wesley*, pages 157–172.

[Bolle *et al.*, 2002] BOLLE, R. M., CONNELL, J. H. et RATHA, N. K. (2002). Biometric perils and patches. *Pattern recognition*, 35(12):2727–2738.

[Boult, 2006] BOULT, T. (2006). Robust distance measures for face-recognition supporting revocable biometric tokens. *In 7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pages 560–566. IEEE.

[Boult *et al.*, 2007] BOULT, T. E., SCHEIRER, W. J. et WOODWORTH, R. (2007). Revocable fingerprint biotokens : Accuracy and security analysis. *In 2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE.

[Bringer *et al.*, 2008] BRINGER, J., CHABANNE, H., COHEN, G., KINDARJI, B. et ZEMOR, G. (2008). Theoretical and practical boundaries of binary secure sketches. *IEEE Transactions on Information Forensics and Security*, 3(4):673–683.

[Campisi, 2013] CAMPISI, P. (2013). *Security and privacy in biometrics*, volume 24. Springer.

[Cappelli *et al.*, 2005] CAPPELLI, R., MAIO, D., MALTONI, D., WAYMAN, J. L. et JAIN, A. K. (2005). Performance evaluation of fingerprint verification systems. *IEEE transactions on pattern analysis and machine intelligence*, 28(1):3–18.

[Cavoukian, 1999] CAVOUKIAN, A. (1999). *Privacy and biometrics*. Information and Privacy Commissioner/Ontario.

[Chang *et al.*, 2006] CHANG, E.-C., SHEN, R. et TEO, F. W. (2006). Finding the original point set hidden among chaff. *In Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, pages 182–188.

[Chang *et al.*, 2004] CHANG, Y.-J., ZHANG, W. et CHEN, T. (2004). Biometrics-based cryptographic key generation. *In 2004 IEEE International Conference on Multimedia and Expo (ICME)(IEEE Cat. No. 04TH8763)*, volume 3, pages 2203–2206. IEEE.

[Chikkerur *et al.*, 2008] CHIKKERUR, S., RATHA, N. K., CONNELL, J. H. et BOLLE, R. M. (2008). Generating registration-free cancelable fingerprint templates. *In 2008 IEEE Second International Conference on Biometrics : Theory, Applications and Systems*, pages 1–6. IEEE.

[Chin *et al.*, 2006] CHIN, C. S., JIN, A. T. B. et LING, D. N. C. (2006). High security iris verification system based on random secret integration. *Computer Vision and Image Understanding*, 102(2):169–177.

[Connie *et al.*, 2004] CONNIE, T., TEOH, A., GOH, M. et NGO, D. (2004). Palmhashing : a novel approach for dual-factor authentication. *Pattern Analysis and Applications*, 7(3):255–268.

[Connie *et al.*, 2005] CONNIE, T., TEOH, A., GOH, M. et NGO, D. (2005). Palmhashing : a novel approach for cancelable biometrics. *Information processing letters*, 93(1):1–5.

[Cooper *et al.*, 2007] COOPER, D., DANG, H., LEE, P., MACGREGOR, W., MEHTA, K. *et al.* (2007). Secure biometric match-on-card feasibility report. *NIST Interagency Report*, 2.

[Cukic et Bartlow, 2005] CUKIC, B. et BARTLOW, N. (2005). Biometric system threats and countermeasures : a risk based approach. *In Proceedings of the Biometric Consortium Conference.*

[Das *et al.*, 2012] DAS, P., KARTHIK, K. et GARAI, B. C. (2012). A robust alignment-free fingerprint hashing algorithm based on minimum distance graphs. *Pattern Recognition*, 45(9):3373–3388.

[Dasgupta et Gupta, 1999] DASGUPTA, S. et GUPTA, A. (1999). An elementary proof of the johnson-lindenstrauss lemma. *International Computer Science Institute, Technical Report*, 22(1):1–5.

[Derman et Keskinöz, 2016] DERMAN, E. et KESKINÖZ, M. (2016). Normalized cross-correlation based global distortion correction in fingerprint image matching. *In 2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 1–4. IEEE.

