REPUBLIQUE ALGERIENNE
DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT
SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ LARBI TEBESSI - TÉBESSA
FACULTÉ DES SCIENCES ET TECHNOLOGIES
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE

MEMOIRE

DE FIN D'ÉTUDES POUR L'OBTENTION DU DIPLOME DE MASTER EN

RÉSEAUX ET TÉLÉCOMMUNICATIONS

THEME

Renforcement de la sécurité des systèmes biométriques à l'aide des caractéristiques profondes de la biométrie de la main

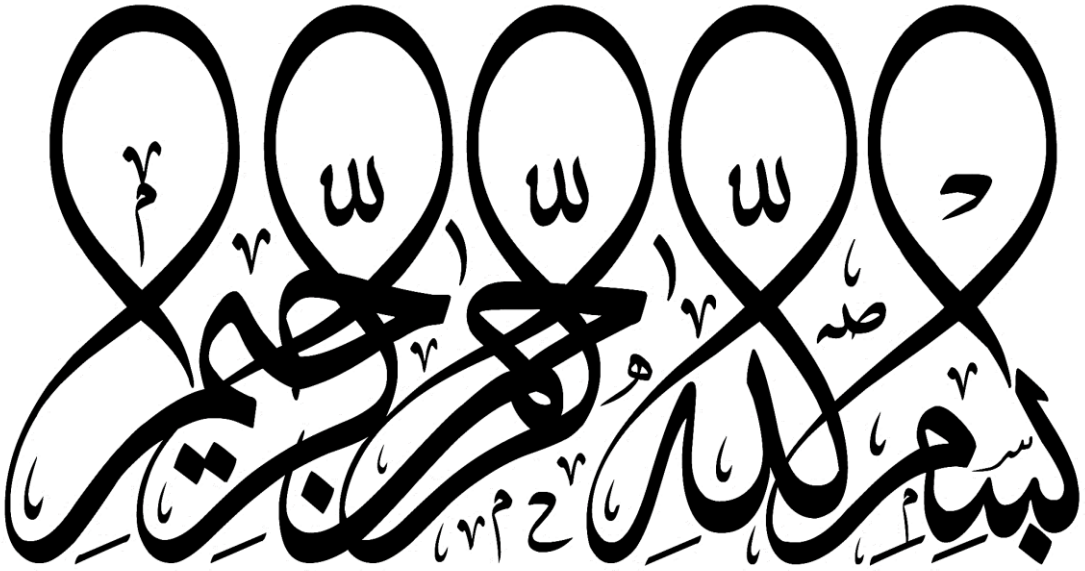
Présenté par le binôme :

- **Saoussen Djeddi**
- **Fatma Zahra Mahdjoub**

Devant le jury :

- **Tarek Bentahar**	**Président**
- **Abdallah Meraoumia**	**Encadreur**
- **Riad Saidi**	**Examinateur**

Année Universitaire 2019 / 2020


بسم الله الرحمن الرحيم

À mes parents qui n'a jamais dit non à mes exigences, pour votre confiance et vos prières.

À mes deux frères, ma sœur et le petit -Ayan- pour votre présence permanente et votre encouragement.

À mes proches, mes chères et à ceux qui me donnent de l'amour, de la motivation et de la vivacité.

À mes amies et compagnons de route que j'ai connu jusqu'à maintenant et qui ont partagé avec moi tous les moments lors durant mes d'études.

À mon binôme -Fatma- pour son soutien moral et sa compréhension tout au long de cette travail.

À mon chat -QQ- pour rester éveillé toutes les nuits me donnant compagnie.

À tous ceux que j'aime.

Djeddi Saoussen

À la plus forte et belle femme, ma mère pour sa tendresse sa douceur sa patience et sa volonté que dieu la protège.

À mon père pour tous les sacrifices.

À mes beaux-frères -Mohamed- et le petit -Youssef El-Amine-.

À ma belle-sœur -Nour-.

À toute ma famille (mes oncles, mes tantes, mes cousins et mes cousines) pour leur soutien tout au long de mon parcours universitaire.

À mes chères amies et mes collègues et toute la promotion de réseaux et télécommunication 2015-2020 pour leurs encouragements.

À ma courageuse binôme -Saoussen- pour la douce amitié et sa présence.

Merci à tous et toutes.

Fatma Zahra Mahdjoub

Remerciement

Avant tout, nous remercions notre Dieu -Allah- pour tout ce qu'il nous a donné.

Nous adressons nos remercîments les plus chaleureuse à l'ensemble des membres du jury :

Dr. Tarek BENTAHAR, d'avoir accepté de présider ce jury,

Dr. Riad SAIDI, d'avoir accepté d'examiner ce travail,

Nous tenons à exprimer notre gratitude à notre directeur de la mémoire,

Dr. Abdallah MERAOUMIA. Nous le remercions de nous encadrer, de nous avoir dirigés, aidés et conseillés.

Nos remercîments les plus sincères, à

Monsieur Ramzi SGHIR, que dieu ait pitié de lui.

Tous les professeurs, collègues, toutes les amis et les personnes qui par leurs paroles, leurs écrits, leurs conseils, leurs critiques, leur soutien moral et intellectuel ont guidé nos réflexions durant nos recherches.

S. Djeddi et F. Mahdjoub

Table des Matières

Chapitre II : Protection de modèle biométrique

Chapitre III : Résultats expérimentaux

Liste des Figures

Liste des tableaux

Introduction Générale

Introduction

RECEMMENT, l'émergence de systèmes informatiques, qui couvrent aujourd'hui presque toute notre vie moderne, a entraîné de nouveaux défis pour la sécurité des données des utilisateurs pour la confidentialité et la vie privée des individus, ce qui augmente sans aucun doute le niveau de confiance des utilisateurs. Malheureusement, plusieurs études montrent que les taux de vol de données, en particulier sur Internet, ont considérablement augmenté ces dernières années, et les systèmes de sécurité sont souvent incapables de fournir la protection nécessaire pour les données des utilisateurs. Certes, les moyens traditionnels, tels que ceux basés sur la possession (ex. cartes à puce) et la connaissance (ex. mots de passe), qui sont les moyens de vérification d'identité les plus utilisés, présentent plusieurs vulnérabilités qui pourraient permettre à des personnes non autorisées de voler les données des utilisateurs et donc de les utiliser illégalement [1]. Plus récemment, la biométrie est devenue une alternative aux méthodes de sécurité traditionnelles, conduisant à l'émergence de plusieurs systèmes automatiques de reconnaissance humaine qui utiliseront des traits humains physiques, comportementaux ou biologiques. Plus précisément, la biométrie est l'un des moyens nouveaux et ambitieux les plus importants pour fournir une protection supplémentaire aux données des utilisateurs, car elles ne peuvent être perdues, oubliées ou volées [2].

Les technologies biométriques se sont avérées efficaces dans les systèmes de sécurité traditionnels, ce qui a incité de nombreux secteurs à intégrer cette technologie dans de nombreuses applications nécessitant un haut niveau de sécurité, mais malgré cela, son utilisation présente encore de nombreux défis, en particulier ceux liés à la sécurité et à la

confidentialité [3]. En général, la biométrie présente des faiblesses importantes qu'il faut prendre en compte avant de les utiliser. En effet, contrairement aux méthodes d'authentification basées sur les connaissances, qui sont totalement confidentielles car le mot de passe est uniquement connu de l'utilisateur, la biométrie est entièrement accessible, par exemple, la voix de l'utilisateur peut être enregistrée, sans son autorisation, puis réutilisée. D'autre part, bien que l'unicité de la biométrie soit une caractéristique très importante, elle peut causer beaucoup de problèmes au fil du temps lorsque nous voulons les changer en cas de vol. Puisqu'une personne possède les mêmes données biométriques qu'elle peut utiliser dans de nombreuses applications, il est très probable que ces données soient volées si l'une de ces applications n'est pas suffisamment sécurisée. Par exemple, si les données biométriques sont compromises dans une application spécifique, toutes les autres applications sont également vulnérables au piratage et, par conséquent, la confidentialité est toujours menacée. Les systèmes biométriques sont relativement coûteux par rapport aux autres systèmes, et comme le piratage de modèles biométriques, extraits par une méthode d'extraction de caractéristiques donnée, nécessite une restructuration de ces systèmes, il peut être très coûteux en particulier pour les petites et moyennes entreprises [4]. Sérieusement, les systèmes biométriques ne peuvent pas traiter les données biométriques volées ou falsifiées et empêcher ainsi diverses attaques, ce qui les rend peu fiables pour la majorité du public.

Afin de sécuriser les données biométriques, plusieurs approches sont proposées dans la littérature, dont les plus importantes sont celles qui utilisent la cryptographie (protection des modèles biométriques) et les transformations de données (transformation irréversible de modèles ou simplement biométrie révocable) [5]. Ainsi, la sécurité dans la première approche reste faible car si la clé de chiffrement est récupérée illégalement, le modèle biométrique peut être déchiffré puis volé [6]. Heureusement, dans la seconde approche, il est impossible de reconstruire le modèle biométrique d'origine car ce modèle a été transformé avec une fonction irréversible, donc, en cas de piratage de données biométriques, il est possible de changer la fonction de transformation pour annuler complètement l'ancien modèle afin qu'il ne soit pas utile. En biométrie révocable et après transformation irréversible du modèle, les performances du système biométrique peuvent être dégradées en utilisant de nouvelles images de modèles [7]. Ainsi, la fonction de transformation doit être soigneusement choisie pour maintenir autant que possible les performances du système. Pour cela, notre réflexion porte sur la manière de choisir une fonction de transformation qui maintient cette performance tout en garantissant une haute sécurité ; un problème qui tournera autour de l'hypothèse suivante : au lieu d'utiliser

la fonction de transformation pour transformer le modèle final (obtenu par une méthode d'extraction de caractéristiques appropriée), il est possible d'impliquer la fonction de transformation dans la tâche d'extraction de caractéristiques.

Dans ce mémoire, nous proposerons une nouvelle méthode d'extraction de caractéristiques biométriques révocables (Cancelable Principal Component Analysis Network (C-PCANet)). Pour que ces caractéristiques soient efficaces et capables de distinguer les personnes, nous avons adopté une approche basée sur une analyse profonde (*Deep Learning*). Par conséquent, notre méthode repose sur l'analyse de l'image selon plusieurs niveaux de convolution et à la fin de ces niveaux, nous projetons le résultat dans un espace irréversible afin d'extraire le vecteur de caractéristiques (modèle biométrique ou Template), et ceci après avoir réduit la taille des données. Pour créer la matrice de projection (utilisée dans la transformation), nous avons utilisé les systèmes chaotiques qui nous permettent d'utiliser une clé secrète pour générer une séquence de valeurs que nous utilisons pour créer une matrice orthogonale. En effet, notre méthode proposée est capable d'extraire des modèles biométriques profonds et révocables (annulables) pour garantir à la fois des performances élevées et une sécurité renforcée. Afin d'évaluer les performances du système proposé, deux modalités biométriques efficaces ont été utilisées, à savoir les empreintes palmaires et les empreintes de veine de paume, en raison d'une part de leur grande acceptabilité par les personnes et d'autre part de l'existence d'une grande base de données disponible et connue.

Ce mémoire est organisé en trois chapitres : le **premier chapitre** présente un aperçu général de la biométrie, traitant en détail les différentes modalités biométriques ainsi que la structure et le fonctionnement des systèmes biométriques. Le **deuxième chapitre** se concentre sur la biométrie révocable. Ce chapitre contient la méthode proposée (C-PCANet), où elle est présentée en détail, ainsi que quelques prérequis sur lesquels le système a été construit. Le **troisième chapitre** présente les résultats expérimentaux liés à la fois aux performances du système biométrique et à l'analyse de sécurité. Enfin, une conclusion générale avec les perspectives que nous envisagerons sont données à la fin de ce mémoire.

Chapitre 1

Sécurité des systèmes biométriques

Résumé

Depuis plusieurs années, l'identité des individus est vérifiée, pour l'accès physique et/ou logique, par des méthodes traditionnelles telles que badges, cartes d'identité, mots de passe, etc. Cependant, ces méthodes présentent certains inconvénients lorsqu'un mot de passe est deviné ou qu'un badge a été volé par un fraudeur. Heureusement, la biométrie a réussi à résoudre ces problèmes; elle est utilisée pour identifier l'identité d'une personne à l'aide de ses caractéristiques physiques, biologiques ou comportementales. Ces caractéristiques qui permettent l'identification des personnes sont appelées des modalités biométriques.

Introduction

La biométrie est une alternative efficace aux méthodes traditionnelles d'identification des personnes, car elles dépendent de leurs caractéristiques intrinsèques représentées par leurs traits biologiques, physiques ou comportementales. Cette technologie a fait ses preuves, en particulier avec l'augmentation récente des problèmes de confidentialité et de sécurité. Dans ce chapitre, nous discuterons du concept de biométrie, où nous nous concentrerons sur les types de modalités biométriques qui existent et la structure et le fonctionnement du système biométrique. De plus, nous présenterons les exigences et les risques qui menacent ces systèmes et nous conclurons tout cela par une analyse de son évaluation de la performance.

I.1 Biométrie

I.1.1 Reconnaissance d'individu

La biométrie est la science qui consiste à établir l'identité d'un individu [8]. La reconnaissance biométrique fait référence à la reconnaissance automatique des individus en fonction de leurs caractéristiques physiologiques et comportementales, c'est-à-dire l'utilisation de ces données biométriques pour confirmer ou identifier l'identité d'une personne [9]. Dans la littérature, il existe d'autres définitions de la biométrie telles que :

La reconnaissance automatique d'une personne à partir de son comportement ou d'une caractéristique physique [10]**.**

La biométrie recouvre l'ensemble des procédés tendant à identifier un individu à partir de la mesure de l'une ou de plusieurs de ses caractéristiques physiques, comportementales ou biologiques [11]**.**

I.1.2 Objectif de la biométrie

Plusieurs raisons peuvent motiver l'utilisation de la biométrie [12] :

- **Haute sécurité :** Associée à d'autres technologies telles que le cryptage, certains systèmes rendent toute tentative de fraude très compliquée.
- **Confort :** Il présente le remplacement des méthodes traditionnelles, par exemple, un mot de passe, la biométrie permet de respecter les règles de base de sécurité. Et lorsque ces règles sont respectées, la biométrie évite aux administrateurs d'avoir répondre aux nombreuses demandes de changement de mot de passe.
- **Sécurité/Psychologie :** C'est la garantie de systèmes efficaces, intégrés et hautement dissuasifs. S'il est techniquement facile de découvrir un mot de passe ou de se procurer frauduleusement un badge d'accès ou une carte magnétique, il est presque impossible de modeler, voler ou copier une caractéristique biométrique physiologique ou comportementale humaine.

I.2 Modalités biométriques

I.2.1 Biométries physiologiques (morphologiques)

Aussi appelées biométrie physique, elles reposent sur l'identification de traits morphologiques particuliers obtenus à partir de plusieurs parties du corps humain.

☑ **Empreinte digitale :** La reconnaissance d'empreintes digitales (en anglais *Fingerprintings*) est la technologie biométrique la plus ancienne et la plus mature (voir Fig. I.1). Cependant, elle est considérée comme une technologie inacceptable par les utilisateurs en raison de son association avec la criminologie.

Fig I.1 : Empreinte digitale

Les empreintes digitales sont formées par les crêtes (*ridge*) et les vallées (*furrow*) présentes à la surface du bout des doigts. Les empreintes digitales ne sont pas totalement déterminées par la génétique puisque même les jumeaux monozygotes ont des empreintes

digitales différentes. Les empreintes digitales sont différentes pour chaque doigt d'une même personne. Il existe de nombreuses méthodes pour acquérir cette empreinte, dont la plus ancienne consiste à recouvrir le bout du doigt d'une fine couche d'encre et à l'imprimer sur une feuille de papier. L'empreinte ainsi imprimée peut ensuite être numérisée. Les dispositifs d'acquisition d'empreintes digitales sont basés sur la capture optique, thermique, électromagnétique ou ultrasonore [13].

☑ **Empreinte palmaire :** Il s'agit d'une technologie biométrique qui utilise la surface interne de la paume pour l'identification et/ou la vérification des personnes, appelée en anglais *Palmprint*. Il est bien adapté aux systèmes de sécurité moyenne tels que le contrôle d'accès physique et/ou logique [14].

Fig I.2 : Empreinte palmaire.

☑ **Empreintes des articulations des doigts :** Cette technologie, en anglais *Finger Knuckle Print (FKP)*, est basée sur la surface du doigt arrière (voir Fig. I.3) car elle contient des caractéristiques distinctives, telles que des lignes principales, des lignes secondaires et des crêtes, qui peuvent être extraites d'images à basse résolution. La main contient plusieurs doigts, pour cela, il est nécessaire de combiner les informations sur chaque doigt pour une reconnaissance précise dans le domaine de l'identification biométrique [14].

Fig I.3 : Empreintes des articulations des doigts.

☑ **Visage :** Le visage est certainement la caractéristique biométrique que les humains utilisent le plus naturellement pour s'identifier entre eux, ce qui peut expliquer pourquoi il est généralement très bien accepté par les utilisateurs. Le système d'acquisition est soit un appareil photo, soit une caméra numérique [15]. Selon le système utilisé, l'individu doit être positionné devant l'appareil ou peut-être en mouvement à distance. Les données biométriques qui sont obtenues sont par la suite comparées au fichier référence [16].

Fig I.4 : Visage.

☑ **Iris :** La reconnaissance de l'iris (voir Fig. I.5) est une technologie plus récente car elle ne s'est vraiment développée que dans les années 1980 [18]. L'iris est la région annulaire entre la pupille et le blanc de l'œil.

Fig I.5 : Iris.

Les iris sont uniques et les deux iris d'un même individu sont différents. Pour le moment, il n'est pas modifiable par chirurgie [15]. L'image de l'iris est capturée par un appareil qui contient une caméra infrarouge lorsque la personne se place à une courte distance de l'appareil et une fois l'image de la texture de l'iris obtenue par le système biométrique, le fonctionnement est identique à celui du système analyser l'empreinte digitale [16].

☑ **Rétine :** La reconnaissance de la rétine (en anglais *retina*) est une méthode assez ancienne puisque les premières études remontent aux années 1930. Les motifs formés par les veines (voir Fig. I.6) sous la surface de la rétine sont uniques et stables dans le temps.

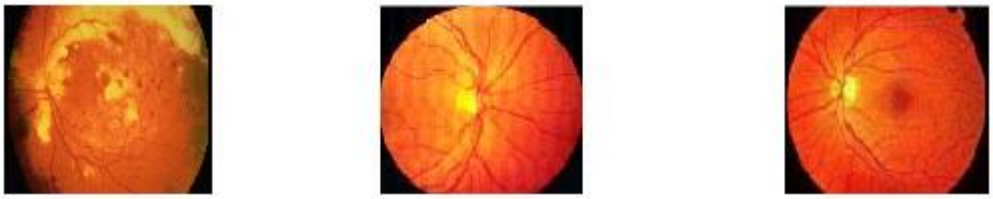


Fig I.6 : La rétine.

Ils ne peuvent être affectés que par certaines maladies. Pour ces raisons, la reconnaissance de la rétine est actuellement considérée comme l'une des méthodes biométriques les plus sûres. Cette technologie est mal acceptée par les utilisateurs, en raison de son caractère trop contraignant : la mesure doit être prise à une très courte distance du capteur pour que le

capteur balaye ensuite la rétine. Il est physiquement impossible de prendre une mesure rétinienne de plus de trente centimètres sur un sujet mobile [15].

I.2.2 Biométries biologiques

Ces types de modalités reposent sur l'étude de traces biologiques particulières.

☑ **A.D.N :** Présent dans les cellules du corps, il est spécifique d'un individu à l'autre et permet de l'identifier avec certitude à partir d'un simple fragment de peau, d'une trace de sang ou d'une goutte de salive.

Fig I.7 : ADN.

Ce procédé représente la technologie d'identification par excellence avec une marge d'erreur bien en dessous des autres moyens biométriques. C'est le moyen le plus précis de déterminer l'identité de la personne. Cette modalité a l'avantage d'être unique et permanente tout au long de la vie [19].

☑ **Thermogramme faciale :** C'est la quantité de chaleur émise par les différentes parties du visage caractérise chaque individu (voir Fig. I.8). Cela dépend de la localisation des veines mais aussi de la quantité de tissu, de muscles, de graisse, de l'épaisseur du squelette, etc.

Fig I.8 : Thermogramme facial.