[Dodis *et al.*, 2008] DODIS, Y., OSTROVSKY, R., REYZIN, L. et SMITH, A. (2008). Fuzzy extractors : How to generate strong keys from biometrics and other noisy data. *SIAM journal on computing*, 38(1):97–139.

[Domeniconi *et al.*, 1998] DOMENICONI, C., TARI, S. et LIANG, P. (1998). Direct gray scale ridge reconstruction in fingerprint images. *In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, volume 5, pages 2941–2944. IEEE.

[Duta, 2009] DUTA, N. (2009). A survey of biometric technology based on hand shape. *Pattern recognition*, 42(11):2797–2806.

[Egan et Egan, 1975] EGAN, J. P. et EGAN, J. P. (1975). *Signal detection theory and ROC-analysis*. Academic press.

[Eriksson et Wretling, 1997] ERIKSSON, A. et WRETLING, P. (1997). How flexible is the human voice ?-a case study of mimicry. *In Fifth European Conference on Speech Communication and Technology*.

[Farooq *et al.*, 2007] FAROOQ, F., BOLLE, R. M., JEA, T.-Y. et RATHA, N. (2007). Anonymous and revocable fingerprint recognition. *In 2007 IEEE conference on computer vision and pattern recognition*, pages 1–7. IEEE.

[Feng et Jain, 2010] FENG, J. et JAIN, A. K. (2010). Fingerprint reconstruction : from minutiae to phase. *IEEE transactions on pattern analysis and machine intelligence*, 33(2):209–223.

[Ferrara *et al.*, 2012] FERRARA, M., MALTONI, D. et CAPPELLI, R. (2012). Noninvertible minutia cylinder-code representation. *IEEE Transactions on Information Forensics and Security*, 7(6):1727–1737.

[Ferrara *et al.*, 2014] FERRARA, M., MALTONI, D. et CAPPELLI, R. (2014). A two-factor protection scheme for mcc fingerprint templates. *In 2014 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–8. IEEE.

[Freire *et al.*, 2007] FREIRE, M., FIERREZ, J., MARTINEZ-DIAZ, M. et ORTEGA-GARCIA, J. (2007). On the applicability of off-line signatures to the fuzzy vault construction. *In Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 1173–1177. IEEE.

[Galbally *et al.*, 2009] GALBALLY, J., FIERREZ, J., ORTEGA-GARCIA, J., MCCOOL, C. et MARCEL, S. (2009). Hill-climbing attack to an eigenface-based face verification system. *In 2009 First IEEE International Conference on Biometrics, Identity and Security (BIdS)*, pages 1–6. IEEE.

[Galbally *et al.*, 2010] GALBALLY, J., MCCOOL, C., FIERREZ, J., MARCEL, S. et ORTEGA-GARCIA, J. (2010). On the vulnerability of face verification systems to hill-climbing attacks. *Pattern Recognition*, 43(3):1027–1038.

[Garris *et al.*, 2001] GARRIS, M. D., WATSON, C. I., MCCABE, R., WILSON, C. L. *et al.* (2001). User's guide to nist fingerprint image software (nfis).

[Ghany *et al.*, 2012] GHANY, K. K. A., HEFNY, H. A., HASSANIEN, A. E. et GHALI, N. I. (2012). A hybrid approach for biometric template security. *In 2012 IEEE/ACM*

International Conference on Advances in Social Networks Analysis and Mining, pages 941–942. IEEE.

[Goel *et al.*, 2005] GOEL, N., BEBIS, G. et NEFIAN, A. (2005). Face recognition experiments with random projection. *In Biometric Technology for Human Identification II*, volume 5779, pages 426–437. International Society for Optics and Photonics.

[Group, 2013] GROUP, I. B. (2013). Biometrics market and industry report 2009-2014. *Technical report*.

[Hämmerle-Uhl *et al.*, 2009] HÄMMERLE-UHL, J., PSCHERNIG, E. et UHL, A. (2009). Cancelable iris biometrics using block re-mapping and image warping. *In International conference on information security*, pages 135–142. Springer.

[Hao *et al.*, 2006] HAO, F., ANDERSON, R. et DAUGMAN, J. (2006). Combining crypto with biometrics effectively. *IEEE transactions on computers*, 55(9):1081–1088.