Pour capturer l'image, il est possible d'utiliser un appareil photo ou une caméra numérique dans le domaine de l'infrarouge [15]. Il peut donc être utilisé même dans l'obscurité ou dans des conditions de mauvaise visibilité. Mais les conditions de prise de vue peuvent conduire à des erreurs [20].

☑ **Empreinte veineuse (veines de la main) :** Les veines de la paume de la main, en anglais *palm-vein*, font partie du réseau sanguin (voir Fig. I.9), qui diffère d'une personne à l'autre, et pour cette raison, ce réseau peut être analysé et les différences peuvent être exploitées pour différencier les personnes [19].

Fig I.9 : Veines de la main.

I.2.3 Biométries comportementales

Ces types de modalités sont basés sur l'analyse de certains comportements d'une personne.

☑ **Voix :** La reconnaissance du locuteur vise à déterminer les caractéristiques uniques de la voix de chaque individu. Bien que généralement classée comme une caractéristique comportementale, la voix (voir Fig. I.10) se trouve à la frontière avec les caractéristiques physiques.

Fig I.10 : La voix.

En effet, une grande partie de cette caractéristique est déterminée par le tractus vocal ainsi que par les cavités buccale et nasale. La voix n'est pas un attribut permanent. Elle change bien entendu avec l'âge, mais peut aussi être temporairement affecté par la santé ou l'état émotionnel du locuteur [15].

☑ **Signature :** Chaque personne a sa propre signature qui peut donc être utilisée pour l'identifier [15]. Les systèmes de reconnaissance analysent la signature (statique ou dynamique) afin d'en déduire les caractéristiques spécifiques qui permettent d'identifier le

signataire, voir Fig.I.11. Pour une signature dynamique, ces caractéristiques peuvent être la vitesse, la pression du crayon, le mouvement, les points et les intervalles de temps lorsque le crayon est soulevé [21].

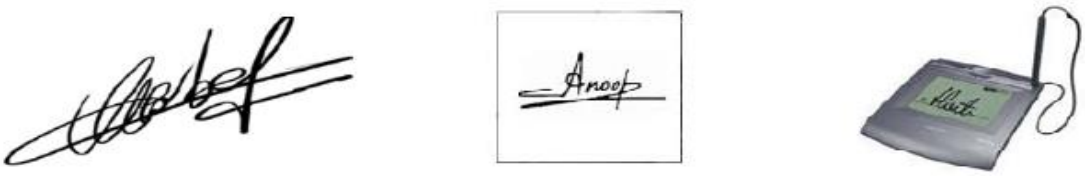


Fig I.11 : Signature manuscrite.

☑ **Démarche :** Il s'agit de reconnaître un individu à sa façon de marcher et de bouger (vitesse, accélération, mouvements du corps…), voir Fig. I.12, en analysant des séquences d'images.

Fig I.12 : La démarche.

La démarche, en anglait *gait*, serait en effet étroitement associée à la musculature naturelle et donc très personnelle. Mais les vêtements amples, par exemple, peuvent compromettre une identification correcte [21].

I.3 Avantage de la biométrie

☑ **Confort :** Au lieu d'avoir à saisir à nouveau le mot de passe et à prendre les cartes à chaque fois, il vous suffit de placer une empreinte digitale ou d'autres données biométriques sur le capteur pour que le système de contrôle d'accès vous reconnaisse et vous permette ainsi d'accéder à l'application.

☑ **Sécurité renforcée :** la biométrie permet de diminuer de la fraude et intransmissible à une autre personne.

☑ **Responsabilité :** La biométrie est le chaînon manquant dans la triade des problèmes de sécurité :

- Diminution de la fraude.
- Rehaussement de l'intégrité des informations et la sécurité.

- Réduction des attaques à l'égard des programmes gouvernementaux.
- Croissance de la confiance envers les systèmes de sécurité.
- Diminution des frais administratifs.
- Accélération des services.

☑ **Reconnaissance négatives :** Une identité vérifiée (Le destinataire est bien la personne autorisée à visualiser ou à utiliser les données). Grande sécurité, intransmissible à une autre personne.

I.4 Caractéristique biométrique

Un certain nombre de caractéristiques biométriques sont utilisées dans diverses applications. Chaque trait biométrique a ses avantages et ses inconvénients, et par conséquent, le choix d'un trait biométrique pour une application particulière dépend de divers problèmes en plus de ses performances de correspondance. Les facteurs qui déterminent l'aptitude d'un trait physique ou comportemental à être utilisé dans l'application de la biométrie sont les suivants [22] :

✍ **Universalité :** chaque individu accédant à l'application doit posséder le trait ;
✍ **Unicité :** le caractère donné doit être suffisamment différent d'un individu à un autre ;
✍ **Permanence :** le trait biométrique doit être inchangé sur une période de temps ;
✍ **Mesurabilité :** Il devrait être possible d'acquérir et de numériser le trait biométrique à l'aide de dispositifs appropriés qui ne causent pas de désagréments indus à l'individu ;
✍ **Acceptabilité :** les individus qui utiliseront l'application devraient être disposés à présenter leur trait biométrique au système.

I.5 Domaine d'application de la biométrie

L'authentification biométrique est utilisée dans tous les domaines, en particulier dans les domaines nécessitant un accès contrôlé tels que les applications bancaires, les lieux hautement sécurisés tels que le siège du gouvernement, l'armée, les services de sécurité, etc. Les applications de la biométrie peuvent être réparties en trois grandes catégories :

☑ **Applications commerciales :** Telles que l'ouverture de réseau informatique, la sécurité de données électroniques, l'e-commerce, l'accès Internet, la carte de crédit, le contrôle d'accès physique, le téléphone cellulaire, la gestion des dossiers médicaux, l'étude à distance, etc.

☑ **Applications gouvernementales :** La biométrie peut empêcher l'utilisation frauduleuse de documents, par exemple la carte d'identité nationale, le permis de conduire, la sécurité sociale, le contrôle aux frontières, le contrôle des passeports, etc.

☑ **Applications légales :** C'est certainement le premier domaine où l'identification biométrique a été appliquée, comme l'identification du corps, la recherche criminelle, l'identification des terroristes, etc.

I.6 Système biométrique

I.6.1 Définition

Un système biométrique est essentiellement un système de reconnaissance de formes qui reconnaît une personne sur la base d'un vecteur de caractéristiques dérivé d'un trait physiologique ou comportementale spécifique que la personne possède. Selon le contexte de l'application, un système biométrique fonctionne généralement selon l'un des deux modes suivants : vérification ou identification [23].

I.6.2 Modes de fonctionnement

Les systèmes biométriques présentent trois modes de fonctionnement principaux qui sont: le mode d'enregistrement (ou enrôlement), le mode de vérification (ou authentification) et le mode d'identification (ou reconnaissance) subdivisée en deux phases [24]:

☑ **Phase d'enregistrement (Enrôlement) :** C'est la première phase de tout système biométrique (voir Fig. I.13). Il s'agit de l'étape au cours de laquelle un utilisateur est enregistré pour la première fois dans le système. Il s'agit d'une phase commune de vérification et d'identification. Lors de l'enrôlement, la caractéristique biométrique est mesurée avec le capteur biométrique pour ensuite extraire la représentation numérique. Cette représentation est ensuite réduite, à l'aide d'un algorithme de sélection bien défini, afin de réduire la quantité de données à stocker et ainsi faciliter la vérification et l'identification. En fonction de l'application et du niveau de sécurité souhaité, le modèle biométrique (Template) sélectionné est stocké soit dans une base de données centrale, soit sur un objet personnel propre à chaque personne [25].

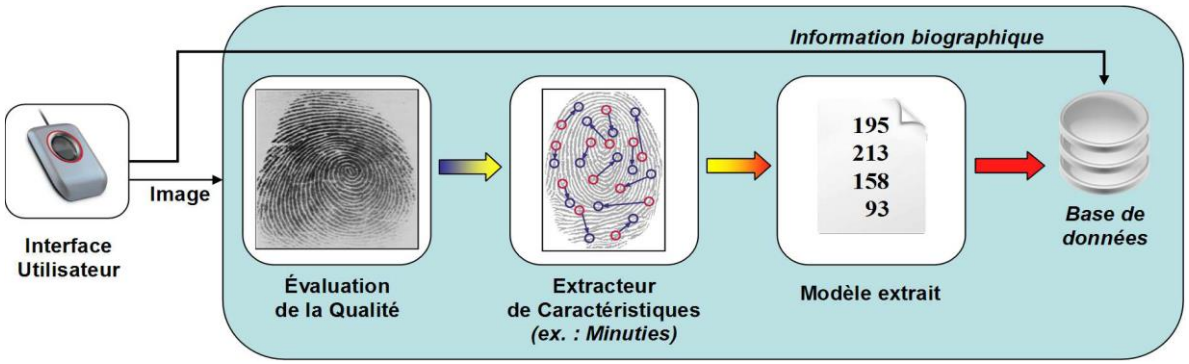


Fig. I.13: Enrôlement d'une personne dans un système biométrique.

☑ **Phase de reconnaissance :** Cette phase est divisée en deux modes comme suite :

– **Identification :** dans ce mode (voir Fig. I.14), le système biométrique détermine l'identité d'un individu à partir d'une base de données d'identités [I_1 ,I_2,…I_N] qui sont enrôlées dans le système. De plus, le système effectue le test ‘’un contre tous‘’ pour attribuer à l'individu inconnue. Pour déterminer l'identité I_k (k∈ [1,2,… N]) , le processus d'identification peut être formalisé comme suit [26] :

$$f(C_U) = \begin{cases} I_K \; Si \; \max \; s(C_U, M_k) \geq \tau \; avec \; 1 \leq k \leq N \\ I_0 \quad\quad sinon \end{cases} \qquad (1)$$

I_0 : L'identité inconnue.

M_k : Le modèle biométrique correspondant à l'identité I_k .

S : fonction de similarité.

τ: Le seuil de décision.

C_U : les caractéristiques biométriques extraites de l'utilisateur u .

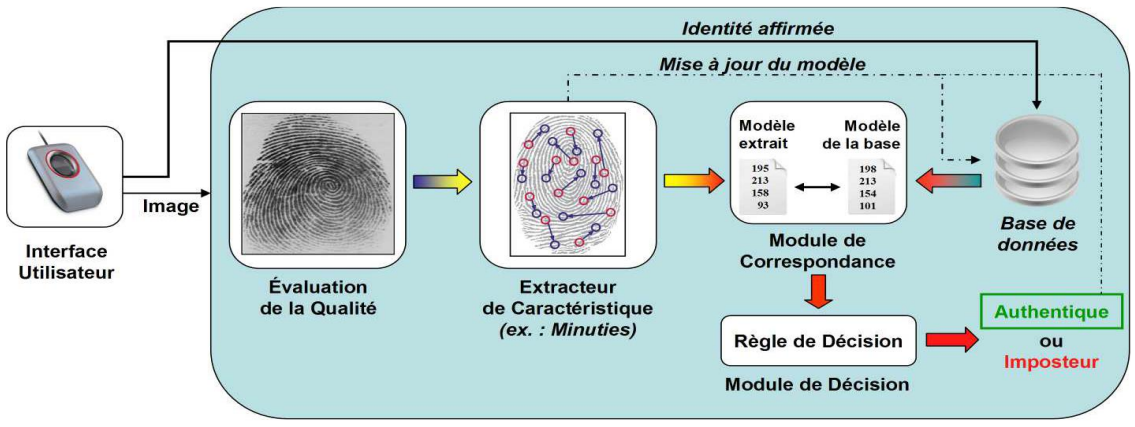


Fig I.14 : Identification d'un individu dans un système biométrique.

– **Vérification :** C'est le module de test qui permet la comparaison entre les données biométriques acquises et les données biométriques correspondantes stockées sur la base de données, on parle d'un test 1 contre 1, voir Fig. I.15. Dans ce cas, le système renvoie uniquement une décision binaire (oui ou non) qui peut être pondéré. Le processus de vérification peut être formalisé comme suit :

Soit le vecteur d'entrée C_U définissant les caractéristiques biométriques de l'utilisateur u extraites par le système, et M_u son modèle biométrique stocké dans la base de données, le système retourne une valeur booléenne suite au calcul de la fonction f définie par [[26]**, [**27]] :

$$f(C_u, M_u) = \begin{cases} 1 \; si \; S(C_U, M_K) \geq \tau \\ 0 \; sinon \end{cases} \qquad (2)$$

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

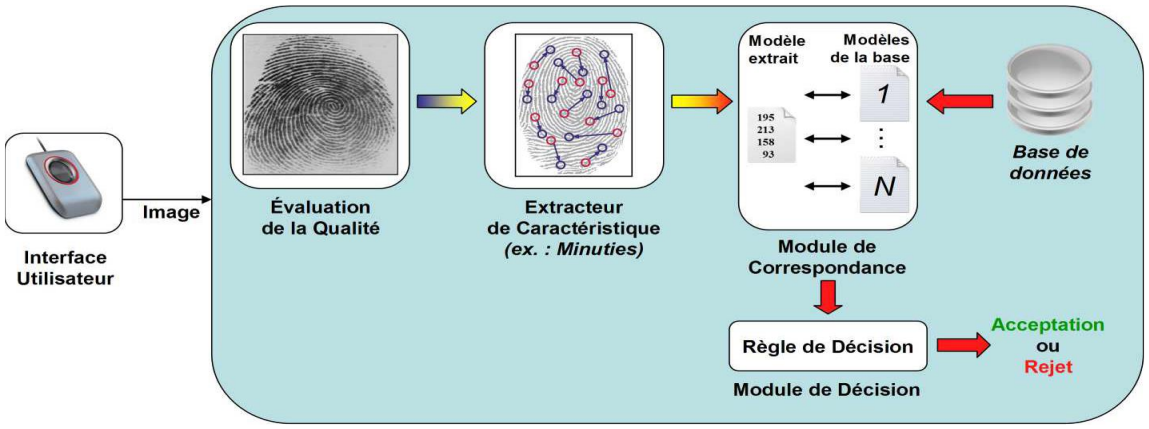


Fig I.15 : Vérification d'un individu dans un système biométrique.

I.6.3 Architecture d'un système biométrique

Le système biométrique contient les cinq modules suivants [28] :

☑ **Module de capture :** Ce module permet l'acquisition de données biométriques pour extraire des informations numériques.

☑ **Module d'extraction de caractéristiques :** Dans lequel les données acquises sont traitées pour extraire un ensemble de caractéristiques saillantes ou discriminatoires.

☑ **Module de stockage :** Qui contient les modèles biométriques des utilisateurs enrôlés du système.

☑ **Module de correspondance :** Dans lequel les caractéristiques extraites au cours reconnaissance sont comparés aux modèles stockés pour générer des scores correspondants.

☑ **Module de décision :** Dans lequel l'identité d'un utilisateur est confirmée (vérification) ou établie (identification) en fonction du score de correspondance.

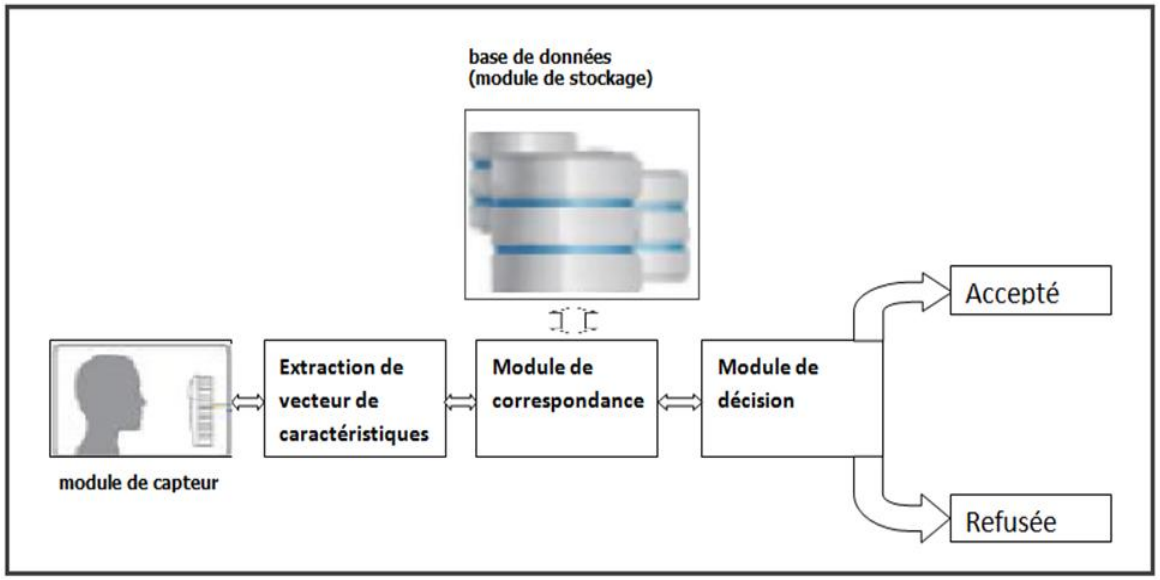


Fig I.16 : Architecture d'un système biométrique.

I.7 critères de performances

Les performances d'un système biométrique sont données en mesurant trois taux d'erreur:

I.7.1 Taux de faux rejet (“**F**alse **R**eject **R**ate” ou **FRR**)

Ce taux représente le pourcentage de personnes censées être reconnues mais rejetées par le système,

$$\mathbf{FRR}=\frac{nombre\ des\ clients\ rejeté(FR)}{nombre\ total\ d'accésde\ clients} \qquad (3)$$

I.7.2 Taux de fausse acceptation (“**F**alse **A**ccept **R**ate” ou **FAR**).

Ce taux représente le pourcentage de personnes censées ne pas être reconnues mais acceptées par le système,

$$\mathbf{FAR}=\frac{nombre\ des\ imposteurs\ accepté(FA)}{nombre\ total\ d'accésimposteurs} \qquad (4)$$

I.7.3 Taux d'égale erreur ("**E**qual **E**rror **R**ate" ou **EER**).

Ce taux est calculé à partir des deux premiers critères et constitue un point commun de mesure de la performance. Ce point correspond à l'endroit où FRR = FAR, c'est-à-dire le meilleur compromis entre faux rejets et fausses acceptations.

$$\mathbf{EER}=\text{FAR}=\text{FRR} \qquad (5)$$

I.7.4 Courbes de performances

Distribution des scores : Pour évaluer la précision d'un système biométrique, nous devons calculer des scores à partir d'échantillons biométriques appartenant à la même personne et des scores à partir d'échantillons biométriques de différentes personnes. La distribution des scores d'échantillons biométriques appartenant à la même personne s'appelle la distribution des personnes clientes. La distribution des scores d'échantillons biométriques de différentes personnes est appelée distribution d'imposteurs. Fig. I.17 illustre les distributions des scores des clients et des scores d'imposteurs.

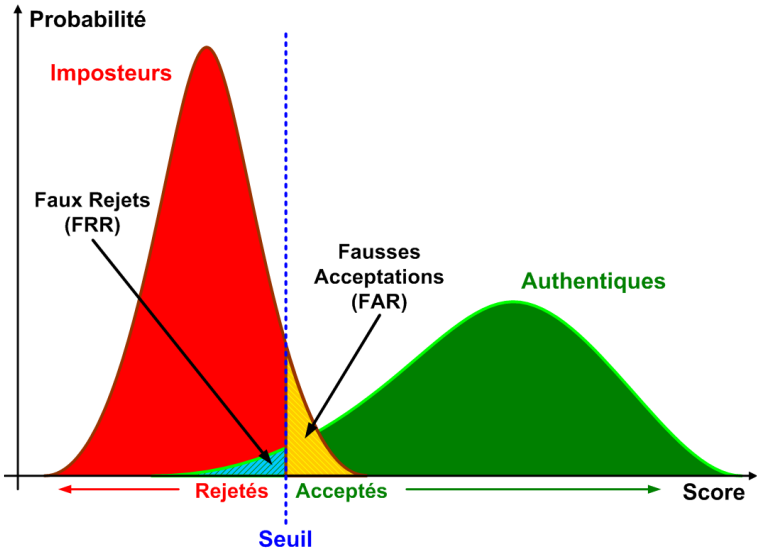


Fig I.17 : illustration du FRR et du FAR.

Les performances d'un système biométrique (vérification ou identification ensemble ouvert) sont généralement évaluées suivant les FAR et FRR. Comme montre la Fig. I.17, ces taux ont une relation de corrélation directe entre eux, c'est-à-dire que si l'un augmente, l'autre diminue. Dans le cas d'un système utilisé dans le processus d'identification, les applications peuvent être déployées dans un monde ensemble fermé (*closed-set identification*) ou dans un monde ensemble ouvert (*open-set identification*) :

– **Mode ensemble ouvert :** Lors de la mesure des performances d'un système biométrique, nous utilisons ce que l'on appelle une courbe ROC. La courbe ROC (Fig. I.18) trace le taux de faux rejet en fonction du taux de fausse acceptation. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé.

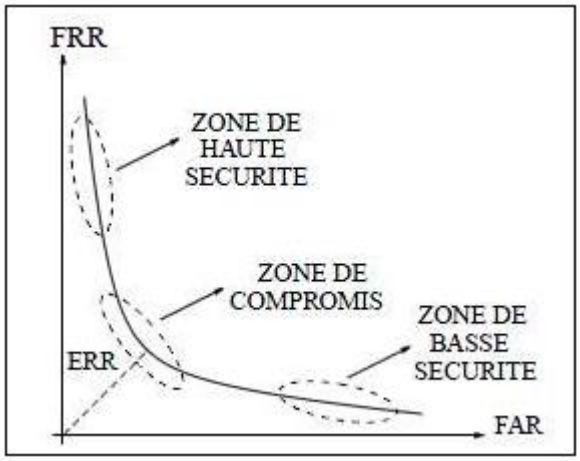


Fig I.18 : courbe ROC.

– **Mode ensemble fermé :** La mesure de ROR est la plus utilisée mais il n'est pas toujours suffisant. En effet, en cas d'erreur, il peut être utile de savoir si le bon choix se trouve les ρ_n premières réponses. Nous traçons alors la courbe des scores cumulées (Cumulative Match Characteristics-CMC) qui représente la probabilité que le bon choix se trouve parmi les premières réponses, comme l'illustre la Fig. I.19. La courbe (CMC) donne le pourcentage de personnes reconnues en fonction d'une variable que l'on appelle le rang.

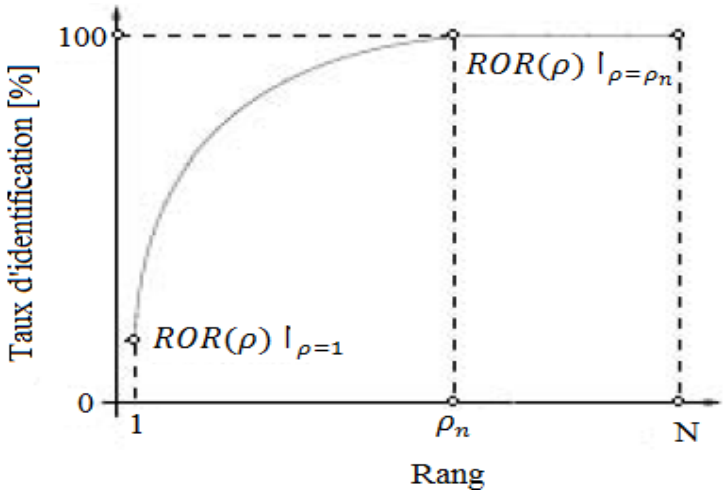


Fig I.19 Courbe de scores cumulés (CMC).

I.8 Exigence du système biométrique :

☑ **Vie privé :** la confidentialité des données clients doit être assurée et sécurisée ;

☑ **Usabilité** : c'est-à-dire facile à utiliser et les utilisateurs doivent être satisfaits lorsqu'ils utilisent ces systèmes ;

☑ **Sécurité :** Le système biométrique doit être robuste contre diverses menaces et l'identité de l'individu doit être protégée.

I.9 Modèle biométrique : vulnérabilités et menaces

L'exposition du modèle biométrique face aux menaces qui concernent aussi bien les risques de violation de la vie privée que les risques d'usurpation d'identité. La sécurité du modèle biométrique est l'une des questions les plus cruciales dans la conception d'un système biométrique sécurisé. Les niveaux de vulnérabilités deviennent plus élevés lorsqu'il s'agit de systèmes biométriques à grande échelle, de bases de données centralisées et de biométrie à distance.

I.9.1 Les risques de violation de la vie privée

La vie privée est relative à la protection des individus à l'égard du traitement des données personnelles et sensibles. Le modèle biométrique est une donnée personnelle car elle identifie son propriétaire.

– Les données biométriques (en particulier basées sur des images) peuvent exposer des informations sensibles telles que des informations sur la santé, l'origine raciale ou ethnique des individus. Ces informations peuvent ensuite servir de base à une discrimination injustifiée ;

– Les données non secrètes sont une modalité à trace. Par conséquent, la collecte et l'utilisation de données biométriques sans le consentement de la personne concernée rend ces informations très sensibles.

– Un autre risque de violation de la vie privée est l'utilisation abusive. Par exemple, une application initialement destinée à un usage spécifique peut être progressivement étendue à un autre usage sans le consentement de la personne [29].

La possibilité de compromission de la base de données, du fait de son unicité, le modèle biométrique ne peut être révoqué et réémis ce qui rend cet identifiant très sensible.

I.9.2 Risques d'usurpation d'identité

Avec le modèle biométrique, un agent attaquant (interne ou externe) pourrait entrer dans le système biométrique en se faisant passer pour un utilisateur légitime. L'attaque de mascarade [29] qui consiste à capturer et envoyer une image contrefaite directement au module d'extraction est un type de menace, généralement plus pernicieuse parce qu'il ne nécessite pas la présence physique de l'attaquant. Ceci est un exemple de la façon dont l'identité peut être volée, qui nécessite une attention urgente et rigoureuse pour fournir des systèmes plus sûrs.

I.10 Conclusion

Aujourd'hui, la biométrie est utilisée dans une variété d'applications pour assurer la sécurité et faciliter l'identification des personnes. Dans ce chapitre, nous nous sommes concentrés sur la biométrie, où nous avons présenté les différentes modalités biométriques et les systèmes biométriques. De plus, nous avons discuté des exigences, de la structure, du fonctionnement et des problèmes qui menacent un système biométrique. Enfin, nous présentons les différents critères couramment utilisés pour évaluer les performances de ces systèmes.

Chapitre 2

Protection de modèle biométrique

Résumé

Ces dernières années, l'utilisation de diverses technologies biométriques pour l'identification automatique des personnes s'est considérablement développée. Avec l'adaptation rapide de ces systèmes, malgré leurs avantages, ils sont menacés par des attaques qui conduisent au vol de la vie privée de l'individu. Pour éviter le vol de modèles biométriques, il est souhaitable de fournir des solutions pour améliorer la sécurité, ce qui conduira nécessairement à la confiance des utilisateurs. Il existe des solutions telles que les cryptosystèmes biométriques ou la biométrie révocable. Dans ce chapitre, nous présenterons un aperçu des méthodes de protection des modèles biométriques, où nous nous concentrerons spécifiquement sur la biométrie révocable.

Introduction

Les modèles biométriques, qui sont utilisés pour identifier les individus, sont des données personnelles et sensibles. Pour assurer leur sécurité, de nombreux systèmes ont été mis en place pour les protéger et, à leur tour, protéger la vie privée et l'identité individuelle. Dans ce chapitre, nous présenterons les solutions disponibles pour protéger les modèles biométriques (biométrie non traçable), en nous concentrant sur celles basées sur la révocabilité. À la fin de ce chapitre, nous présenterons une présentation complète et détaillée de notre méthode proposée qui est basée sur des cartes chaotiques et peut produire des modèles biométriques profonds pour augmenter la précision du système et révocables pour augmenter la sécurité de la vie privée individuelle.

II.1 Renforcement de la sécurité des systèmes biométriques

Comme déjà mentionné, malgré les avantages des systèmes biométriques, ils ne sont pas toujours fiables et le modèle biométrique stocké peut être volé par un fraudeur afin de le réutiliser illégalement. En effet, dans la littérature sur la sécurité des systèmes biométriques, il existe trois technologies de protection des modèles biométriques.

☑ Crypto-systèmes biométriques.
☑ Biométrie révocable.
☑ Biométrie anonyme.

En général, la biométrie non traçable doit présenter cinq caractéristiques pour une protection adéquate des données :

- Les données biométriques brutes et le modèle biométrique ne sont pas conservés ;
- Il est impossible de recréer une image ou un modèle des données biométriques à partir des informations stockées, les rendant ainsi non traçables ;
- Un grand nombre de modèles biométriques différents non traçables peuvent être créés à partir des mêmes données biométriques pour différentes applications ;
- Les modèles biométriques non traçables ainsi créés pour différentes applications ne peuvent pas être liés ensemble ;
- Un modèle biométrique non traçable peut être annulé ou renouvelé.

II.1.1 Crypto-systèmes biométriques

Un crypto-système biométrique combine la biométrie et la cryptographie qui libère des clés cryptographiques associées à des modèles biométriques (voir Fig. II.1). Ainsi, les crypto-systèmes biométriques offrent des solutions pour sécuriser la gestion des clés cryptographiques ainsi que la protection des modèles biométriques.

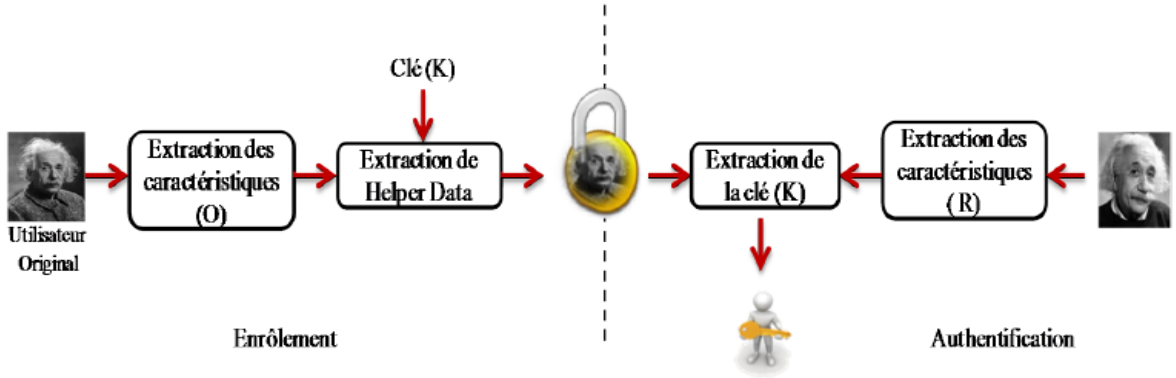


Fig II.1 : Principe de fonctionnement de crypto-systèmes

En fonction de la façon dont les données d'assistance (*helper data*) sont dérivées, les approches de cryptosystème biométrique peuvent être divisées en deux catégories : les types de liaison de clé (*key binding*) et de génération de clé (*key generation*) [30]. Dans la première catégorie, il est possible de lier une clé sécrète aux données biométriques pour obtenir un croquis sécurisé (*secure sketch*) à partir duquel aucune information concernant les données biométriques ou la clé ne peut être récupérée. Dans cette catégorie, deux exemples bien connus sont utilisés, à savoir l'engagement flou (*fuzzy commitment*) [31] et la voûte floue (*fuzzy vault*) [32]. Basé sur une clé cryptographique, le premier schéma sécurise les modèles biométriques en tant que vecteurs binaires, tandis que le second les sécurise en tant

qu'ensemble non ordonné de points. Dans la deuxième catégorie (*key generation*), les données d'assistance sont dérivées uniquement du modèle biométrique de sorte que la clé cryptographique est directement générée à partir des données d'assistance et d'un modèle biométrique donné.

II.1.2 Biométrie anonyme

La biométrie anonyme [33] est un dispositif dans lequel les données biométriques ne sont reliées à aucune donnée personnelle permettant d'identifier la personne, et qui ne permet aucune interconnexion avec un autre système où elle pourrait être identifiée. Elle peut être utilisée pour les dispositifs d'authentification seulement, en utilisant par exemple un tiers de confiance qui authentifierait la donnée biométrique à la demande de la personne ou du fournisseur de service. La seule donnée communiquée serait un numéro de transaction. Cela demanderait de la part du tiers de confiance qu'il mette en place des procédures sécurisées et adéquates d'enrôlement, de conservation, de comparaison.

II.1.3 Biométrie révocable

La biométrie révocable [34] fait référence à la distorsion intentionnelle et systématiquement répétable des caractéristiques biométriques pour protéger les données sensibles spécifiques à l'utilisateur. Si une fonction annulable est comprise, les caractéristiques de distorsion sont modifiées et la même biométrie est mappée vers un nouveau modèle, qui est utilisé plus tard.

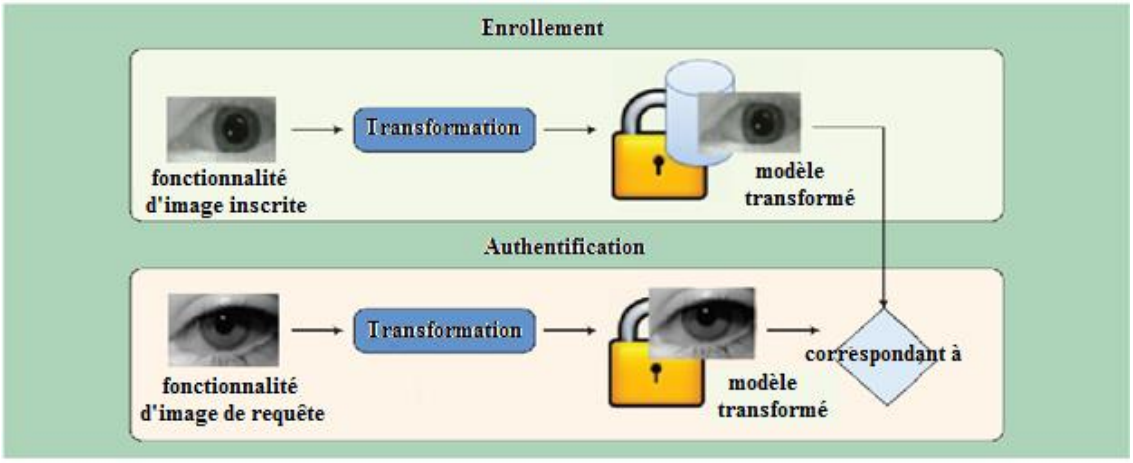


Fig II.2 : Transformation révocable

Pour sécuriser un système biométrique avec le principe de révocabilité, les objectifs suivants doivent être pris en compte :

- **Diversité :** aucune même caractéristique révocable ne peut être utilisée dans diverses applications.
- **Réutilisation/révocabilité :** révocation et réémission simples en cas de compris.
- **Performances :** la formulation ne doit pas affecter négativement les performances de reconnaissance.

Enfin, il est à noter que ces méthodes peuvent être combinées pour obtenir une méthode hybride capable d'améliorer la sécurité du système. Par exemple, les méthodes de protection basées sur la révocabilité peuvent être combinées avec celles des cryptosystèmes biométriques pour construire un système robuste capable d'exploiter les avantages des deux méthodes.

Par exemple, *Feng et al.* [35] ont proposé une approche hybride basée sur la reconnaissance faciale en utilisant d'abord une projection aléatoire du modèle biométrique, puis le résultat a été chiffré à l'aide d'un cryptosystème biométrique basé sur l'engagement flou (*Fuzzy Commitment*). En outre, *Song et al.* [36]] ont proposé une méthode hybride basée sur la génération de la clé secrète lors de l'inscription à partir de données biométriques en appliquant le hachage discret et le code correcteur d'erreur de *Reed Solomon* [37].

II.2 Méthodes de la biométrie révocable :

II.2.1 Salage biométrique

Le salage biométrique est similaire au salage de mot de passe en cryptographie, qui consiste en des bits aléatoires *r* utilisés comme facteur d'entrée à concaténer avec une clé secrète, *k*. La sortie est souvent stockée sous forme de hachage *H(r+k)* dans la base de données.

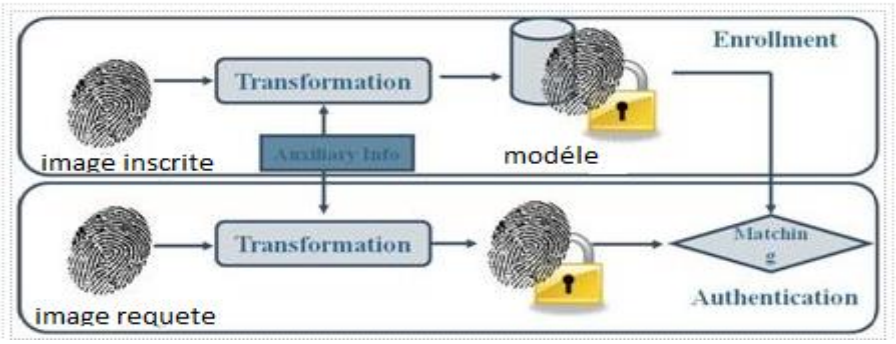


Fig.II.3 : Schéma fonctionnel du salage biométrique.

Le salage biométrique suit un principe selon lequel un facteur d'entrée indépendant spécifique à l'utilisateur (des données auxiliaires telles qu'un mot de passe ou des nombres aléatoires spécifiques à l'utilisateur) est mélangé avec des données biométriques pour dériver une version déformée du modèle biométrique.

Puisque les données auxiliaires sont dérivées de l'extérieur et interagissent directement avec les données biométriques, elles peuvent être modifiées et révoquées facilement mais doivent être gardées secrètes pour une protection maximale. Cependant, étant donné que les clés ou mots de passe externes confidentiels peuvent facilement être perdus, volés ou compromis, l'exactitude et les vulnérabilités des systèmes existants doivent être justifiées.

Dans ce contexte, il y a eu un certain nombre d'efforts de recherche visant à résoudre les problèmes liés à la sécurité des systèmes biométriques. *Savvides et al*. (2004) [38**]** ont proposé un exemple de méthode de salage biométrique qui crypte les images d'apprentissage en synthétisant un filtre de convolution pour la reconnaissance faciale. Ils ont montré que différents modèles peuvent être obtenus à partir des mêmes données biométriques en faisant varier les noyaux de convolution aléatoire permettant ainsi des modèles révocables. Notez que les noyaux aléatoires ont été générés à partir de différentes matrices aléatoires créées à l'aide d'un code PIN spécifique à l'utilisateur. Leurs résultats ont démontré que la convolution des images d'apprentissage avec n'importe quel noyau de convolution aléatoire avant la construction du filtre biométrique ne modifie pas les rapports de crête de sortie de corrélation résultants sur les lobes latéraux, préservant ainsi les performances existantes. Cependant, la sécurité pourrait être compromise par une dé-convolution déterministe avec un noyau aléatoire connu.

Hirata et al. (2009) [39] ont proposé une nouvelle méthode utilisant la mise en correspondance basée sur la corrélation. Dans leur travail, l'image biométrique a été transformée par transformation théorique de nombre, puis le résultat obtenu à été masqué avec un filtre aléatoire. Il est possible de calculer la corrélation entre l'image enregistrée et l'image de d'entrée dans le champ masqué (c.-à-d. le champ crypté) sans connaître les images d'origine. Les auteurs ont appliqué la méthode proposée à la vérification des réseaux veineux de doigts et ont obtenu des performances de vérification très élevées.

Jeong et al. (2006) [40] ont proposés un schéma de salage biométrique pour un modèle de visage basé sur l'apparence. Deux vecteurs de caractéristiques ont été extraits avec PCA (analyse en composantes principales) et ICA (analyse en composantes indépendantes) à partir

d'une image faciale, et ces vecteurs ont ensuite été normalisés. Les vecteurs résultants ont ensuite été permutés à l'aide d'une matrice de permutation dérivée de jetons et fusionnés au niveau des caractéristiques via la règle SUM. Si ces vecteurs ont été compromis, un nouveau vecteur de caractéristiques peut être généré en changeant la matrice de permutation. Pour les minuties d'empreintes digitales, *Lee et al.* (2007) [41] ont introduit les valeurs invariantes de translation et de rotation, qui ont été extraites à l'aide d'informations d'orientation autour de chaque minutie. Ces valeurs ont ensuite été utilisées comme entrées pour deux fonctions de transformation spécifiques à l'utilisateur chargées de générer les paramètres de translation et de rotation. Les modèles biométriques ont ensuite été construits en changeant chaque minute selon lesdits paramètres. Lorsque le modèle révocable a été compromis, un nouveau modèle peut être régénéré en remplaçant les fonctions de transformation.

Farooq et al. (2007) [42] ont présenté une méthode en convertissant les minuties d'empreintes digitales en une chaîne binaire annulable. L'idée est basée sur le fait que les empreintes digitales peuvent être représentées par un ensemble de triangles dérivés d'ensembles de trois minuties qui peuvent être utilisés directement dans une correspondance basée sur un modèle. La méthode proposée s'est avérée irréversible en termes de calcul et satisfait aux critères de réutilisabilité et de diversité. Notez que la réutilisabilité est obtenue en attribuant une clé unique à chaque utilisateur de la base de données pour randomiser le modèle utilisateur, et en cas de compromis, le modèle biométrique peut être révoqué en attribuant simplement une clé différente.