[Hill, 2001] HILL, C. J. (2001). Risk of masquerade arising from the storage of biometrics. *Bachelor of Science thesis, The Department of Computer Science, Australian National University*.

[Holland et Komogortsev, 2013] HOLLAND, C. D. et KOMOGORTSEV, O. V. (2013). Complex eye movement pattern biometrics : Analyzing fixations and saccades. *In 2013 International conference on biometrics (ICB)*, pages 1–8. IEEE.

[Hong *et al.*, 1998] HONG, L., WAN, Y. et JAIN, A. (1998). Fingerprint image enhancement : algorithm and performance evaluation. *IEEE transactions on pattern analysis and machine intelligence*, 20(8):777–789.

[Information *et al.*, 2008] INFORMATION, COMMISSIONER/ONTARIO, P. et CAVOUKIAN, A. (2008). *Fingerprint biometrics : Address privacy before deployment*. Information and Privacy Commissioner of Ontario.

[Jain *et al.*, 2006a] JAIN, A. K., BOLLE, R. et PANKANTI, S. (2006a). *Biometrics : personal identification in networked society*, volume 479. Springer Science & Business Media.

[Jain *et al.*, 2007] JAIN, A. K., FLYNN, P. et ROSS, A. A. (2007). *Handbook of biometrics*. Springer Science & Business Media.

[Jain *et al.*, 1997] JAIN, A. K., HONG, L., PANKANTI, S. et BOLLE, R. (1997). An identity-authentication system using fingerprints. *Proceedings of the IEEE*, 85(9):1365–1388.

[Jain *et al.*, 2008] JAIN, A. K., NANDAKUMAR, K. et NAGAR, A. (2008). Biometric template security. *EURASIP Journal on advances in signal processing*, 2008:1–17.

[Jain et Ross, 2004] JAIN, A. K. et ROSS, A. (2004). Multibiometric systems. *Communications of the ACM*, 47(1):34–40.

[Jain *et al.*, 2006b] JAIN, A. K., ROSS, A. et PANKANTI, S. (2006b). Biometrics : a tool for information security. *IEEE transactions on information forensics and security*, 1(2):125–143.

[Jain *et al.*, 2004] JAIN, A. K., ROSS, A. et PRABHAKAR, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1):4–20.

[Jiang et Yau, 2000] JIANG, X. et YAU, W.-Y. (2000). Fingerprint minutiae matching based on the local and global structures. *In Proceedings 15th international conference on pattern recognition. ICPR-2000*, volume 2, pages 1038–1041. IEEE.

[Jin *et al.*, 2004] JIN, A. T. B., LING, D. N. C. et GOH, A. (2004). Biohashing : two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 37(11):2245–2255.

[Jin *et al.*, 2014] JIN, Z., LIM, M.-H., TEOH, A. B. J. et GOI, B.-M. (2014). A non-invertible randomized graph-based hamming embedding for generating cancelable fingerprint template. *Pattern Recognition Letters*, 42:137–147.

[Jin *et al.*, 2012] JIN, Z., TEOH, A. B. J., ONG, T. S. et TEE, C. (2012). Fingerprint template protection with minutiae-based bit-string for security and privacy preserving. *Expert systems with applications*, 39(6):6157–6167.

[Juels *et al.*, 2005] JUELS, A., MOLNAR, D. et WAGNER, D. (2005). Security and privacy issues in e-passports. *In First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05)*, pages 74–88. IEEE.

[Juels et Sudan, 2006] JUELS, A. et SUDAN, M. (2006). A fuzzy vault scheme. *Designs, Codes and Cryptography*, 38(2):237–257.

[Juels et Wattenberg, 1999] JUELS, A. et WATTENBERG, M. (1999). A fuzzy commitment scheme. *In Proceedings of the 6th ACM conference on Computer and communications security*, pages 28–36.

[Kelkboom *et al.*, 2007] KELKBOOM, E. J., GÖKBERK, B., KEVENAAR, T. A., AKKERMANS, A. H. et van der VEEN, M. (2007). "3d face" : biometric template protection for 3d face recognition. *In International Conference on Biometrics*, pages 566–573. Springer.

[Kho *et al.*, 2019] KHO, J. B., KIM, J., KIM, I.-J. et TEOH, A. B. (2019). Cancelable fingerprint template design with randomized non-negative least squares. *Pattern Recognition*, 91:245–260.

[Kong *et al.*, 2006] KONG, A., CHEUNG, K.-H., ZHANG, D., KAMEL, M. et YOU, J. (2006). An analysis of biohashing and its variants. *Pattern recognition*, 39(7):1359–1368.

[Kumar et Kumar, 2009] KUMAR, A. et KUMAR, A. (2009). Development of a new cryptographic construct using palmprint-based fuzzy vault. *EURASIP Journal on Advances in Signal Processing*, 2009:1–11.

[Kumar *et al.*, 2010] KUMAR, G., TULYAKOV, S. et GOVINDARAJU, V. (2010). Combination of symmetric hash functions for secure fingerprint matching. *In 2010 20th International Conference on Pattern Recognition*, pages 890–893. IEEE.

[Lahmidi *et al.*, 2021] LAHMIDI, A., MINAOUI, K., MOUJAHDI, C. et RZIZA, M. (2021). Fingerprint template protection using irreversible minutiae tetrahedrons. *The Computer Journal*.

[Lahmidi *et al.*, 2022] LAHMIDI, A., MOUJAHDI, C., MINAOUI, K. et RZIZA, M. (2022). On the methodology of fingerprint template protection schemes conception : meditations on the reliability. *EURASIP Journal on Information Security*, 2022(1):1–13.

[Lee *et al.*, 2007a] LEE, C., CHOI, J.-Y., TOH, K.-A., LEE, S. et KIM, J. (2007a). Alignment-free cancelable fingerprint templates based on local minutiae information. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(4):980–992.

[Lee *et al.*, 2007b] LEE, Y. J., BAE, K., LEE, S. J., PARK, K. R. et KIM, J. (2007b). Biometric key binding : Fuzzy vault based on iris images. *In International Conference on Biometrics*, pages 800–808. Springer.

[Li *et al.*, 2008] LI, Q., GUO, M. et CHANG, E.-C. (2008). Fuzzy extractors for asymmetric biometric representations. *In 2008 IEEE computer society conference on computer vision and pattern recognition workshops*, pages 1–6. IEEE.

[Lumini et Nanni, 2006] LUMINI, A. et NANNI, L. (2006). A clustering method for automatic biometric template selection. *Pattern Recognition*, 39(3):495–497.

[Lumini et Nanni, 2007] LUMINI, A. et NANNI, L. (2007). An improved biohashing for human authentication. *Pattern recognition*, 40(3):1057–1065.

[Ma *et al.*, 2004] MA, L., TAN, T., WANG, Y. et ZHANG, D. (2004). Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image processing*, 13(6):739–750.

[Maio *et al.*, 2002] MAIO, D., MALTONI, D., CAPPELLI, R., WAYMAN, J. L. et JAIN, A. K. (2002). Fvc2000 : Fingerprint verification competition. *IEEE transactions on pattern analysis and machine intelligence*, 24(3):402–412.

[Maiorana et Campisi, 2009] MAIORANA, E. et CAMPISI, P. (2009). Fuzzy commitment for function based signature template protection. *IEEE signal processing letters*, 17(3): 249–252.

[Maiorana *et al.*, 2010] MAIORANA, E., CAMPISI, P., FIERREZ, J., ORTEGA-GARCIA, J. et NERI, A. (2010). Cancelable templates for sequence-based biometrics with application to on-line signature recognition. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 40(3):525–538.

[Maiorana *et al.*, 2008a] MAIORANA, E., CAMPISI, P. et NERI, A. (2008a). User adaptive fuzzy commitment for signature template protection and renewability. *Journal of Electronic Imaging*, 17(1):011011.

[Maiorana *et al.*, 2008b] MAIORANA, E., CAMPISI, P., ORTEGA-GARCIA, J. et NERI, A. (2008b). Cancelable biometrics for hmm-based signature recognition. *In 2008 IEEE*

Second International Conference on Biometrics : Theory, Applications and Systems, pages 1–6. IEEE.