II.2.2 Transformations non inversibles

La transformation non inversible est une fonction conçue pour modifier intentionnellement une image biométrique brute en une nouvelle forme dans le contexte de la fonction ou de l'espace de signal. Cette fonction sert d'agent dans le cadre de la sécurité des modèles permettant la non-inversibilité, la réutilisabilité et la diversité des modèles. Comme cette fonction n'interagit pas directement avec la biométrie brute, le principal avantage de cette approche est que cette fonction n'a pas besoin d'être tenue secrète [[43, 44]].

La réalisation d'une transformation non inversible a été rapportée par *Ratha et al*. (2007) dans [45] dans lequel les données d'empreintes digitales sont transformées par une séquence de trois fonctions de transformation non inversibles, voir Fig. II.4. Comme le montre cette figure, les trois fonctions de transformation sont basées sur la transformation cartésienne, polaire et par pliage de surface des positions des minuties.

☑ **Transformation polaire :** Dans la méthode de transformation polaire, les positions des minuties sont mesurées en coordonnées polaires par rapport à la position centrale. Les angles sont mesurés par rapport à l'orientation du noyau. En conséquence, l'espace de coordonnées est divisé en régions polaires. La transformation non inversible consiste à changer les positions des coins polaires. Les angles des minuties changent également avec les différences de positions des coins avant et après transformation [[45, 46].

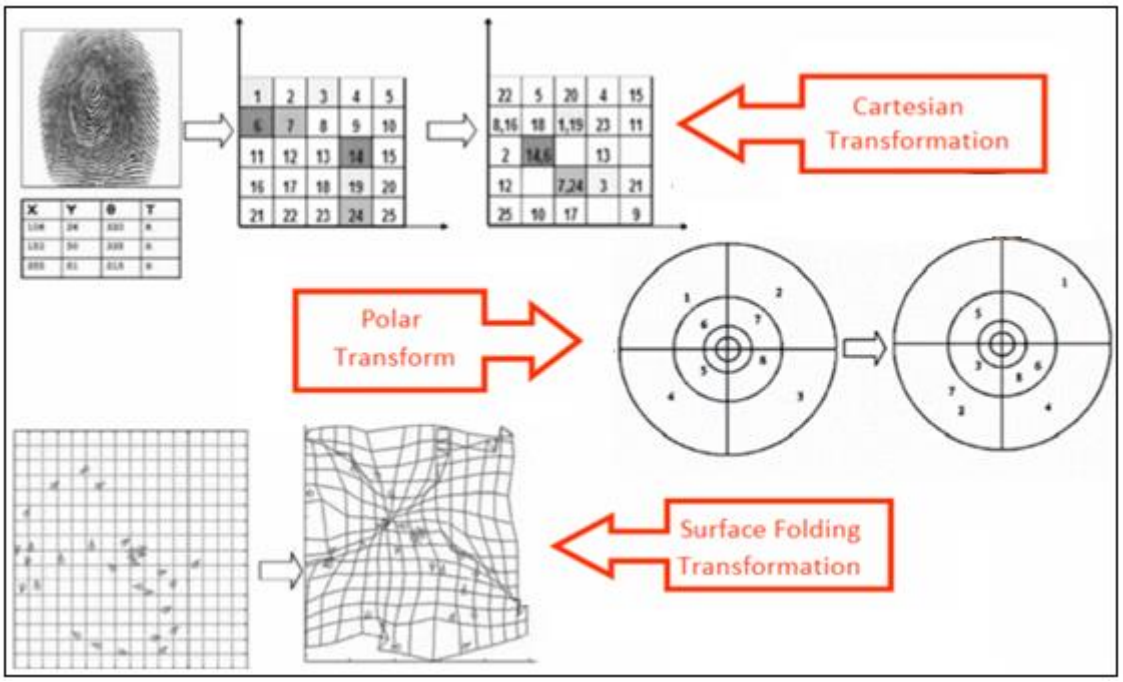


Fig II.4 : Modèles annulables de différentes transformations

☑ **Transformation cartésienne** : Dans cette transformation, les positions des minuties sont mesurées en coordonnées rectangulaires par rapport à la position du point singulier en alignant l'axe x avec son orientation, le système de coordonnées est divisé en cellules de taille fixe. Notez que cette transformation n'est pas une simple permutation car chaque cellule, qui contient éventuellement quelques minuties, est déplacée vers une nouvelle position par une transformation non inversible [45], [46].

☑ **Transformation de pliage de surface :** Dans la méthode de transformation par pliage de surface, un mélange gaussien 2D est utilisé pour traduire les points de minutie. Etant donné que les transformations utilisées dans le mélange sont localement lisses, cela n'aura qu'un effet minimal sur les taux d'erreur et ne réduira pas dans une large mesure la discrimination des minuties par rapport aux deux transformées précédentes. Cependant, comme un petit changement dans la position des minuties de l'empreinte digitale originale peut conduire à une grande transformation après transformation, en particulier si le point franchit une frontière

nette, un pré-alignement approprié en référence à la position du point central est nécessaire pour s'assurer que la caractéristique biométrique est transformée de manière cohérente à travers plusieurs instances de minuties.

II.2.3 Méthodes de bio-hachage

Les schémas de cette méthode sont principalement appliqués à l'empreinte digitale, elle a d'abord été proposée en 2003 pour la reconnaissance faciale puis en 2004 pour les empreintes digitales. Son principe est généralement généré un bio code qui est utilisé lors de l'enrôlement et de la vérification. L'utilisateur présente son empreinte et un nombre aléatoire stockés sur une clé USB, une carte à puce, un jeton, etc. pour extraire les paramètres de l'empreinte sous la forme d'un code de doigt.

Pour générer le bio-code, la fonction de transformation prend le code du doigt et le nombre aléatoire comme entrée, voir Fig. II.5. Ce code est ensuite stocké dans la base de données. Lors de la vérification, le bio-code est recalculé à chaque vérification, il nécessite le stockage sécurisé de nombre aléatoire. Le résultat de la vérification se fait en calculant la distance de *Hamming* entre le Bio-code de référence et le bio-code calculé [47].

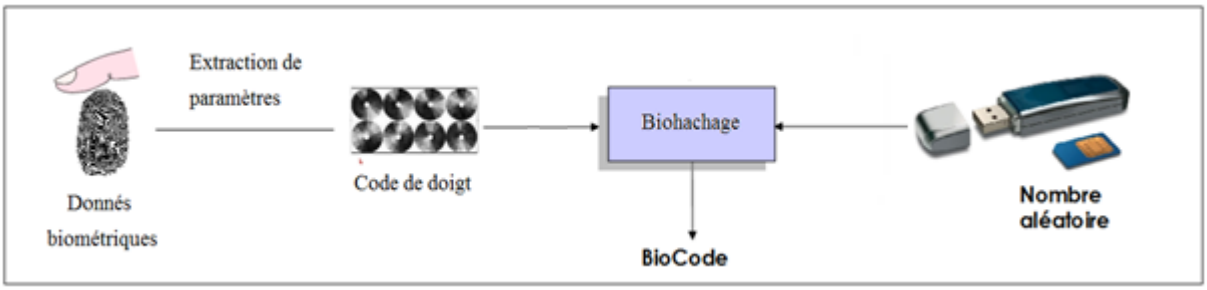


Fig II.5 : Schéma général de protection d'une donnée biométrique

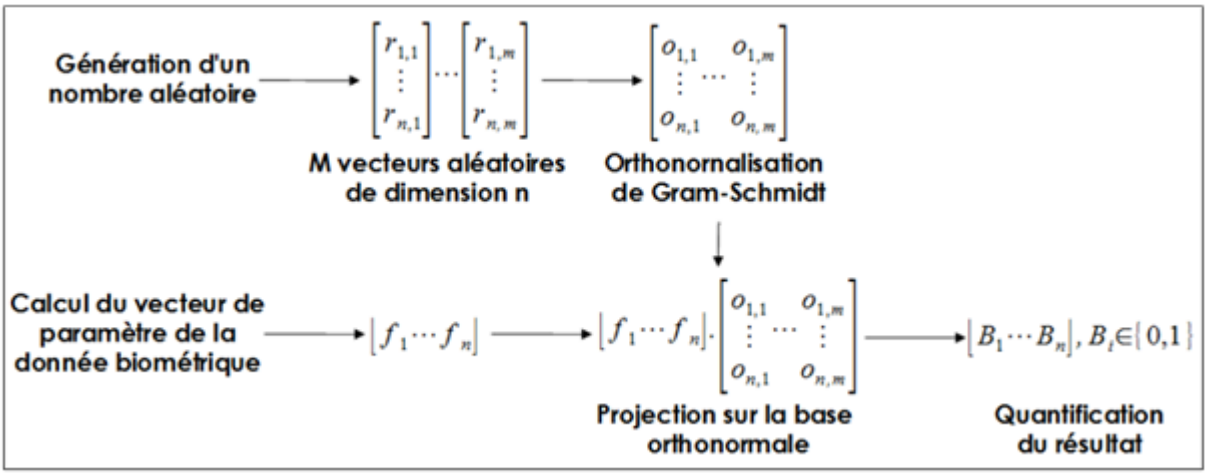


Fig II.6 : Description du procédé de génération d'un Biocode

II.3 Méthode proposée (C-PCANet)

Pendant longtemps, les chercheurs ont tenté de concevoir de nombreux systèmes biométriques révocables et ont réussi dans une certaine mesure, mais un examen rapide des publications récentes de la dernière décennie ne montre pas une grande partie du travail basé sur l'apprentissage en profond *(deep learning*). Dans la plupart des systèmes biométriques révocables basés sur une transformation irréversible, qu'elle soit basée sur les méthodes d'extraction de caractéristiques classiques (*hand-crafted*) ou des méthodes d'apprentissage profond, le vecteur de caractéristiques final est directement converti à l'aide d'une fonction de transformation, ce qui peut donner de mauvais résultats en raison du processus de quantification. Nous avons donc pensé intégrer cette fonction dans la méthode d'extraction de caractéristiques, surtout si l'on prend en compte les avantages offerts par la technique PCANet. En tant que l'une des méthodes d'extraction de caractéristiques les meilleures et les plus simples, le PCANet a montré des résultats significatifs, en particulier dans les systèmes biométriques. Cette méthode est principalement conçue pour extraire les caractéristiques profondes des images. Il est rapide et léger par rapport aux autres méthodes d'apprentissage profond. De plus, étant donné que la plupart des systèmes biométriques, qui implémentent la biométrie révocable, souffrent de performances dégradées en raison du processus de bio-hash, la structure PCANet contient déjà ce processus et son effet sur les performances du système est donc moindre que s'il avait été ajouté récemment. Ainsi, la motivation derrière cette recherche est de développer et de tester une nouvelle méthode d'extraction de caractéristiques biométriques révocables (C-PCANet) et profondes pour le développement de systèmes biométriques sécurisés.

II.3.1 Fondements préliminaires

Cette étude vise à développer une méthode sûre et fiable pour protéger le modèle biométrique des risques de tentatives de fraude lors de la transmission ou du stockage afin qu'il puisse être annulé à tout moment. En général, tous les problèmes associés à la conception finale du système sont liés à la tâche d'extraction de caractéristiques. Dans cette section, la technique d'analyse en composantes principales (Principal Component Analysis-PCA) et les systèmes chaotiques utilisés dans la méthode d'extraction de caractéristiques proposée sont expliqués à l'aide de quelques principes mathématiques.

☑ **Analyse en composantes principales :** En raison de l'intérêt croissant pour l'utilisation d'images dans les applications de vision artificielle, un certain nombre de méthodes ont été

utilisées pour analyser le contenu de ces images afin de les comprendre et donc de les utiliser de manière optimale. La PCA [48] est l'une des méthodes d'analyse de données les plus efficaces pouvant être utilisées dans l'analyse d'images. Cette technique utilise une transformation orthogonale statistique pour convertir un ensemble de données corrélées en données linéairement non corrélées appelées composantes principales.

Soit les N vecteurs d'entrée $\{V_i \ \epsilon R^{n\times 1}\}$ sont concaténés pour former une base d'apprentissage comme suit :

$$V = \{v_1, v_2, \dots v_N\} \epsilon R^{n\times N} \quad (1)$$

Où *n* désigne la longueur du vecteur et *N* le nombre total de vecteurs d'apprentissage. Pour décomposer cette base (*V*) en composantes principales, il faut d'abord la normaliser en supprimant les moyens de chaque vecteur pour former la base d'apprentissage normalisée $\bar{V}$.

$$\bar{V} = \{\bar{v}_1, \bar{v}_2, \dots \bar{v}_N\} \epsilon R^{n\times N} \quad (2)$$

Où $\bar{v}_i$ est défini comme :

$$\bar{v}_i \ = v_i - \frac{1}{N}\sum_{j=1}^{N} v_i \quad (3)$$

Deuxièmement, les valeurs propres et les vecteurs propres de *V* sont obtenus à l'aide de l'algorithme PCA en minimisant l'erreur objective (ε) définie ci-dessous :

$$\varepsilon = \|\bar{V} - UU^T\bar{V}\|^2 \quad / U^T U = I_m \quad /U \in \mathbb{R}^{n\times m} \quad (4)$$

Où *m* ($m \leq N$) désigne le nombre de composants principaux qui ont les valeurs propres les plus élevées et $I_m$est une matrice d'identité de taille $m \times m$. L'équation (4) peut être résolue en utilisant la décomposition des valeurs propres définie comme suit :

$$C = U\Lambda \mathrm{U}^T \quad (5)$$

Où C est une matrice de covariance définie comme :

$$C = \frac{1}{N}\ \bar{V}.\bar{V}^T \epsilon \mathbb{R}^{n\times m} \quad (6)$$

Et Λ est une matrice diagonale composée des *m* premières valeurs propres les plus grandes de *C* :

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_m \end{pmatrix} \quad (7)$$

Les valeurs de $\lambda_{i|i=1..m}$ sont triées sous forme décroissante, c'est-à-dire $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. Enfin, la matrice obtenue $U \epsilon R^{n \times m}$ est principalement composé des *m* premiers vecteurs propres principaux (composantes principales):

$$U = \{u_1, u_{2,\dots,} u_m\} \epsilon \mathbb{R}^{n \times m} \quad (8)$$

En générale, le PCA fait pivoter le système d'axes initial car les vecteurs propres sont orthogonaux les uns par rapport aux autres, formant ainsi un nouveau système de coordonnées. Les composants des vecteurs propres sont le cosinus entre les nouveaux et les anciens axes. En effet, le premier vecteur propre est le vecteur qui comprend la majorité des informations de la base d'apprentissage. Le deuxième vecteur propre est celui qui explique le mieux tous les résidus obtenus, et ainsi de suite ...

☑ **Systèmes chaotiques :** Ces dernières années, le comportement dynamique des systèmes non linéaires a suscité un grand intérêt pratique dans de nombreuses applications en raison de leur simplicité, complexité et richesse [49]. Les systèmes chaotiques sont parmi les plus importants de ces systèmes, qui se caractérisent par leur extrême sensibilité aux conditions initiales, leur périodicité, leur comportement pseudo-aléatoire et leur grande complexité [50]. En effet, dans un système chaotique, la sensibilité aux conditions initiales est sans aucun doute la caractéristique essentielle d'un comportement chaotique dont l'évolution est imprévisible sur le long terme. Il est donc sensible à une très faible perturbation de la condition initiale (état initial). Même si les points de départ sont presque identiques, les trajectoires se séparent rapidement.

Un système chaotique en temps discret est défini par l'équation suivante :

$$x_{n+1} = \Gamma(x_n)\,, \quad n = 0,1,2 \dots \quad (9)$$

Où $x_n \epsilon R^n$ est appelés état, et Γ trace l'état suivant x_{n+1}. A partir d'un état initial x_0, l'application répétée de cette fonction (Γ) provoque une séquence de *N* points ($\{x_n\}_{n=0}^{N}$) appelée orbite du système à temps discret.

Sans aucun doute, ces systèmes ont été utilisés avec succès dans des applications de sécurité de l'information, pour la génération de clés secrètes dynamiques dans des algorithmes

de cryptage, de stéganographie et de tatouage numérique. Les cartes chaotiques sont l'un des systèmes les plus simples à utiliser pour générer une séquence chaotique. Dans la littérature, plusieurs cartes chaotiques à une dimension (1-D), deux dimensions (2-D) et trois dimensions (3-D) sont proposées. Dans cette sous-section, nous décrirons brièvement quelques cartes chaotiques, telles que la logistique et le tente.

- **Cartes logistiques :** La carte logistique discrète unidimensionnelle [51] est un système chaotique défini par l'équation de récurrence suivante :

$$\begin{aligned} x_{n+1} &= \Gamma_\iota(x_n, u) \\ &= \mu x_n(1 - x_n), \qquad \mu \epsilon[0,4], \qquad x_n \epsilon[0,1] \end{aligned} \tag{10}$$

x_n où x_{n+1} est l'état du système pour n = 0, 1, 2, · · · et μ est le paramètre de contrôle. Selon les valeurs de μ, Γ_l peut être une séquence convergente, une séquence oscillante ou une séquence chaotique :

- 0 ≤ μ ≤ 1 : les états de Γ_l seront proches de zéro, quelle que soit la population de départ.
- 1 <μ ≤ 3 : les états de Γ_l se stabiliseront à la valeur $\frac{\mu-1}{\mu}$ quelle que soit la population initiale.
- 3 <μ ≤ 3,45 : les états de Γ_l oscilleront entre deux valeurs de la population initiale.
- 3,45 <μ <3,57 : les états de Γ_l oscilleront entre 4 valeurs, puis 8, 16, 32, etc., qui sont également indépendantes de la population initiale.
- 3,57 ≤ μ ≤ 4 : les états de Γ_l seront chaotiques et de légères variations dans la population initiale conduiront à des résultats radicalement différents.
- 4 <μ : les états de Γ_l quittent l'intervalle [0,1] et divergent pour toutes les valeurs initiales.

Généralement, dans un système de sécurité d'information, l'état initial x_0 et le paramètre de contrôle μ du système chaotique peuvent être utilisés comme clés secrètes $K \equiv \{x_0, \mu\}$.

- **Cartes des tentes :** La carte de tente [52] est un système dynamique caractérisé par deux lignes simples, ce qui rend son analyse simple par rapport aux systèmes non linéaires. Son comportement peut être défini par l'équation récurrente suivante :

$$\begin{aligned} x_{n+1} &= \Gamma_t(x_n, u) \\ &= \mu \min(x_n, 1 - x_n), \quad \mu \epsilon[0,2], \quad x_n \epsilon[0,1] \end{aligned} \tag{11}$$

Où x_n est l'état du système pour n = 0, 1, 2, · · · et μ est le paramètre de contrôle. Heureusement, malgré la simplicité et la linéarité de cette équation, pour certains paramètres,

ce système peut fournir des comportements très complexes et des phénomènes chaotiques. Cependant, en fonction des valeurs de μ et de la valeur initiale x_0, ce système présente des comportements très différents :

- μ <1 : les états de Γ_t convergeront vers zéro, quelle que soit la valeur initiale (x_0);
- μ = 1 : les états de Γ_t convergeront à chaque valeur de $0 \leq x \leq 0.5$, quelle que soit la valeur initiale (x_0);
- $1 \leq \mu < \sqrt{2}$: les états de Γ_t apparaissent périodiquement ;
- $\sqrt{2} \leq \mu < 2$: les états de Γ_t deviennent un système chaotique avec une périodicité disparue;
- μ > 2 : les états de Γ_t divergent presque pour toutes les valeurs initiales.

II.3.2 Structure de C-PCANet

La méthode d'extraction de caractéristiques basée sur C-PCANet conserve la simplicité de PCANet, mais avec la possibilité de produire des caractéristiques biométriques révocables (cancelable), grâce à la couche supplémentaire de transformation, ce qui la rend plus sûre contre toute attaque ou intrusion. La figure ci-dessous (Fig. II.7) montre notre structure proposée du C-PCANet qui comprend deux stages. En général, cette structure peut être divisée en quatre étapes principales : la convolution, la transformation cachée, le hachage binaire et l'histogramme par blocs.