[Maiorana *et al.*, 2008c] MAIORANA, E., MARTINEZ-DIAZ, M., CAMPISI, P., ORTEGA-GARCIA, J. et NERI, A. (2008c). Template protection for hmm-based on-line signature authentication. *In 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–6. IEEE.

[Maltoni *et al.*, 2009] MALTONI, D., MAIO, D., JAIN, A. K. et PRABHAKAR, S. (2009). *Handbook of fingerprint recognition*. Springer Science & Business Media.

[Martinez-Diaz *et al.*, 2011] MARTINEZ-DIAZ, M., FIERREZ, J., GALBALLY, J. et ORTEGA-GARCIA, J. (2011). An evaluation of indirect attacks and countermeasures in fingerprint verification systems. *Pattern Recognition Letters*, 32(12):1643–1651.

[Matsumoto *et al.*, 2002] MATSUMOTO, T., MATSUMOTO, H., YAMADA, K. et HOSHINO, S. (2002). Impact of artificial" gummy" fingers on fingerprint systems. *In Optical Security and Counterfeit Deterrence Techniques IV*, volume 4677, pages 275–289. International Society for Optics and Photonics.

[Matyáš et Rıha, 2000] MATYÁŠ, V. et RIHA, Z. (2000). Biometric authentication systems. *In verfügbar über : http ://grover. informatik. uni-augsburg. de/lit/MM-Seminar/Privacy/riha00biometric. pdf*. Citeseer.

[Modi, 2011] MODI, S. K. (2011). *Biometrics in identity management : Concepts to applications*. Artech House.

[Mohanty *et al.*, 2007] MOHANTY, P., SARKAR, S. et KASTURI, R. (2007). From scores to face templates : a model-based approach. *IEEE transactions on pattern analysis and machine intelligence*, 29(12):2065–2078.

[Monrose *et al.*, 2000] MONROSE, F., REITER, M. K., LI, Q. et WETZEL, S. (2000). Cryptographic key generation from voice. *In Proceedings 2001 IEEE Symposium on Security and Privacy. S&P 2001*, pages 202–213. IEEE.

[Moujahdi *et al.*, 2014] MOUJAHDI, C., BEBIS, G., GHOUZALI, S. et RZIZA, M. (2014). Fingerprint shell : Secure representation of fingerprint template. *Pattern Recognition Letters*, 45:189–196.

[Muramatsu, 2008] MURAMATSU, D. (2008). Online signature verification algorithm using hill-climbing method. *In 2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, volume 2, pages 133–138. IEEE.

[Nagar *et al.*, 2010] NAGAR, A., NANDAKUMAR, K. et JAIN, A. K. (2010). Biometric template transformation : a security analysis. *In Media Forensics and Security II*, volume 7541, page 75410O. International Society for Optics and Photonics.

[Nandakumar *et al.*, 2007a] NANDAKUMAR, K., JAIN, A. K. et PANKANTI, S. (2007a). Fingerprint-based fuzzy vault : Implementation and performance. *IEEE transactions on information forensics and security*, 2(4):744–757.

[Nandakumar *et al.*, 2007b] NANDAKUMAR, K., NAGAR, A. et JAIN, A. K. (2007b). Hardening fingerprint fuzzy vault using password. *In International conference on Biometrics*, pages 927–937. Springer.

[Neurotechnology, 2010] NEUROTECHNOLOGY (2010). Verifinger sdk. http://www.neurotechnology.com/megamatcher.html.

[Nyang et Lee, 2007] NYANG, D. et LEE, K. (2007). Fuzzy face vault : how to implement fuzzy vault with weighted features. *In International Conference on Universal Access in Human-Computer Interaction*, pages 491–496. Springer.

[Ong *et al.*, 2008] ONG, T. S., JIN, A. T. B. et NGO, D. C. L. (2008). Application-specific key release scheme from biometrics. *Int. J. Netw. Secur.*, 6(2):127–133.

[Patel *et al.*, 2015] PATEL, R. N., CHAUHAN, S. P., PANWALA, K. C., PRAJAPATI, H. D. et KARIYA, S. L. (2015). Dynamic signature recognition and verification using pixel based approach. *International Journal of Computer Science and Information Technologies*, 6(2):1497–1499.