Afin de décrire le cadre du système, supposons que nous avons une entrée des N images d'apprentissage ($\psi \equiv \{L_i\}_{i=1}^{N}$) avec une taille de $h \times w$ et que la taille du patch, c'est-à-dire la taille du filtre de convolution (2D), pour le stage ℓ est

$$W_i^l = k_1^l \times k_2^l, \quad i \in [1 \cdots L_\ell], \; l \in [1 \cdots S_\ell] \qquad (12)$$

Où L_ℓ désigne le nombre de filtres dans la couche de convolution ℓ et S_ℓ est le nombre de couches de convolution. Il est important de noter que $k_j^\ell|_{j=1,2}$ est un nombre entier impair satisfaisant les conditions suivantes : $k_j \leq h$ et $k_j^\ell \leq w$.

☑ **Couche de convolution :** L'objectif principal de la couche de convolution est d'extraire les caractéristiques de l'image d'entrée. En pratique, ce processus examine d'abord un sous-ensemble de l'image, puis le filtre traverse l'image entière. La couche de convolution dans le C-PCANet comporte deux stages de sorte qu'à chaque stage, les filtres sont d'abord sélectionnés puis convolés avec l'image d'entrée.

- **Première stage :** Dans cette étape, tout d'abord, l'ensemble d'images Ψ_0 est utilisé pour créer les filtres de convolution L_1. Ainsi, pour chaque pixel dans chaque image $I_i \subset \Psi_0$, nous utilisons le patch de taille $k_1^1 \times k_2^1$ pour tronquer les voisins de ce pixel comme un segment $k_1^1 \times k_2^1$, puis les réorganiser en un vecteur (v) de taille $(k_1^1.k_2^1)\times 1$. Touts les vecteurs (colonnes) extraits de l'image i sont concaténés pour obtenir une matrice V_i :

$$V_i^1 = \left[v_1^{1,i}, v_2^{1,i}, \cdot\cdot, v_{n1}^{1,i}\right] \in \mathbb{R}^{k_1^1 \cdot k_2^1 \times n_1}, \quad i = 1,2,\cdot\cdot N \qquad (13)$$

Où $v_j^{1,i}\,|_{j=1,2,\ldots,\eta_1}$désigne le patch vectorisé j^{th} dans l'image I_i, N représente le nombre d'images d'apprentissage et η_1 est le nombre de vecteurs extraits, qui varie en fonction de la taille du patch ainsi que du chevauchement entre les zones tronquées.

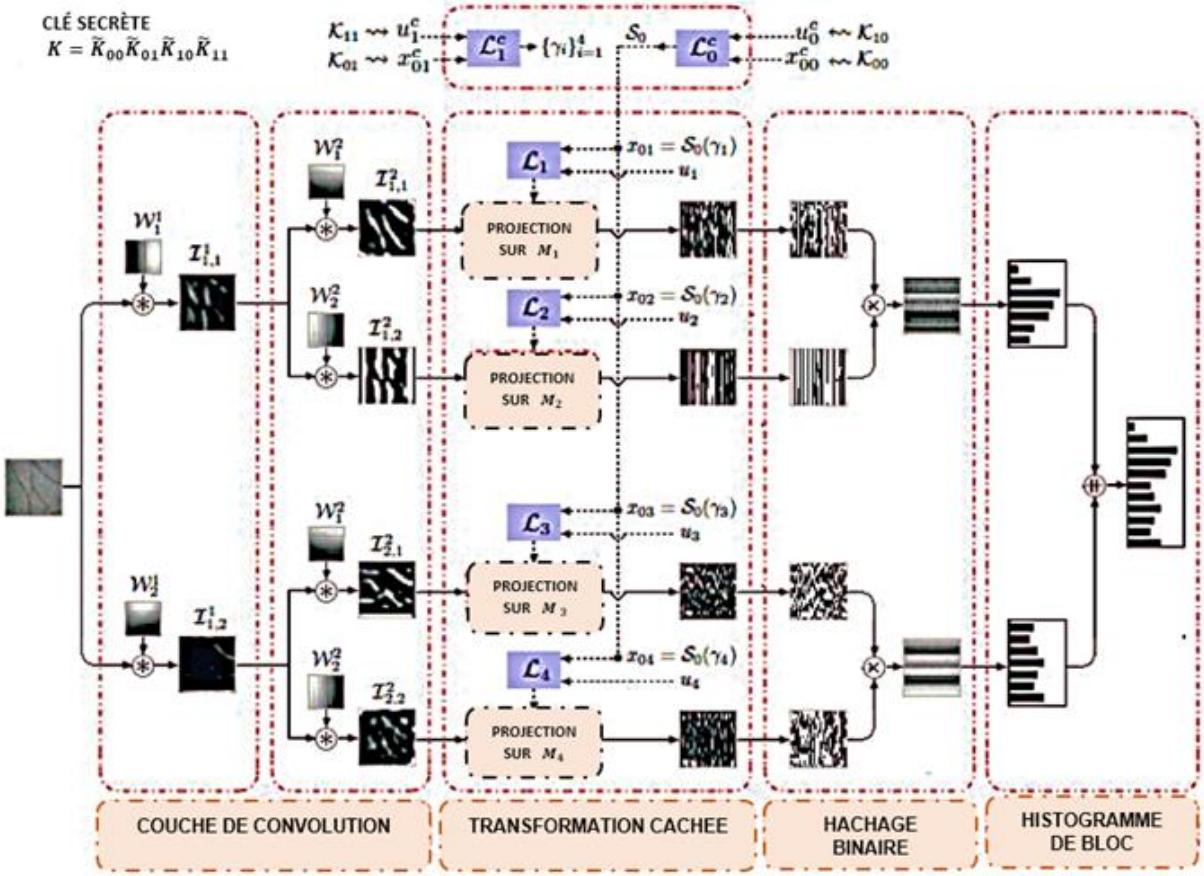


Fig II.7 : Schéma fonctionnel de la méthode d'extraction de caractéristiques profondes et révocables basée sur les cartes chaotiques. Un exemple de structure C-PCANet à 2 stages avec 2 filtres de convolution sur chaque stage.

Après avoir appliqué ce processus à toutes les images d'apprentissage Ψ_0, nous obtenons la matrice suivante :

$$V^1 = [v_1^1, v_2^1, \cdot, v_N^1.] \in \mathbb{R}^{k_1^1 \cdot k_2^1 \times N\, n_1} \qquad (14)$$

Appliquer maintenant la technique PCA sur la matrice V_1 pour obtenir la matrice U_1 composée des principaux vecteurs propres L_1 :

$$U^1 = \{u_1^1, u_2^1, \cdots u_{L_1}^1\} \in \mathbb{R}^{k_1^1 \cdot k_2^1 \times L_1} \quad (15)$$

Enfin, réorganiser les vecteurs propres $u_i^1|_{\mathrm{i=1,2,...L_1}}$pour obtenir les filtres (W_i^1) du premier stage :

$$W_i^1 = F_{k_1^1, k_2^1}(u_i^1) \in \mathbb{R}^{k_1^1 \times k_2^1},\ i = 1,2,\cdots L_1 \quad (16)$$

où $\mathcal{F}_{k_1^1\ k_2^1}$ est une fonction qui réorganise le vecteur $u_i \epsilon\ R^{k_1^1\ k_2^1}$ en une matrice $W_i^1\ \epsilon\ R^{k_1^1 \times k_2^1}$. Les sorties du premier stage de C-PCANet sont obtenues en filtrant toutes les images dans Ψ_0 par les filtres W_i^1:

$$I_{j,i}^1 = I_j \circledast W_i^1, \quad i = 1,2,\cdots L_1, \quad j = 1,2,\cdots N \quad (17)$$

Où $\circledast$ désigne un processus de convolution 2D. Il est important de noter que pour obtenir des images filtrées de même taille que I_j, une interpolation de frontière (à remplissage nul) est appliquée. Enfin, en utilisant les L_1 filtres, nous pouvons obtenir L_1 images filtrées pour chaque image d'entrée, donc pour toutes les images de la base d'apprentissage (Ψ_0), nous obtenons un ensemble d'images$\Psi \equiv \{I_i^1\}_{i=1}^{N.L_1}$.

- **Deuxième stage :** Comme pour le premier stage, les mêmes opérations sont effectuées dans le deuxième stage mais en utilisant les sorties du premier stage (Ψ_1, au lieu de Ψ_0,) comme base d'apprentissage. Dans le deuxième stage, l'ensemble d'images Ψ_1, est utilisé pour créer les filtres de convolution L_2, Ainsi, pour chaque pixel de chaque $I_i^1 \subset \Psi_1$, un patch de taille $k_1^1 \times k_2^1$ est utilisé pour tronquer les voisins de ce pixel en un segment $k_1^1 \times k_2^1$, puis réorganiser en un vecteur (v) de taille $(k_1^1 \times k_2^1)\times 1$. Tous les vecteurs obtenus sont concaténés comme suit :

$$V_i^2 = \left[v_1^{2,i}, v_2^{2,i}, \cdot, v_{n2}^{2,i}\right] \in \mathbb{R}^{k_1^2 \cdot k_2^2 \times n_2}, i = 1,2,\cdot N \cdot L_1 \quad (18)$$

Où $v_j^{2,i}\ |_{j=1,2,\ldots,\eta_1}$désigne le patch vectorisé j^{th} dans l'image I_i^1et η_2 est le nombre de vecteurs extraits. Pour toutes les images d'apprentissage en Ψ_1, nous obtenons la matrice suivante :

$$V^2 = \left[v_1^2, v_2^2, \cdot, v_{N \cdot L_1}^2\right] \in \mathbb{R}^{k_1^2 \cdot k_2^2 \times N \cdot L_1 \cdot n_2} \quad (19)$$

Comme pour le premier stage, la technique de PCA est appliquée sur la matrice V_2 pour obtenir la matrice U_2 qui se compose des vecteurs propres L_2:

$$U^2 = \{u_1^2, u_2^2, \cdots u_{L_2}^2\} \in \mathbb{R}^{k_1^1 \cdot k_2^1 \times L_2} \quad (20)$$

Réorganiser maintenant les vecteurs propres $v_1^2 \mid_{j=1,2,\ldots,L_2}$ pour obtenir les filtres PCA (W_i^2) du deuxième stage :

$$W_i^2 = F_{k_1^2,k_2^2}(u_i^2) \in \mathbb{R}^{k_1^2 \times k_2^2},\ i = 1,2,\cdots L_2 \quad (21)$$

Les sorties du deuxième stage sont obtenues en filtrant toutes les images de Ψ_2 par les filtres W_i^2:

$$I_{j,i}^2 = I_j \circledast W_i^2, \quad i = 1,2,\cdots L_2, \quad j = 1,2,\cdots N \cdot L_1 \quad (22)$$

Enfin, en utilisant les L_2, filtres, nous pouvons obtenir L_2 images filtrées pour chaque image d'entrée, donc pour toutes les images de Ψ_1, nous obtenons un ensemble d'images $\Psi_2 \equiv \{I_i^2\}_{i=1}^{N.L_1.L_2}$

☑ **Transformation cachée :** En tant que solution de sécurité, cette couche masque les modèles biométriques afin qu'ils puissent être annulés et remplacés par un autre à tout moment. Par conséquent, les modèles résultants changent avec le changement de la clé secrète tout en préservant, dans la mesure du possible, les performances du système biométrique. Dans notre système, nous avons adopté le principe de la transformation cachée en projetant des modèles sur une matrice orthogonale.

- **Génération de matrice :** La transformation cachée commence par générer les matrices de projection pour chaque sortie du dernier stage de convolution (dans notre système, le deuxième stage). Ainsi, notre système utilise plusieurs systèmes chaotiques, dont deux sont des principaux (clés secrètes) et les autres (qui sont utilisés pour générer les matrices de projection) changent en fonction du nombre de filtres dans le premier et deuxième stage. Il est important de souligner que la clé secrète de notre système peut être représentée par une valeur réelle ou entière codée en hexadécimal sur *M-bits*. Dans le paragraphe suivant, nous donnerons un exemple de clé secrète représentée sous forme de valeur entière.

Premièrement, utiliser la clé secrète (*K*) et les deux principaux systèmes chaotiques ($\mathcal{L}_0^C$ et $\mathcal{L}_1^C$) pour générer deux séquences (S_0^C et S_1^C) qui sont utilisées pour contrôler les systèmes chaotiques auxiliaires. Soit *K* une clé secrète représentée par :

$$K = \uplus_{i=0}^{1} \uplus_{j=0}^{1} \widetilde{K}_{ij} = \widetilde{K}_{00}\widetilde{K}_{01}\widetilde{K}_{10}\widetilde{K}_{11} \quad (23)$$

Où $\widetilde{K}_{ij}$est une valeur hexadécimale codée sur *M-bits* et ⊎ est une fonction de concaténation qui permet à une nouvelle valeur de rejoindre la chaîne hexadécimale. Ainsi, les paramètres des deux systèmes principaux ($\mathcal{L}_0^C$ et $\mathcal{L}_1^C$) sont définis comme suit:

$$\begin{cases} x_{0i}^c = \frac{K_{0i}}{2^M} \\ u_i^c = \alpha + \beta \left(\frac{K_{1i}}{2^M}\right) \end{cases}, \qquad i = 0{,}1 \tag{24}$$

Où K_{ij} est la représentation décimale de la valeur hexadécimale $\widetilde{K}_{ij}$ et la paire (α, β) est égale à (3.57, 0.43) et (1.41, 0,58) pour, respectivement, la carte logistique et la carte tente. Ces valeurs sont choisies pour que les deux systèmes chaotiques conservent toujours leur comportement chaotique.

Maintenant, pour chaque image filtrée en Ψ_0, génère deux séquences (S_0^C en utilisant $\mathcal{L}_0^C$ et S_1^C en utilisant $\mathcal{L}_1^C$). Ces deux séquences sont utilisées pour déterminer les états initiaux des systèmes chaotiques auxiliaires. La séquence S_1^C a $L_1.L_2$ éléments et est définie comme suit :

$$S_1^c = \Gamma_v(x_{01}^c, u_1^c) = \{s_{1j}\}_{j=1}^{L_1 \cdot L_2} \tag{25}$$

Les éléments de S_1^C (générés par $\mathcal{L}_1^C$) sont utilisés comme des coordonnées dans S_0^C, ils doivent donc devenir des entiers. En effet, la séquence S_1^C est normalisée dans l'intervalle [1 500], comme suit :

$$S_1^c = 1 + \lfloor 10^5 \cdot S_1^c \rfloor (\text{mod } 500) \in [1 \cdots 500] \tag{26}$$

Où $\lfloor \bullet \rfloor$ désigne la partie entière. Pour déterminer les états initiaux des systèmes chaotiques auxiliaires, la séquence S_0^C (générée par $\mathcal{L}_0^C$) est utilisée :

$$S_0^c = \Gamma_v(x_{00}^c, u_0^c) = \{s_{0j}\}_{j=1}^{500} \tag{27}$$

Les états initiaux des systèmes chaotiques auxiliaires sont définis comme suit :

$$\begin{cases} x_{0i} = S_0^c(\gamma_i) \\ \gamma_i = S_1^c(i) \end{cases} \qquad i = 1{,}2{,}\cdots L_1 L_2 \tag{28}$$

De plus, les paramètres de contrôle des systèmes chaotiques auxiliaires sont définis comme suit :

$$u_i = \alpha + \beta\left(\frac{\widehat{K_i}}{2^M}\right) \tag{29}$$

Il est important de noter qu'une méthode d'optimisation est utilisée pour sélectionner les différents $\widetilde{K}_i$(donc u_i) afin de maximiser le taux d'identification du système biométrique.

Parce que, dans notre implémentation, les différents K_i sont binaires, un algorithme génétique (GA) est utilisé pour l'opération d'optimisation. Enfin, chaque système chaotique auxiliaire génère une séquence de longueur $h \cdot w$:

$$S_i = \Gamma_v(x_{0i}, u_i) = \{s_j\}_{j=1}^{h \cdot w} \quad i = 1,2,\cdots L_1 L_2 \tag{30}$$

Chaque séquence (S_i) est ensuite réorganisée pour former une matrice (M_i) de même taille que l'image d'entrée :

$$M_i = F_{h,w}(S_i) \in \mathbb{R}^{h \times w} \quad i = 1,2,\cdots L_1 L_2 \tag{31}$$

Une fois les différentes matrices (M_i) sont générées et pour garantir des sorties de projection non corrélées, une opération de factorisation est appliquée à chaque matrice. Pour ce faire, nous avons utilisé la factorisation *QR* [53] qui est l'une des opérations importantes de l'analyse matricielle dans le traitement du signal / image et les statistiques.

Soit M_i une matrice composée des différentes colonnes v_i définies comme suit :

$$M_i = \left[v_0^i, v_1^i, v_2^i, \cdots v_w^i\right] \tag{32}$$

La factorisation QR effectue la décomposition orthogonale-triangulaire de la matrice M_i, où cette matrice est décomposée en deux matrices, dont l'une est une matrice unitaire réelle (Q) et l'autre est une matrice triangulaire supérieure (R).

$$M_i = Q_i \cdot R_i, \quad R_i \in \mathbb{R}^{w \times w}, \; Q_i \in \mathbb{R}^{h \times w} \tag{33}$$

La matrice résultante Q_i a la même dimension que M_i mais avec des colonnes orthogonales. Enfin, ces matrices orthogonales ($Q_i|_{i=1}^{L_1 L_2}$) sont utilisées pour transformer les sorties du dernier stage de convolution dans un autre espace pour leur permettre d'être cachées.

- **Processus de projection :** En utilisant les matrices Q_i, nous pouvons projeter les sorties du dernier stage de convolution de chaque image :

$$\hat{I}_i^2 = I_i^2 \cdot Q_i \in \mathbb{R}^{h \times w}, \; i = 1,\cdots L_1 L_2 \tag{34}$$

Après avoir projeté toutes les images en Ψ_2, nous obtenons un ensemble d'images de sorties $\Psi_3 \equiv \{\hat{\mathrm{i}}_{\mathrm{i}}^2\}_{\mathrm{i}=1}^{\mathrm{N.L_1.L_2}}$.

☑ **Couche de hachage binaire :** La condition la plus importante pour le système biométrique révocable est la non-réversibilité du vecteur caractéristique biométrique résultant. Heureusement, cette couche permet de vérifier cette condition, dans laquelle les images

projetées (en Ψ_3) sont binarisées. Dans cette couche, les L_2 sorties du dernier couche , pour chaque image, sont converties en une image à valeur entière.

- **Quantification binaire :** Le processus de quantification binaire transforme une valeur réelle en valeur binaire. En fait, le principe de seuillage est appliqué, comme suit :

$$I_{i,j}^{b} = \begin{cases} 1 & \text{if } \hat{I}_{i,j}^{2} \geq \tau_0 \\ 0 & \text{if } \hat{I}_{i,j}^{2} < \tau_0 \end{cases}, \qquad i = 1,\cdot\cdot L_1, \quad j = 1,\cdot\cdot L_2 \tag{35}$$

Où τ_0 est le seuil de binarisation. En pratique, ce seuil est choisi égal à 0 ($\tau_0 = 0$) car les résultats de la projection ont la même probabilité d'être négatifs ou positifs. Après avoir converti toutes les images en Ψ_3, nous obtenons un ensemble d'images binaires $\Psi_4 \equiv \{\mathrm{I}_{\mathrm{i}}^{\mathrm{b}}\}_{\mathrm{i}=1}^{\mathrm{N.L_1.L_2}}$.

- **Conversion binaire en décimal :** Dans cette étape, la chaîne de binaires (L_2-bits) autour de chaque pixel est convertie en valeur entière. Pour cela nous utilisons le polynôme de décodage suivant :

$$I_i^3 = \sum_{j=1}^{L_2} I_{i,j}^{b} \cdot 2^{j}, \qquad i = 1,\cdot\cdot L_1 \tag{36}$$

Après l'opération de décodage de chaque groupe L_2 dans Ψ_4 séparément, nous obtenons un ensemble d'images à valeurs entières $\Psi_5 \equiv \{\mathrm{I}_3^{\mathrm{b}}\}_{\mathrm{i}=1}^{\mathrm{N.L_1}}$.

☑ **Histogramme par bloc :** Afin de réduire la taille du vecteur de caractéristiques de chaque image d'entrée, la représentation avec l'histogramme est utilisée comme opération de regroupement. Ainsi, l'histogramme de chaque image (parmi les L_1 images) est calculé et tous ces histogrammes sont concaténés pour former le vecteur de caractéristiques.