[Pötzsch *et al.*, 1996] PÖTZSCH, M., MAURER, T., WISKOTT, L. et vd MALSBURG, C. (1996). Reconstruction from graphs labeled with responses of gabor filters. *In International Conference on Artificial Neural Networks*, pages 845–850. Springer.

[Quan *et al.*, 2008] QUAN, F., FEI, S., ANNI, C. et FEIFEI, Z. (2008). Cracking cancelable fingerprint template of ratha. *In 2008 International Symposium on Computer Science and Computational Technology*, volume 2, pages 572–575. IEEE.

[Ratha *et al.*, 2006] RATHA, N., CONNELL, J., BOLLE, R. M. et CHIKKERUR, S. (2006). Cancelable biometrics : A case study in fingerprints. *In 18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 370–373. IEEE.

[Ratha *et al.*, 1995] RATHA, N. K., CHEN, S. et JAIN, A. K. (1995). Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition*, 28(11): 1657–1672.

[Ratha *et al.*, 2007] RATHA, N. K., CHIKKERUR, S., CONNELL, J. H. et BOLLE, R. M. (2007). Generating cancelable fingerprint templates. *IEEE Transactions on pattern analysis and machine intelligence*, 29(4):561–572.

[Ratha *et al.*, 2001a] RATHA, N. K., CONNELL, J. H. et BOLLE, R. M. (2001a). An analysis of minutiae matching strength. *In International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 223–228. Springer.

[Ratha *et al.*, 2001b] RATHA, N. K., CONNELL, J. H. et BOLLE, R. M. (2001b). Enhancing security and privacy in biometrics-based authentication systems. *IBM systems Journal*, 40(3):614–634.

[Ratha *et al.*, 2003] RATHA, N. K., CONNELL, J. H. et BOLLE, R. M. (2003). Biometrics break-ins and band-aids. *Pattern Recognition Letters*, 24(13):2105–2113.

[Rathgeb et Uhl, 2009] RATHGEB, C. et UHL, A. (2009). Systematic construction of iris-based fuzzy commitment schemes. *In International Conference on Biometrics*, pages 940–949. Springer.

[Rathgeb et Uhl, 2010] RATHGEB, C. et UHL, A. (2010). Secure iris recognition based on local intensity variations. *In International Conference Image Analysis and Recognition*, pages 266–275. Springer.

[Rathgeb et Uhl, 2011] RATHGEB, C. et UHL, A. (2011). A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security*, 2011(1):1–25.

[Roberts, 2007] ROBERTS, C. (2007). Biometric attack vectors and defences. *Computers & Security*, 26(1):14–25.

[Ross *et al.*, 2007] ROSS, A., SHAH, J. et JAIN, A. K. (2007). From template to image : Reconstructing fingerprints from minutiae points. *IEEE transactions on pattern analysis and machine intelligence*, 29(4):544–560.

[Ross *et al.*, 2006] ROSS, A. A., NANDAKUMAR, K. et JAIN, A. K. (2006). *Handbook of multibiometrics*, volume 6. Springer Science & Business Media.

[Ross *et al.*, 2005] ROSS, A. A., SHAH, J. et JAIN, A. K. (2005). Toward reconstructing fingerprints from minutiae points. *In Biometric Technology for Human Identification II*, volume 5779, pages 68–80. International Society for Optics and Photonics.

[Sakata *et al.*, 2006] SAKATA, K., MAEDA, T., MATSUSHITA, M., SASAKAWA, K. et TAMAKI, H. (2006). Fingerprint authentication based on matching scores with other data. *In International Conference on Biometrics*, pages 280–286. Springer.

[Sandhya et Prasad, 2015] SANDHYA, M. et PRASAD, M. V. (2015). k-nearest neighborhood structure (k-nns) based alignment-free method for fingerprint template protection. *In 2015 International Conference on Biometrics (ICB)*, pages 386–393. IEEE.

[Sandhya *et al.*, 2016] SANDHYA, M., PRASAD, M. V. et CHILLARIGE, R. R. (2016). Generating cancellable fingerprint templates based on delaunay triangle feature set construction. *IET Biometrics*, 5(2):131–139.

[Savvides *et al.*, 2004] SAVVIDES, M., KUMAR, B. V. et KHOSLA, P. K. (2004). Cancelable biometric filters for face recognition. *In Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 3, pages 922–925. IEEE.

[Shin *et al.*, 2009] SHIN, S. W., LEE, M.-K., MOON, D. et MOON, K. (2009). Dictionary attack on functional transform-based cancelable fingerprint templates. *ETRI journal*, 31(5):628–630.

[Soutar, 1999] SOUTAR, C. (1999). Biometric system performance and security. *IEEE Auto. Identification Advanced Technol.*

[Sutcu *et al.*, 2007] SUTCU, Y., LI, Q. et MEMON, N. (2007). Protecting biometric templates with sketch : Theory and practice. *IEEE Transactions on Information Forensics and Security*, 2(3):503–512.

[Sutcu *et al.*, 2005] SUTCU, Y., SENCAR, H. T. et MEMON, N. (2005). A secure biometric authentication scheme based on robust hashing. *In Proceedings of the 7th Workshop on Multimedia and Security*, pages 111–116.

[Syverson, 1994] SYVERSON, P. (1994). A taxonomy of replay attacks [cryptographic protocols]. *In Proceedings The Computer Security Foundations Workshop VII*, pages 187–191. IEEE.

[Teoh *et al.*, 2006] TEOH, A. B., GOH, A. et NGO, D. C. (2006). Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE transactions on pattern analysis and machine intelligence*, 28(12):1892–1901.

[Teoh *et al.*, 2007] TEOH, A. B. J., TOH, K.-A. et YIP, W. K. (2007). 2ˆ n discretisation of biophasor in cancellable biometrics. *In International Conference on Biometrics*, pages 435–444. Springer.

[Testoni et Kirovski, 2010] TESTONI, V. et KIROVSKI, D. (2010). On the inversion of biometric templates by an example. *In 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1830–1833. IEEE.

[Thieme, 2003] THIEME, M. (2003). Identifying and reducing privacy risks in biometric systems. *In 13 th Annual Conference on Computers, Freedom & Privacy< http ://www. biometricgroup. com.*

[Tran *et al.*, 2018] TRAN, Q. N., HU, J. et WANG, S. (2018). Alignment-free cancellable template with clustered-minutiae local structure. *In 2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE.

[Tuyls *et al.*, 2005] TUYLS, P., AKKERMANS, A. H., KEVENAAR, T. A., SCHRIJEN, G.-J., BAZEN, A. M. et VELDHUIS, R. N. (2005). Practical biometric authentication with template protection. *In International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 436–446. Springer.

[Tuyls *et al.*, 2004] TUYLS, P. T., VERBITSKIY, E., IGNATENKO, T., SCHOBBEN, D. et AKKERMANS, T. H. (2004). Privacy-protected biometric templates : Acoustic ear identification. *In Biometric Technology for Human Identification*, volume 5404, pages 176–182. International Society for Optics and Photonics.

[Uludag et Jain, 2004a] ULUDAG, U. et JAIN, A. K. (2004a). Attacks on biometric systems : a case study in fingerprints. *In Security, steganography, and watermarking of multimedia contents VI*, volume 5306, pages 622–633. International Society for Optics and Photonics.

[Uludag et Jain, 2004b] ULUDAG, U. et JAIN, A. K. (2004b). Fuzzy fingerprint vault. *In Proc. Workshop : Biometrics : Challenges arising from theory to practice*, pages 13–16.

[Uludag *et al.*, 2004] ULUDAG, U., PANKANTI, S., PRABHAKAR, S. et JAIN, A. K. (2004). Biometric cryptosystems : issues and challenges. *Proceedings of the IEEE*, 92(6):948–960.

[Vizcaya et Gerhardt, 1996] VIZCAYA, P. R. et GERHARDT, L. A. (1996). A nonlinear orientation model for global description of fingerprints. *Pattern Recognition*, 29(7): 1221–1231.

[Voderhobli *et al.*, 2006] VODERHOBLI, K., PATTINSON, C. et DONELAN, H. (2006). A schema for cryptographic keys generation using hybrid biometrics. *In 7th annual postgraduate symposium : The convergence of telecommunications, networking and broadcasting*.