- **Partition de blocs :** Pour obtenir le vecteur de caractéristiques de chaque image d'entrée, nous partitionnons d'abord chacune des L_1 images en blocs. Ainsi, en supposant que la taille du bloc utilisé (B) est $b_1 \times b_2$ avec un taux de chevauchement o, chaque image est partitionnée en $\mathcal{N}_B$ blocs comme suit :

$$\mathcal{N}_B = \left\lfloor \frac{h - b_1}{o} + 1 \right\rfloor \times \left\lfloor \frac{w - b_2}{o} \right\rfloor \tag{37}$$

Pour toutes les L_1 images de chaque image d'entrée (i), on obtient un ensemble de blocs Φ_{i} définis comme suit :

$$\begin{aligned} \Phi_i = \quad & \{B_1, B_2,\cdot\cdot B_{L_1.\mathcal{N}_B}\} \in \mathbb{R}^{(b_1 \times b_2) \times L_1.\mathcal{N}_B} \\ & B \in \mathbb{R}^{b_1 \times b_2}, \quad i = 1 \cdot\cdot N \end{aligned} \tag{38}$$

Où B_c désigne le bloc c^{th}

- **Histogramme :** Dans cette étape, un histogramme, avec 2^{L_2} cases (*bins*), pour chaque bloc (B_c) est calculé. Cela signifie que chaque valeur entière est définie comme un *bin* et qu'un vecteur représentant l'histogramme est construit.

$$\mathcal{H}_c(i) = \sum_{x=0}^{b_1-1}\sum_{y=0}^{b_2-1}[\sum_{i=0}^{2^{L_2}-1}\mathcal{G}(B_c(x,y)=i)] \quad c = 1,2,\cdots L_1\mathcal{N}_B \tag{39}$$

Où $\mathcal{G}(\sigma)$ = 1 lorsque σ est vrai et $\mathcal{G}(\sigma)$= 0, sinon. Tous les histogrammes calculés sont donc concaténés en un seul vecteur pour obtenir le vecteur de caractéristiques biométriques (modèle biométrique) de chaque image d'entrée (*i*) :

$$\mathcal{T}_i = \left[\mathcal{H}_2, \mathcal{H}_2, \cdots \mathcal{H}_{L_1\mathcal{N}_B}\right] \epsilon \mathbb{R}^{2^{L_2}L_1\mathcal{N}_B} \quad i = 1,\cdots N \tag{40}$$

Enfin, les vecteurs de caractéristiques de toutes les images d'apprentissage en entrée sont obtenus comme suit :

$$\mathcal{T} = [\mathcal{T}_1, \mathcal{T}_2, \cdot \cdot \mathcal{T}_N\ ] \epsilon \mathbb{R}^{2^{L_2}NL_1\mathcal{N}_B} \tag{41}$$

Il est important de noter que la longueur et la précision du vecteur ($\mathcal{T}_i$), de chaque image d'entrée, change selon à la taille du bloc ($b_1 \times\ b_2$) et le taux de chevauchement (*o*).

II.4 Conclusion

Ces dernières années, la disponibilité de dispositifs d'acquisition de données biométriques à faible coût et le développement impressionnant de la technologie numérique, en plus de la haute sécurité requise pour les données sensibles, ont conduit à une croissance significative de l'utilisation des technologies biométrique pour la reconnaissance automatique de l'identité humaine. Malheureusement, les données biométriques d'un individu sont très sensibles en raison de leur association permanente avec l'utilisateur, ce qui justifie des préoccupations croissantes concernant la vie privée et l'anonymat des individus face à toute tentative de piratage. Par conséquent, de nombreuses études se sont concentrées sur la recherche de moyens d'extraire des caractéristiques biométriques qui peuvent être annulées et remplacées à tout moment lorsqu'elles sont compromises. Dans ce chapitre, nous avons présenté quelques solutions proposées pour protéger le modèle biométrique. En outre, une nouvelle méthode d'extraction de caractéristiques profondes et révocables (C-PCANet) utilisant des cartes chaotiques est proposée.

Chapitre 3

Résultats expérimentaux

Résumé

Afin de représenter efficacement les caractéristiques biométriques d'une empreinte, les images capturées par le dispositif de capture sont généralement analysées par la fonction d'extraction de caractéristiques. D'un point de vue sécurité, la nouvelle représentation produite par cette fonction doit être unique à chaque personne et peut être révoquée en cas de piratage. En fait, ce chapitre vise à évaluer la méthode d'extraction de caractéristiques proposée (C-PCANet) sous deux aspects très importants, à savoir la précision du système d'identification des personnes et la robustesse contre les tentatives de piratage. Tous les tests ont été effectués à l'aide d'une base de données biométrique connue et disponible.

Introduction

Étant donné que la méthode proposée dans ce travail est une version améliorée de la méthode d'extraction de caractéristiques profondes connue sous le nom de *PCANet deep learning*, dans ce chapitre, nous présenterons d'abord un système biométrique utilisant la méthode d'extraction de caractéristiques profondes basées sur PCANet et un classifier SVM. Ensuite, nous testons les performances et la robustesse de notre système (C-PCANet) contre les tentatives d'attaques. Ces tests seront effectués sur une base de données très connue et accessible au public.

III.1 Modalités biométriques utilisées

Le système biométrique révocable proposé dans ce travail est basé sur deux empreintes obtenues de la paume de la main, à savoir l'empreinte palmaire et l'empreinte du réseau veineux. Récemment, ces deux empreintes ont été largement utilisées pour améliorer la reconnaissance des personnes et les distinguer l'une de l'autre dans diverses applications en raison de la variété des caractéristiques qui peuvent en être extraites. Les caractéristiques de la texture ainsi que les propriétés de la structure veineuse sont les traits distinctifs extraits de l'image de la paume, capturée sous lumière visible et infrarouge.

III.1.1 Empreint palmaire (palmprint)

L'empreinte de la paume, qui est la surface interne de la main entre le poignet et les doigts, est connue sous le nom d'empreinte palmaire [14] et se compose principalement de lignes principales et secondaires (voir Fig. III.1). Les systèmes biométriques qui utilisent ces empreintes sont parmi les systèmes les plus populaires utilisés dans les applications de

contrôle d'accès physiques et/ou logiques en raison de leur discrimination et de leur grande acceptabilité par les personnes. En plus des lignes, cette empreinte contient des caractéristiques distinctives supplémentaires telles que des ridules (*wrinkles*) et des crêtes (*ridges*), qui peuvent être extraites à partir d'images à basse résolution.

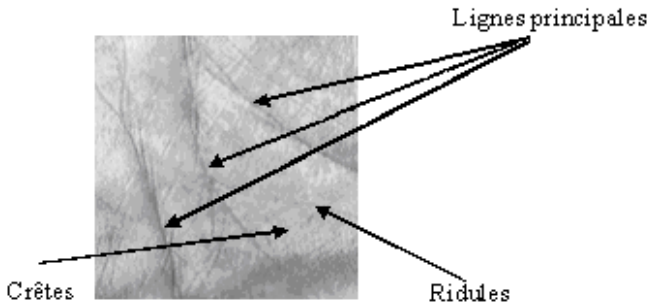


Fig III.1 : Empreinte palmaire.

III.1.2 Empreinte de réseau veineux de la paume (palm-vein)

La technologie d'imagerie infrarouge est utilisée pour capturer le motif du réseau veineux dans la paume. La structure de ce réseau peut être utilisée pour distinguer les personnes et ainsi être utilisée dans le système de reconnaissance. La figure III.2 montre l'empreinte du réseau veineux de la paume [19].

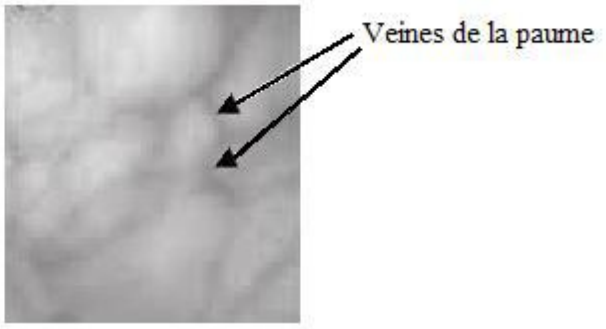


Fig III.2 : Empreinte du réseau veineux de la paume.

La technologie d'acquisition de cette empreinte repose sur des diodes qui émettent une lumière proche infrarouge (IR) sur la paume de la main, où le tissu cutané réfléchit ces rayons et l'hémoglobine dans les veines palmaires les absorbe, réduisant leur réflexion, ce qui donne une apparence noire du réseau veineux sur l'image capturée. Ce type de biométrie est bien utilisé pour gérer l'accès dans des lieux protégés et pour remplacer les mots de passe pour se connecter à un réseau. De toute évidence, la veine de la paume est beaucoup plus difficile à falsifier que l'empreinte palmaire.

III.2 Description de la base de données

Le système proposé dans ce travail utilise deux modalités biométriques, pour cela, une base de données multi-biométrique appropriée doit être utilisée pour évaluer ses performances. Pour des raisons de crédibilité de nos résultats, cette base de données doit être fiable et reconnue. Ainsi, la base de données utilisée contient 250 personnes différentes représentant les étudiants et le personnel de l'Université de PolyU [54]. Dans cette base, 195 personnes sont des hommes, la majorité entre 20 et 60 ans. La collecte d'images a été réalisée en deux sessions à 9 jours d'intervalle. Au cours de chaque session, six images de la paume des mains (gauche et droite) ont été capturées de chaque personne. Par conséquent, cette base de données contient 6000 images pour bande spectrale de 500 paumes différentes. Dans chaque acquisition, le système collecte quatre bandes spectrales {rouge *(Red)*, verte *(Green)*, bleue *(Blue)* et proche infrarouge (*Near-Infrared-NIR*)}. La bande NIR représente les réseaux veineux de la paume (*Palm-vein (PLV)*), tandis que les trois bandes spectrales (rouge, verte et bleue) après conversion en une image en niveaux de gris représentent l'empreinte palmaire (palmprint (PLM)). Fig. III.3 montre un exemple de l'image multi-spectrale capturée.

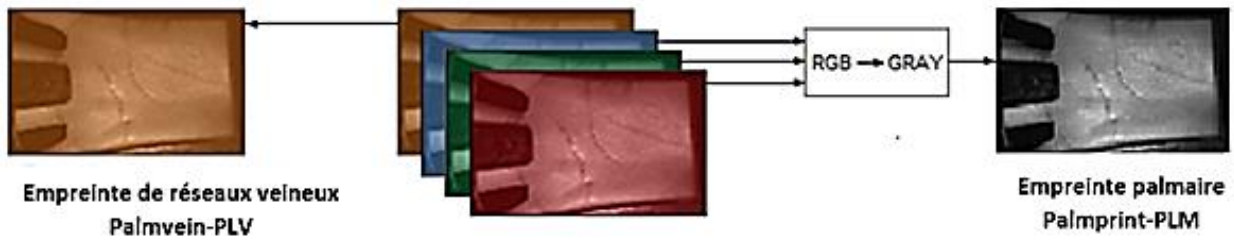


Fig III.3 : Exemple d'une image multi-spectrale dans PolyU.

III.3 Performance du système biométrique

Dans ce travail, nous nous intéressons particulièrement aux modes d'identification (ensemble ouvert et ensemble fermé). Les expériences menées dans cette étude ont utilisé une base de données de 400 personnes (12 images pour chaque modalité biométrique). Quatre images, avec un total de 1 200 images, sont utilisées pour construire la base de données de référence et les 3 200 images restantes pour tester les performances des systèmes biométriques basés sur PCANet/C-PCANet. Pour la mesure des performances, les distributions imposteurs et authentiques sont générées par 638400 et 3200 comparaisons, respectivement. Dans l'ensemble, les tests expérimentaux sont divisés en trois parties. Dans la première partie, le système biométrique basé sur PCANet est testé et évalué à l'aide de la base de données décrite ci-dessus. La deuxième partie de ces tests présente et discute les performances du système biométrique révocable basé sur C-PCANet. Enfin, la troisième

partie se concentre sur l'analyse de sécurité fournie par la couche supplémentaire de transformation.

III.3.1 Système biométrique non sécurisé (PCANet)

Avant de commencer à tester le système basé sur C-PCANet (système sécurisé), qui dépend principalement du schéma PCANet (non sécurisé), il faut d'abord sélectionner la meilleure configuration du PCANet en choisissant ses meilleurs paramètres, qui ont un impact significatif sur les performances du système biométrique. Par conséquent, ils doivent être choisis séparément et soigneusement pour les deux modalités biométriques (PLM et PLV) qui sont l'objectif principal de la première partie. Puisque l'efficacité de PCANet dépend non seulement du nombre des stages mais aussi du nombre de filtres (k_ℓ) et de la taille des filtres ($k_1^\ell \times k_2^\ell$) à chaque stage (ℓ), nous pouvons réaliser une série d'expériences (utilisant les modalités PLM et PLV) pour sélectionner efficacement les paramètres appropriés. Ainsi, la variation de ces paramètres permet de donner plusieurs vecteurs de caractéristiques biométriques, et il est donc possible de choisir expérimentalement et empiriquement les paramètres appropriés, qui peuvent effectivement améliorer la précision du système d'identification biométrique, en modifiant à chaque fois ces paramètres et choisissez le meilleur qui offre les meilleures performances.

Avec la variation des paramètres, à chaque fois les performances du système biométrique, basé sur PCANet, sont évaluées. Ainsi, dans le premier stage ($\ell = 1$), nous essayons de sélectionner le nombre de filtres (L_1) parmi cinq nombres prédéfinis (2, 4, 6, 8, 10). De plus, la taille des filtres de convolution ($k_1^\ell \times k_2^\ell$) est choisie parmi cinq tailles prédéfinies (9 × 9, 11 × 11, 13 × 13, 15 × 15, 17 × 17). Pour examiner l'effet de ces paramètres sur la précision du système d'identification biométrique (en mode ensemble ouvert), nous illustrons, pour les deux modalités biométriques, les performances du système sous forme d'intervalle de confiance (d_{CI}) et les résultats obtenus sont présentés dans les figures III.4 et III.5 pour le PLM et le PLV, respectivement. Ainsi, à partir des courbes de la Fig. III.4.*(a)* et Fig. III.5.*(a)*, nous pouvons tirer deux remarques importantes:

i) Avec l'utilisation de deux modalités biométriques (PLM ou PLV), une configuration PCANet à un stage peut donner une performance parfaite (séparation des deux distributions, imposteur et authentique). Par conséquent, il n'est pas nécessaire de réexaminer le système biométrique en deux ou plusieurs stages.

ii) Presque pour les deux modalités biométriques, la plupart des paramètres conduisent à des performances élevées, même celles qui n'ont pas donné une performance parfaite, leurs performances dépassant en fait 99,998%.

Dans les deux modalités biométriques (PLM et PLV), on observe qu'en général, l'intervalle de confiance (d_{CI}) devient positif lorsque l'on utilise un plus grand nombre de filtres (séparation totale des deux distributions). Pour la modalité PLM (voir Fig. III.4.*(a)*), lorsque nous utilisons 4, 6, 8 ou 10 filtres, la taille du filtre de convolution de $k_1^1 \times k_2^1 = 15 \times 15$ donne la meilleure performance. Ainsi, un grand d_{CI} = 0.171) est obtenu dans le cas de huit filtres de convolution (voir figure III.4.*(b)*). Bien sûr, dans ce cas, le système fonctionne avec un taux d'acceptation authentique (GAR) maximum égal à 100,00% et *zéro FAR* (FAR = 0,000%) ainsi qu'un *zéro FRR* (FRR = 0,000%) au seuil de sécurité (T_o) de 0,350 (voir Fig. III.4.*(c)*). Il est important de noter que tous les scores obtenus dans le système d'identification biométrique basé sur le PLM sont normalisés à l'intervalle [0,1] en le divisant par un facteur de 3.

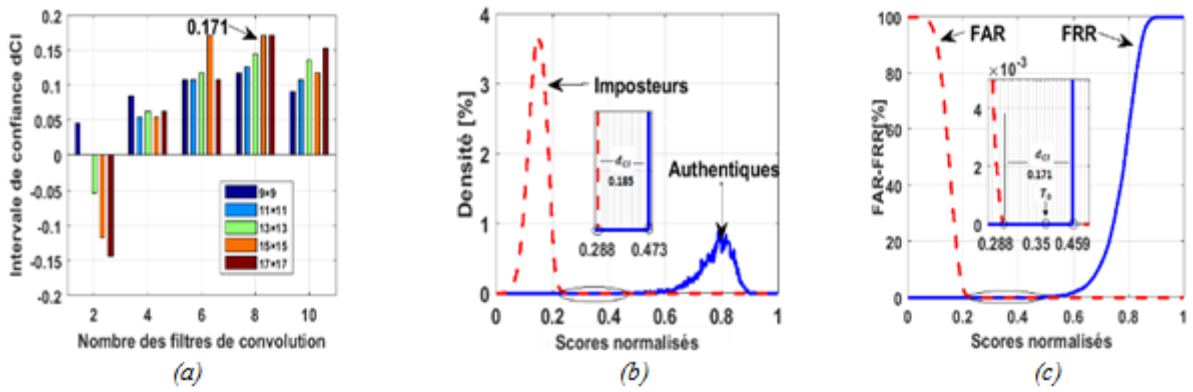


Fig III.4 : Sélection des paramètres PCANet d'un système basé sur PLM. *(a)* Intervalle de confiance (d_{CI}), *(b)* Distributions authentiques et imposteurs, and *(c)* Courbes FAR et FRR.

De même, six filtres à convolution de taille 9 × 9 fournissent toujours la meilleure performance dans un système biométrique basé sur PLV (voir Fig. III.5. *(a)*). Cette similitude est due au fait que les deux modalités biométriques sont similaires, mais la supériorité de la modalité PLM réside dans sa richesse en caractéristiques intrinsèques telles que les lignes principales et les rides. Dans cette configuration, le système fonctionne parfaitement avec un d_{CI} acceptable égal à 0,081 (voir Fig. III.5. *(b)*) et *zéro FRR*, *zéro FAR* et GAR parfait à un seuil (T_o) égal à 0,340 (voir Fig. III.5. *(c)*). Nous avons également examiné le mode d'identification ensemble fermé et pour les deux modalités, le système fonctionne toujours efficacement avec un taux d'identification (ROR) égal à 100,00% et un rang d'identification

parfaite (RPR) égal à 1 dans toutes les configurations testées. Les résultats des tests expérimentaux obtenus dans cette section indiquent clairement que l'utilisation des modalités biométriques basées sur la paume en conjonction avec la méthode d'extraction de caractéristiques basée sur l'apprentissage profond (PCANet) peut fournir de meilleures performances, ce qui peut permettre d'obtenir un système biométrique efficace qui peut être utilisé dans de nombreux applications nécessitant une sécurité élevée, en particulier si ces applications traitent de petites ou moyennes bases de données.

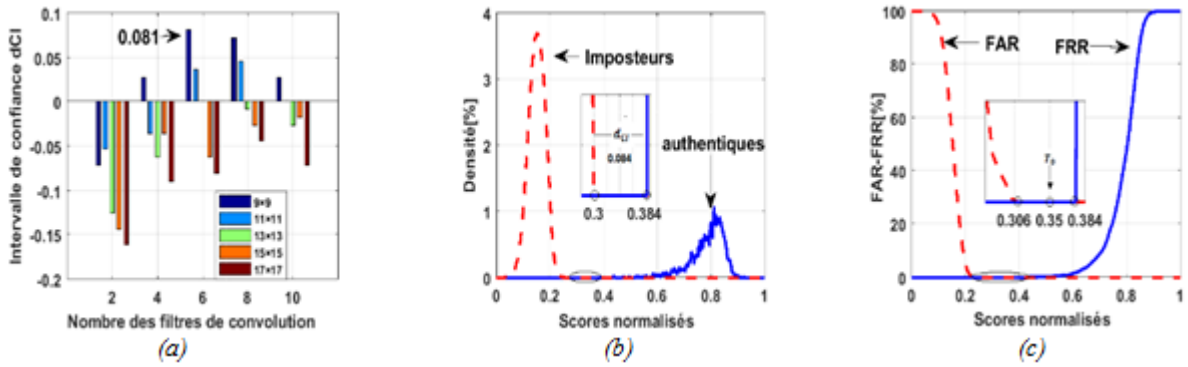


Fig III.5 : Sélection des paramètres PCANet d'un système basé sur PLV. *(a)* Intervalle de confiance (d_{CI}), *(b)* Distributions authentiques et imposteurs, and *(c)* Courbes FAR et FRR.