[Wang et Hu, 2012] WANG, S. et HU, J. (2012). Alignment-free cancelable fingerprint template design : A densely infinite-to-one mapping (ditom) approach. *Pattern Recognition*, 45(12):4129–4137.

[Wang et Hu, 2016] WANG, S. et HU, J. (2016). A blind system identification approach to cancelable fingerprint templates. *Pattern Recognition*, 54:14–22.

[Wang *et al.*, 2017] WANG, S., YANG, W. et HU, J. (2017). Design of alignment-free cancelable fingerprint templates with zoned minutia pairs. *Pattern Recognition*, 66:295–301.

[Wang et Plataniotis, 2007] WANG, Y. et PLATANIOTIS, K. (2007). Face based biometric authentication with changeable and privacy preservable templates. *In 2007 Biometrics Symposium*, pages 1–6. IEEE.

[Wilson *et al.*, 2004] WILSON, C., HICKLIN, A., BONE, M., KORVES, H., GROTHER, P., ULERY, B., MICHEALS, R., ZOEPFL, M., OTTO, S. et WATSON, C. (2004). Fingerprint vendor technology evaluation 2003 : Summary of results and analysis report. *NIST Technical Report NISTIR*, 7123.

[Wong *et al.*, 2014] WONG, W. J., WONG, M. D. et TEOH, A. B. J. (2014). A security- and privacy-driven hybrid biometric template protection technique. *In 2014 International Conference on Electronics, Information and Communications (ICEIC)*, pages 1–5. IEEE.

[Xi et Hu, 2009] XI, K. et HU, J. (2009). Biometric mobile template protection : a composite feature based fingerprint fuzzy vault. *In 2009 IEEE International Conference on Communications*, pages 1–5. IEEE.

[Yamazaki *et al.*, 2005] YAMAZAKI, Y., NAKASHIMA, A., TASAKA, K. et KOMATSU, N. (2005). A study on vulnerability in on-line writer verification system. *In Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, pages 640–644. IEEE.

[Yang *et al.*, 2009] YANG, H., JIANG, X. et KOT, A. C. (2009). Generating secure cancelable fingerprint templates using local and global features. *In 2009 2nd IEEE International Conference on Computer Science and Information Technology*, pages 645–649. IEEE.

[Yang et Verbauwhede, 2005] YANG, S. et VERBAUWHEDE, I. (2005). Automatic secure fingerprint verification system based on fuzzy vault scheme. *In Procee-*

dings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005., volume 5, pages v–609. IEEE.

[Yang *et al.*, 2012] YANG, W., HU, J. et WANG, S. (2012). A delaunay triangle-based fuzzy extractor for fingerprint authentication. *In 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, pages 66–70. IEEE.

[Yang *et al.*, 2014] YANG, W., HU, J., WANG, S. et STOJMENOVIC, M. (2014). An alignment-free fingerprint bio-cryptosystem based on modified voronoi neighbor structures. *Pattern Recognition*, 47(3):1309–1320.

[Yang *et al.*, 2013] YANG, W., HU, J., WANG, S. et YANG, J. (2013). Cancelable fingerprint templates with delaunay triangle-based local structures. *In International Symposium on Cyberspace Safety and Security*, pages 81–91. Springer.

[Ye *et al.*, 2016] YE, Y., ZHENG, H., NI, L., LIU, S. et LI, W. (2016). A study on the individuality of finger vein based on statistical analysis. *In 2016 international conference on biometrics (ICB)*, pages 1–5. IEEE.

[Zhou, 2007] ZHOU, X. (2007). Template protection and its implementation in 3d face recognition systems. *In Biometric technology for human identification IV*, volume 6539, page 65390L. International Society for Optics and Photonics.

[Zhu *et al.*, 2012] ZHU, H.-H., HE, Q.-H. et LI, Y.-X. (2012). A two-step hybrid approach for voiceprint-biometric template protection. *In 2012 International Conference on Machine Learning and Cybernetics*, volume 2, pages 560–565. IEEE.

[Zuo *et al.*, 2008] ZUO, J., RATHA, N. K. et CONNELL, J. H. (2008). Cancelable iris biometric. *In 2008 19th International conference on pattern recognition*, pages 1–4. IEEE.