Dans la partie suivante et pour sécuriser le système biométrique, nous sélectionnons la meilleure configuration qui est: $\ell = 1$, $k_1^1 \times k_2^1 = 15 \times 15$ et $L_1 = 8$, $\ell = 1$, $k_1^1 \times k_2^1 = 9 \times 9$ et $L_1 = 6$, pour les deux modalités biométriques PLM et PLV, respectivement, afin de construire un système biométrique révocable (C-PCANet).

III.3.2 Système biométrique sécurisé (C-PCANet)

Dans cette sous-section, nous examinons en détail les performances du système biométrique révocable proposé (C-PCANet). Avant de commencer, nous devons mentionner que contrairement à notre proposition de système biométrique révocable, tous les systèmes développés dans la littérature ne discutent que du mode d'identification en ensemble fermé. La force de notre système réside donc dans sa validité à la fois dans les deux modes d'identification (ensemble ouvert et fermé).

***a)* Mode d'identification ensemble ouvert :** dans ce mode et afin d'évaluer sérieusement le système biométrique révocable proposé, trois points principaux liés à son comportement doivent être examinés. Ces points sont :

i) Les scores trouvés par le système biométrique basé sur C-PCANet devraient présenter des distributions authentiques et imposteurs presque similaires à celles présentées par le système biométrique basé sur PCANet.

ii) Les distributions authentiques et imposteurs trouvées par le système biométrique basé sur C-PCANet ne devraient pas changer de manière significative lors de l'utilisation de différentes clés secrètes.

iii)Un changement soudain des scores authentiques en dessous du seuil de sécurité (T_o) lors de l'utilisation d'une fausse clé de sécurité (tentative d'attaque). Le meilleur cas est lorsque tous les scores d'attaque sont inférieurs au point *zéro FAR*.

Pour le premier point, nous avons comparé les performances de PCANet et C-PCANet pour les deux modalités biométriques afin de voir le changement qui peut se produire sur le comportement de PCANet lors de l'intégration de la couche de sécurité (voir Fig. III.6. *(a)* et Fig. III.6. *(b)* pour les modalités PLM et PLV, respectivement). A partir de ces figures, on peut clairement voir que le comportement général du système PCANet après l'intégration de la couche de sécurité n'a pas changé.

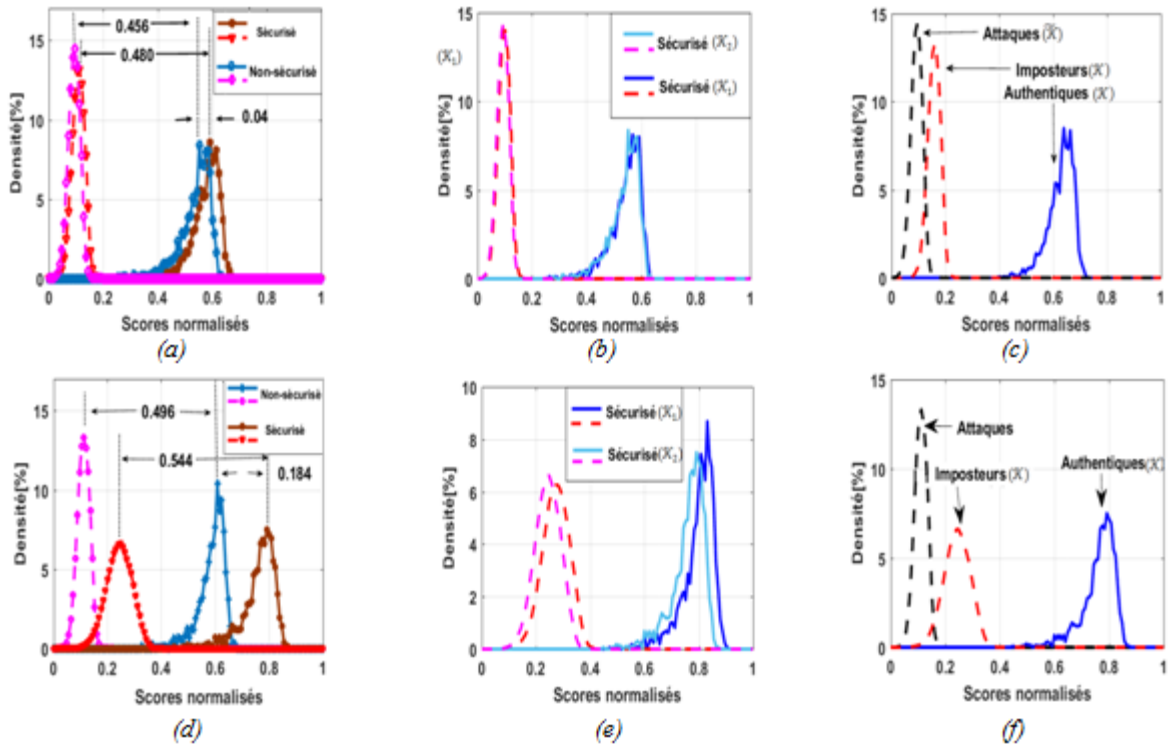


Fig III.6 : Comportement du système d'identification biométrique basé sur C-PCANet, en mode ensemble ouvert. *(a), (b), (c)* Système biométrique utilisant la modalité biométrique PLM et, *(d), (e), (f)* Système biométrique utilisant la modalité biométrique PLV.

Pour la modalité PLM, la distance moyenne entre les deux distributions (authentique et imposteur) est de 0,450 pour PCANet, ce qui est approximativement la même pour C-PCANet (0,480) et les deux distributions ont été légèrement décalées vers la gauche d'une légère distance de 0,04. De plus, pour la modalité PLV, la distance moyenne entre les deux distributions était de 0,496 et 0,544 pour PCANet et C-PCANet, respectivement, avec un léger décalage vers la droite de 0,184. Il est important de noter que les distributions du système d'identification biométrique révocable basé sur C-PCANet, pour les modalités biométriques PLM et PLV, ont été obtenues par des clés secrètes codées en hexadécimal avec M = 16 *bits* ($K = \widetilde{K_{00}}\widetilde{K_{01}}\ \widetilde{K_{10}}\ \widetilde{K_{11}}$), K_{PLM1} = "1A02-1D20 -0395-ACFF » et K_{PLV1} = « 0ABC2226-A12C-100D », respectivement.

Pour le deuxième point, il suffit de prouver que le système C-PCANet peut fonctionner avec un seul seuil de sécurité, quel que soit le changement de clé secrète, et cela ne peut être réalisé que si les différents scores changent dans le même intervalle. Par conséquent, nous avons exécuté le système C-PCANet avec deux clés secrètes différentes et comparé les différentes distributions obtenues (pour les deux modalités biométriques) et les résultats sont présentés dans la Fig. III.6. *(b)* et la Fig. III.6. *(e)*. En effet, en plus des clés secrètes mentionnées ci-dessus, nous avons utilisé deux autres clés, K_{PLM2} = "BA0F-20BB-2C3A-9568" (pour PLM) et K_{PLV2} = "1BD3-FF20-CCF4-0625" (pour PLV). On peut voir sur ces figures que les distributions authentiques et imposteurs ont varié dans le même intervalle pour les deux modalités biométriques et, par conséquent, le système C-PCANet peut effectivement utiliser le même seuil (T_0) pour toutes les clés secrètes. Par conséquent, ces résultats prouvent que le système proposé peut être utilisé dans le système d'identification biométrique qui fonctionne en mode ensemble ouvert.

Dans le troisième point, nous cherchons à examiner le comportement de C-PCANet lors d'une attaque. Ainsi, dans nos tests, tous les utilisateurs sont enregistrés dans la base de données par K_{PLM_1} et K_{PLV_1} et dans le test d'identification nous utilisons d'autres clés, K_e (K_{PLM3} = "5E3C-AAAA-11FF-1E36" et K_{PLV3} = "000B-10AA-32FC-CC0D"). Ainsi, pour voir les performances des systèmes d'identification (en mode ensemble ouvert) vis-à-vis des attaques, sur la Fig. III.6.*(c)* et la Fig. III.6.*(f)*, nous illustrons les résultats des modalités PLM et PLV. Dans ces figures, il est clair que tous les scores d'attaque sont complètement décalés en dessous du seuil de sécurité (en dessous de *zéro FAR*). Il est très remarquable que le point, où le taux d'attaque est nul (ASR = 0), soit beaucoup plus petit que le point *zéro FAR*, ce qui reflète l'efficacité et la robustesse de notre système contre toute attaque éventuelle.

Enfin, afin de montrer les performances du système d'identification biométrique révocable (en mode ensemble ouvert) basé sur C-PCANet, un tableau présentant les résultats expérimentaux, exprimés en taux d'erreur égal (EER), FAR, FRR et ASR en fonction du seuil (T_0) est généré (voir le tableau 1).

Tableau III-1 : Résultats du test de système d'identification biométrique révocable (système unimodal)

Modalités	ERR		FAR à FRR = 0		FRR à FAR = 0		ASR à FAR = 0	
	T_o	EER	T_o	FAR	T_o	FRR	T_o	ASR
PLM (K_{PLM1})	0.0670	0.0000	0.2610	0.0000	0.2070	0.0000	0.2070	0.000
PLV (K_{PLV1})	0.0230	0.0010	0.4320	0.0001	0.4590	0.1250	0.4590	0.000

À partir de ce tableau, nous pouvons clairement voir l'efficacité de notre système proposé, qui peut fonctionner en toute sécurité avec un EER minimum égal à 0,0000% et 0,001% à un seuil égal à 0,0670 et 0,0230 pour les modalités biométriques PLM et PLV, respectivement, pour un taux d'attaque nul (ASR = 0,0000%).

***b)* Mode d'identification ensemble fermé :** dans ce mode, le système ne doit traiter que les utilisateurs déjà enrôlés dans la base de données du système. Pendant le processus d'identification, le système calcule les scores de similitude d'un vecteur de caractéristiques d'entrée avec tous les vecteurs de caractéristiques pré-stockés pour prendre la décision appropriée. Soit un système biométrique contenant *N* vecteurs de caractéristiques pré-stockés ($T_i\,|_{i=1}^{N}$) et un utilisateur (*P*), qui a un vecteur de caractéristiques *T*, veut accéder au système. En effet, le système calcule *N* scores de similarité (*d*) comme suit :

$$V_d = [d_1, d_2, \cdots, d_N] \qquad (1)$$

Après cela, le système détermine l'identité de la personne en fonction du score de similitude le plus élevé comme suit :

$$P \equiv P_j,\ j = \arg \max_{J\epsilon[1\cdots N]} (V_d) \qquad (2)$$

Dans ce mode, le système détermine l'identité de la personne en fonction de la personne la plus proche stockée dans la base de données même s'il s'agit d'une attaque. Pour pallier cette faiblesse, deux solutions peuvent être proposées :

i) Avec la première solution, un processus d'identification en mode ensemble ouvert est d'abord exécuté, si la personne est acceptée (un client dans la base de données système), le processus d'identification en mode ensemble fermé est exécuté pour déterminer son identité exacte.

ii) Avec la seconde solution, un processus d'identification en mode ensemble fermé est effectué, puis un processus de seuillage est utilisé pour déterminer l'identité de la personne.

L'objectif de cette sous-section est d'évaluer le système d'identification en mode ensemble fermé lors d'une tentative d'attaque. Ainsi, le système est exécuté dans sa structure sécurisée et non sécurisée avec des clés secrètes correctes et erronées et tous les scores obtenus sont ensuite comparés. Ainsi, pour les deux modalités biométriques, la figure III.7 montre les résultats de PCANet et C-PCANet avec les clés secrètes utilisées précédemment (clé correcte $\{K_1, K_2\}$, clé erronée $\{\widetilde{K} = K_3\}$). De ces figures, nous pouvons tirer deux remarques importantes:

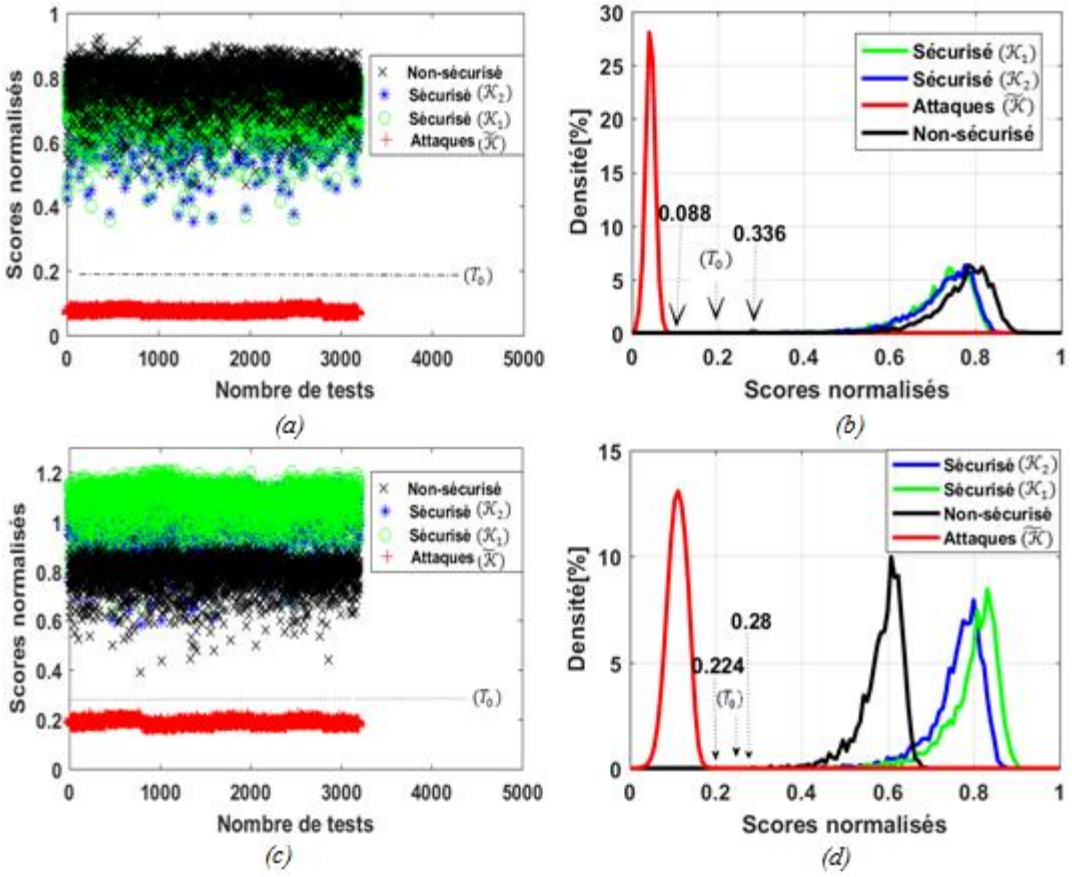


Fig. III.7 Comportement du système d'identification biométrique, en mode ensemble fermé, basé sur C-PCANet. *(a), (b)* Système biométrique basé sur C-PCANet utilisant la modalité biométrique PLM et, *(c), (d)* Système biométrique basé sur C-PCANet utilisant la modalité biométrique PLV.

Tout d'abord, l'examen visuel de la Fig. III.7.*(a)* et la Fig. III.7.*(c)* montre qu'il n'y a pas de dégradation des performances dans le système C-PCANet pour les deux clés secrètes correctes car les résultats obtenus par PCANet et C-PCANet sont approximativement dans la même plage et ceci est pour les deux modalités biométriques. Deuxièmement, de petites

valeurs des scores sont obtenues lorsque clé erronée est utilisée (lors d'une attaque), ce qui indique clairement l'efficacité de notre système contre toute attaque éventuelle.

Il est important de noter que le système d'identification proposé en mode ensemble fermé, sous clés secrètes correctes, produit un ROR parfait (ROR = 100,00%) avec un RPR de 1. Les résultats obtenus en utilisant la modalité PLM (Fig. III.7.*(b)*) indiquent que les scores d'attaque (obtenus par la clé erronée) sont concentrés dans une plage étroite de sorte que le score le plus élevé (qui est de 0,088) est beaucoup plus petit que le petit score obtenu par une clé correcte (qui est de 0,336) . Par conséquent, le seuil peut être choisi dans cette plage (T_0 = [0.088..0.336]). De même, un système d'identification basé sur PLV (en mode ensemble fermé), Fig. III.7.*(d)*, utilisant la clé erronée, produit de petits scores d'attaque par rapport au plus petit score qui peut être obtenu en utilisant les bonnes clés (correctes) et donc le seuil (T_0) peut être choisi dans la plage [0.280..0.224]. Une analyse sérieuse des résultats précédents montre que dans l'ensemble, la discrimination qui peut être obtenue par des caractéristiques profondes basées sur PCANet a été considérablement améliorée grâce à l'utilisation d'une couche de transformation supplémentaire qui leur permet d'être utilisées dans des applications nécessitant un haut niveau de sécurité.

III.3.3 Analyse de sécurité

L'objectif principal de notre système est de sécuriser les modèles biométriques, dans cette partie nous effectuerons une analyse de sécurité pour tester la robustesse de cette méthode face aux attaques potentielles. En règle générale, pour garantir la sécurité, trois points de la conception du système doivent être vérifiés, à savoir [55] :

i) Même si la structure du système est connue, on ne trouve pas le modèle biométrique ;

ii) L'impossibilité de trouver des modèles biométriques à travers une recherche exhaustive, donc l'espace de clés doit être très grande.

iii) Un petit changement dans la clé sécrète produit des modèles biométriques complètement différents.

Tout d'abord, avant de commencer à analyser la sécurité, il convient de noter que notre système biométrique révocable fonctionne avec deux systèmes chaotiques principaux ($\mathcal{L}_0^c \; et \; \mathcal{L}_1^c$) et $\xi = \Pi_{i=1}^{\ell} L_i$ systèmes chaotiques auxiliaires ($\mathcal{L}_i|_{i=1}^{\xi}$). Les principaux systèmes chaotiques ont pour rôle de faire varier l'état initial des systèmes chaotiques auxiliaires. En

général, les paramètres qui contrôlent la sécurité de notre système sont les états initiaux de $\mathcal{L}_0^c$ et $\mathcal{L}_1^c$ ($x_{0i}|_{i=1}^2$) ainsi que les paramètres de contrôle de $\mathcal{L}_0^c$, $\mathcal{L}_1^c$ et $\mathcal{L}^i|_{i=1}^{\xi}$ (u_0^c, u_1^c, $u_i|_{i=1}^{\xi}$).

III.3.3.1 Analyse de l'espace clé

Dans cette partie, nous allons calculer l'espace de tentatives (en utilisant tous les systèmes chaotiques) qui permet à l'attaquant de récupérer le modèle biométrique. Comme mentionné précédemment, la sous-clé secrète ($\widetilde{K}$) peut être une valeur réelle ou entière codé en hexadécimal sur M bits (dans notre cas M = 16). Nous présentons donc ci-dessous l'analyse de sécurité dans les deux cas.

a) **Clé secrète entière :** dans ce cas, l'espace des tentatives pour chaque paramètre (x_{ij} et u_i) est de 2^M. Donc l'espace de tentative pour les systèmes chaotiques principaux et auxiliaires est :

$$S_p(\mathcal{L}_0^c, \mathcal{L}_1^c) = 2^M \cdot 2^M \cdot 2^M \cdot 2^M = 2^{4M} \quad (3)$$

$$S_a(\mathcal{L}_0^c, \mathcal{L}_1^c) = 2^M \cdot 2^M \cdot 2^M = 2^{\xi M} \quad (4)$$

Alors, l'espace des tentatives globale est :

$$S_g = S_p \cdot S_a = 2^{(4+\varepsilon)M} \quad (5)$$

Dans notre travail, chaque sous-clé secrète est codée sur M = 16 bits et la dernière couche de convolution a ξ = 8 filtres ($\xi = \Pi_{i=1}^{\ell} L_i = L_1 = 8$) pour la modalité PLM, donc l'espace total des clés est :

$$S_g(PLM) = 2^{192} = 6{,}277 \cdot 10^{57} \quad (6)$$

Alors que, pour la modalité PLV, la dernière couche de convolution a ξ = 6 filtres ($\xi = \Pi_{i=1}^{\ell} L_i = L_1 = 6$), donc l'espace total des clés est :

$$S_g(PLV) = 2^{160} = 2{,}462 \cdot 10^{48} \quad (7)$$

Le système annulable développé fonctionne avec cet espace quel que soit le système chaotique utilisé (logistique ou tente). Mais malheureusement, cet espace n'est pas suffisant pour assurer une haute sécurité du système [56], il doit donc être augmenté en codant la sous-clé de sécurité sur M ≫ 16 bits au lieu de 16 bits. Pour un exemple simple, avec la modalité PLM si M = 32, l'espace de clés devient ≥ $3{,}940 \cdot 10^{115}$, ce qui est suffisant pour sécuriser efficacement notre système. De plus, les systèmes biométriques basés sur l'apprentissage profond utilisent généralement un grand nombre des stages et de filtres de convolution, ce qui

augmente considérablement l'espace de clés. Dans le tableau III.2, l'espace de clés pour les différentes architectures possibles est indiqué. Ce tableau montre clairement l'efficacité de notre schéma, en particulier lorsque le nombre de filtres est augmenté. Enfin, il est important de mentionner que le nombre maximum de bits (*M*) utilisés pour coder la sous-clé secrète ($\widetilde{K}_{ij}$) est lié à la sensibilité du système chaotique (voir la partie suivante).

Tableau III.2 : Espace clé pour certaines configurations C-PCANet

	Nombre des stages ($\ell = 1$)		
Nombre de filtres	$L_1 = 8$	$L_1 = 16$	$L_1 = 32$
Longueur de sous-clé (M-bits)	S_g	S_g	S_g
16	$6{,}277 \cdot 10^{57}$	$2{,}136 \cdot 10^{96}$	$2{,}473 \cdot 10^{137}$
24	$4{,}973 \cdot 10^{86}$	$3{,}122 \cdot 10^{144}$	$1{,}230 \cdot 10^{260}$
32	$3{,}940 \cdot 10^{115}$	$4{,}562 \cdot 10^{194}$	$6{,}117 \cdot 10^{346}$

b) **Clé secrète réelle :** c'est le cas où les clés secrètes sont représentées par des valeurs réelles, l'espace de tentative est calculé en utilisant toutes les erreurs absolues moyennes entre deux séquences générées par deux clés secrètes voisins [57].

Soit $S^x, \tilde{S}^x, S^u, \tilde{S}^u$ quatre séquences générées par le même système chaotique dans les conditions suivantes :

$$\begin{cases} S^x = \Gamma_\alpha(x_0, u) = s_i^{x\,L_i^\ell}{}_{i=1} \\ \check{S}^x = \Gamma_\alpha(x_0 + d, u) = \tilde{s}_i^{x\,L_i^\ell}{}_{i=1} \end{cases} \quad (8)$$

$$\begin{cases} S^u = \Gamma_\alpha(x_0, u) = s_i^{u\,L_i^\ell}{}_{i=1} \\ \check{S}^u = \Gamma_\alpha(x_0, u + d) = \tilde{s}_i^{u\,L_i^\ell}{}_{i=1} \end{cases} \quad (9)$$

Où Γ_α désigne le système chaotique (logistique ou tente), x_0 l'état initiale, u le paramètre de contrôle, d une très petite valeur et L^ℓ désigne la longueur de la séquence S^ℓ. L'erreur absolue moyenne $\varepsilon_\ell|_{\ell=\{x,u\}}$ pour le système chaotique est définie comme suit :

$$\varepsilon_\ell(S^\ell, \tilde{S}^\ell) = \frac{1}{L^\ell}\sum_{j=1}^{L^\ell}\left|S^\ell(j) - \tilde{S}^\ell(j)\right| \quad (10)$$

Ainsi, l'espace des clés pour x_0, appelé s_x qui vaut $1/d_x$, où d_x est la valeur de d pour laquelle $\mathcal{E}_\ell = 0$. La même chose pour l'espace des clés de u qui appelé s_u, il est égal à $1/d_u$, où d_u est la valeur de d pour laquelle $\mathcal{E}_\ell = 0$. Comme on a déjà mentionné, notre système

utilise deux systèmes chaotiques principaux ($\mathcal{L}_0^c \; et \; \mathcal{L}_1^c$) et ($\prod_{i=1}^{\varepsilon} L_i$) systèmes chaotiques auxiliaires$\left(\mathcal{L}_i|_{i=1}^{\varepsilon}\right)$, ainsi, l'espace des clés total de chaque groupe devient :

$$\mathcal{S}_p(\mathcal{L}_0^c, \mathcal{L}_1^c) = s_x^{c0} \cdot s_u^{c0} \cdot s_x^{c1} \cdot s_u^{c1} \quad (11)$$

$$\mathcal{S}_a(\mathcal{L}_i|_{i=1}^{\varepsilon}) = \prod_{i=1}^{\varepsilon} s_u^i \quad (12)$$

Donc, l'espace des clés totale est :

$$S_g = \mathcal{S}_p \cdot \mathcal{S}_a = s_x^{c0} \cdot s_u^{c0} \cdot s_x^{c1} \cdot s_u^{c1} \cdot \prod_{i=1}^{\varepsilon} s_u^i \quad (13)$$

Pour tout système logistique, la valeur de s_x est égale à $2.0920 \cdot 10^{16}$ et la valeur de su est égale à $0.5248 \cdot 10^{16}$. Ces valeurs deviennent $3{,}6310 \cdot 10^{16}$ et $0{,}9127 \cdot 10^{16}$ respectivement pour tous les systèmes de tentes. Par conséquent, l'espace des clés total de notre schéma devient :

☑ Pour les systèmes logistiques:

$$\mathcal{S}_p = (2.0920)^2 \cdot 10^{32} \cdot (0.5248)^2 \cdot 10^{32} = 1{,}2053 \cdot 10^{64} \quad (14)$$

$$\mathcal{S}_a = (0.5248)^{\xi} \cdot 10^{16\xi} \quad (15)$$

$$\mathcal{S}_g = 1.2053 \cdot (0.5248)^{\xi} \cdot 10^{16(\xi+4)} \quad (16)$$

☑ Pour les systèmes tentes:

$$\mathcal{S}_p = (3.6310)^2 \cdot 10^{32} \cdot (0.9127)^2 \cdot 10^{32} = 1{,}0983 \cdot 10^{65} \quad (17)$$

$$\mathcal{S}_a = (0.9127)^{\xi} \cdot 10^{16\varepsilon} \quad (18)$$

$$\mathcal{S}_g = 10.9827 \cdot (0.9127)^{\xi} \cdot 10^{16(\xi+4)} \quad (19)$$

Dans notre travail, $\xi = \prod_{i=1}^{1} L_i = L_1 = 8$ pour le système qui utilise la modalité PLM, et par conséquent, l'espace total des clés (S_g) devient $6{,}9350 \cdot 10^{189}$ et $5{,}2885 \cdot 10^{192}$ pour les systèmes chaotique logistique et tente, respectivement. Alors que pour le système qui utilise la modalité PLV, $= \prod_{i=1}^{1} L_i = L_1 = 6$, donc l'espace total des clés (S_g) devient $2{,}5180 \cdot 10^{158}$ et $6{,}3486 \cdot 10^{160}$, pour les systèmes chaotique logistique et tente, respectivement.

Avant de comparer les deux représentations de clés secrètes (codages entiers ou réels), nous avons d'abord trouvé le nombre maximum de bits (M) pouvant être utilisé pour coder la clé dans le premier cas. Dans une clé secrète représentée par un entier, la distance minimale *d* pour les deux systèmes chaotiques est à β^{2M}, donc :

$$\frac{\beta}{2^M} \cong d_\ell \Longrightarrow M \cong log_2\left(\frac{\beta}{d_\ell}\right) = log_2(\beta s_\ell) \quad (20)$$

M est utilisé pour coder la distance minimale dans l'état initiale (d_x) et dans le paramètre de contrôle (d_u), il doit donc être choisi comme suit :

$$M = min\{\lfloor log_2(\beta s_x)\rfloor, \lfloor log_2(\beta s_u)\rfloor\} \quad (21)$$

Dans le système logistiques (où β = 0.43), la valeur de *M* devient :

$$M = \min\{\lfloor 52{,}9981\rfloor, \lfloor 51{,}0031\rfloor\} = \min\{52{,}51\} = 51\ bits \quad (22)$$

Ainsi, l'espace des clés devient égal à :

✍ Pour système qui utilise la modalité PLM :

$$S_g = 2^{(4+\xi)M} = 2^{(4+8)51} = 1{,}6996 \cdot 10^{184} \quad (23)$$

✍ Pour système qui utilise la modalité PLV :

$$S_g = 2^{(4+\xi)M} = 2^{(4+6)51} = 3.3520 \cdot 10^{153} \quad (24)$$

Dans le système tente (où β = 0,58), la valeur de *M* devient :

$$M = \min\{\lfloor 54{,}2253\rfloor, \lfloor 52{,}2332\rfloor\} = \min\{54{,}52\} = 52\ bits \quad (25)$$

Ainsi, l'espace clé devient égal à :

✍ Pour système qui utilise la modalité PLM :

$$S_g = 2^{(4+\xi)M} = 2^{(4+8)52} = 6{,}9617 \cdot 10^{187} \quad (26)$$

✍ Pour système qui utilise la modalité PLV :

$$S_g = 2^{(4+\xi)M} = 2^{(4+6)52} = 3.4323 \cdot 10^{156} \quad (27)$$

En comparant les deux représentations de la clé secrète, il est clair que la représentation par une valeur réelle est plus efficace car elle donne un espace plus grand que celui de la représentation par une valeur entière d'environ 10^5 et 10^4 pour les systèmes logistiques et tente, respectivement, mais heureusement, les deux représentations sont bien meilleures que de nombreuses méthodes dans la littérature.

III.3.3.2 Analyse de sensibilité des clés

Dans cette partie, nous avons examiné le vecteur de caractéristique résultant de plusieurs clés secrètes les plus proches pour tester la sensibilité de notre système à une légère variation de clé. Par conséquent, nous avons utilisé trois clés différentes : une clé correcte (K_c) et deux

clés incorrectes plus proches de la clé correcte par $d_x = 10^{-16}$ ($\widetilde{K}_c^1$) et $d_u = 10^{-16}$ ($\tilde{k}_c^2$). Pour examiner les vecteurs de caractéristiques obtenus, nous avons calculé la corrélation entre ces vecteurs, qui est définie comme suit :

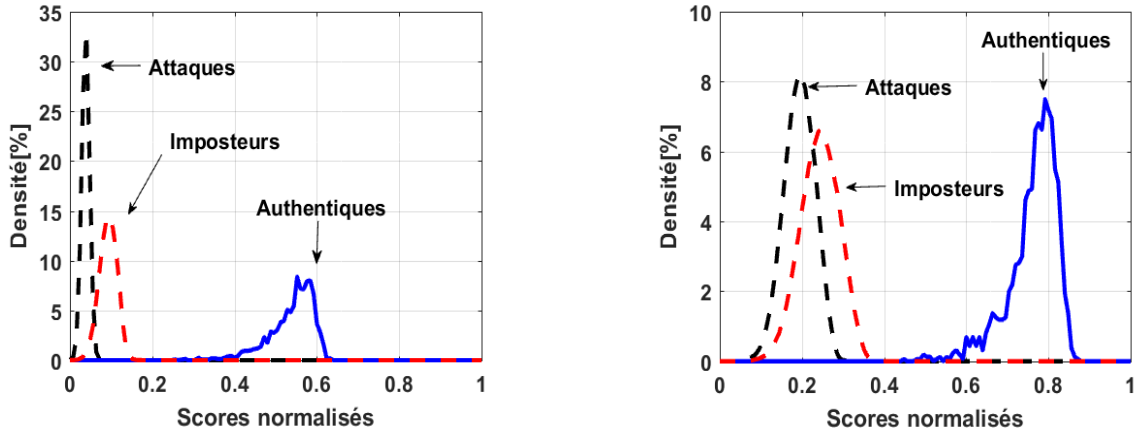
$$\rho_c[\%] = 100 \cdot \frac{c_{ij}}{\sigma_i \sigma_j} \qquad (28)$$

Où C_{ij}est la covariance entre les vecteurs de caractéristiques $\mathcal{T}_i$ et $\mathcal{T}_j$, qui ont des écarts-types de σ_i et σ_j. De plus, pour une comparaison équitable, nous avons sélectionné aléatoirement deux personnes différentes (*i* et *j*) dans la base de données. Après avoir extrait le vecteur de caractéristique de la première personne (*i*) en utilisant la bonne clé, nous avons extrait les vecteurs de caractéristiques des deux personnes (*i* et *j*) en utilisant toutes les clés, puis nous avons calculé la corrélation entre tous les vecteurs de caractéristiques obtenus et les résultats obtenus sont présentés dans le tableau III.3.

Tableau III-3 : Corrélation entre les vecteurs caractéristiques produits par deux personnes

			Personne i			Personne j		
			K_c	$\widetilde{K}_c^1$	$\tilde{k}_c^2$	K_c	$\widetilde{K}_c^1$	$\tilde{k}_c^2$
Personne *i*	K_c	PLM	100.00	6.810	5.310	28.540	7.120	6.910
		PLV	100.00	4.780	3.110	31.701	3.56	2.172

De ce tableau, on peut clairement extraire deux remarques importantes : Premièrement, l'utilisation de la même clé (clé secrète correcte) donne une corrélation totale pour une même personne, cette corrélation devient un peu considérable pour deux personnes différentes du fait de la similitude des traits biométriques de l'empreinte pour les deux personnes, mais bien sûr, le système biométrique capable de différencier ces deux vecteurs de caractéristiques. Deuxièmement, une légère modification d'un paramètre du système chaotique provoque une divergence notable entre les vecteurs de caractéristiques soit pour la même personne, soit pour deux personnes. Aussi, pour voir les performances des systèmes d'identification, qui fonctionnent en mode ouvert, vis-à-vis des attaques (avec une clé très proche de la bonne clé), sur la Fig. III.8.*(a)* et Fig. III.8.*(b)*, nous illustrons les résultats obtenus, sous forme de distributions, pour les deux modalités biométriques (PLM et PLV).



(a) *(b)*

FIG III.8 : Comportement du système d'identification biométrique basé sur C-PCANet (mode d'identification ensemble ouvert) avec deux clés voisines. (*a)* Système biométrique basé sur C-PCANet utilisant la modalité biométrique de PLM, et *(b)* Système biométrique basé sur C-PCANet utilisant la modalité biométrique de PLV.

D'après ces figures, il est clair que tous les scores d'attaque sont complètement décalés en dessous du seuil (en dessous de *zéro FAR*). Il est très remarquable aussi que le point, où le taux d'attaque est nul (ASR = 0), soit beaucoup plus petit que le point *zéro FAR*.

III.4 Conclusion

L'identification biométrique s'est avérée supérieure aux moyens traditionnels d'authentification. Malheureusement, ces systèmes sont vulnérables à une variété d'attaques, dont peut-être la plus grave est l'attaque du modèle biométrique stocké ou transmis, ce qui rend la sécurité de ce modèle plus importante dans la conception des systèmes biométriques. Ce travail suggère donc une méthode d'extraction de caractéristiques efficace qui peut fournir une caractéristique biométrique profonde et révocable. Dans ce contexte, nous testons notre système proposé en utilisant deux modalités biométriques. Les tests dans ce chapitre comprenaient la précision du système biométrique et le niveau de sécurité. Les résultats obtenus sont très excellents et peuvent être encore améliorés si le nombre des stages dans le C-PCANet et le nombre de filtres dans chaque stage augmentent.

Conclusion Générale

Conclusion

Aujourd'hui, la biométrie est largement utilisée dans de nombreuses applications comme moyen de sécuriser l'accès logique/physique. Malgré les résultats impressionnants de ces techniques, mais malheureusement, elles peuvent être exposées à de nombreuses attaques, et donc la vie privée de l'individu est menacée car les traits biométriques sont intrinsèques et ne peuvent être remplacés. Pour surmonter ce grave problème, les données biométriques de la personne doivent être protégées, soit dans la base de données, soit sur le réseau lors de la transmission. L'une des méthodes les plus efficaces consiste à extraire des caractéristiques qui peuvent être révoquées lorsqu'une attaque se produit, et bien sûr, les traits biométriques ne doivent pas être retrouvés à l'aide de ces vecteurs de caractéristiques. Dans ce contexte, nous avons proposé un système d'identification biométrique efficace (*C-PCANet*) capable de produire des caractéristiques révocables. Notre système est basé sur deux technologies prometteuses, à savoir le deep learning, pour extraire des caractéristiques discriminantes, et des systèmes chaotiques pour améliorer la sécurité du système.

Des résultats expérimentaux utilisant une base de données publique et disponible de 400 personnes montrent la robustesse de notre méthode vis-à-vis aux attaques. De plus, un taux d'identification élevé a été obtenu, qui peut également être amélioré en augmentant les nombre des stages de C-PCANet. En effet, notre système peut fonctionner efficacement avec de très grands espaces des clés, surtout si le nombre de stages de convolution et de filtres utilisés augmente. De manière générale, à partir des résultats obtenus, il est clair que notre système peut être utilisé dans des applications qui nécessitent un très haut niveau de sécurité. Les travaux futurs de cette étude se concentreront sur l'utilisation d'autres techniques d'apprentissage profond telles que la *DCTNet* et le *ICANet* et leur utilisation potentielle dans l'Internet des objets (*IoT*) ainsi que dans les applications mobiles basées sur le cloud.

Bibliographies

[1] J. Blasco, T. M. Chen, J. Tapiador, and P. Peris-Lopez."A survey of wearable biometric recognition systems", In: *ACM Comput. Surv.*, 2016, vol. 49, no. 3, p. 43.

[2] C. Jonietz, E. Monari, H. Widak, C. Qu. "Towards mobile and touchless fingerprint verification", In: *Int Conf. on Ad- vanced Video and Signal Based Surveillance (AVSS)*, pp.1 6, Aug 2015.

[3] Hadid, A., Evans, N., Marcel, S., Julian, F. "Biometrics systems under spoofing attack: an evaluation methodology and lessons learned", In: *IEEE Signal Process. Mag.*,2015, 32(5), pp. 2030.

[4] Chen PT, Wu SC, Hsieh JH. "A Cancelable Biometric scheme based on multi-lead ECGs", In: Annual international conference of the engineering in medicine and biology society *(EMBC)*, 2017, vol 39, pp 34973500.

[5] Dang T, Truong Q, Le T, Truong H. "Cancelable fuzzy vault with periodic transformation for biometric template protection", In:*IET* Biometrics, 2016, 5(3), pp. 229235.

[6] Jin Z, Teoh A, Goi B, Tay Y. "Biometric cryptosystems: a new biometric key binding and its implementation for fingerprint minutiae-based representation", *Pattern Recogn*, 2016, 1(56), pp. 5062.

[7] Nandakumar K, Jain A. "Biometric template protection: bridging the performance gap between theory and practice",In: *IEEE Signal Process Mag*, 2015, 32(5), pp. 88100.

[8] Jain, A. K., Flynn, P., & Ross, A. A. (Eds.). (2008). Handbook of Biometrics.

[9] A. Ross, A. Jain, S. Prabhakar. An introduction to biometric recognition. IEEE transactions on circuits and systems for video technology, 2004, Vol. 14, No 1, pp. 4 20.

[10] Y. Lee, K. Lee, H. yung keun Jee, Youn-Hee Gil, Woo-Yong Choi, Dosung Ahn, Sung Bum Pan, "Fusion for Multimodal Biometric Identification", 5th International conference on Audio and video-based biometric person authentication-AVBPA, Hilton Rye Town, N.Y. USA, 2005, pp. 1071-1079.

[11] S.Jidong, L. Xiaoming, "Fusion of Radar and AIS Data", 7th International Conference on Signal Processing-ICSP'04, Beijing, China, 2004, pp, 2604-2607.

[12] N. V. Boulgouris, K. N. Plataniotis and E. Micheli-Tzanakou., \Biometrics : Theory, Methods, and Applications", David B. Fogel, Series Editor, Willy publisher, IEEE Press on Computational Intelligence, 2010.

[13] P. MEENEN, R. ADHAMI, “Fingerprinting for Security”, IEEE potentials, Aout-Septembre 2001, Vol. 20, No. 3, p. 33-38.

[14] K. Nanda kumar, A. Ross, and A. K. Jain, "Biometric Fusion: Does Modeling Correlation Really Matter? ", Proc. 3rd Int`l Conf. on Biometrics: Theory, Applications and Systems, Washington DC, Sept. 2009.

[15] Florent Perronnin and Jean-Luc Dugelay, “Introduction `a la Biométrie’’, Authentification des Individus par Traitement Audio-Vidéo, 2002, vol.19 No.4.

[16] A. Ross and A. K. Jain, "Information Fusion in Biometrics", Pattern Recognition Letters, September, 2003, Vol. 24, Issue 13, pp. 2115-2125.

[17] J. Daughman,” High Confidence Visual Recognition of Persons by a Test of Statistical Independence”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, Vol. 15, p. 1148-1161.

[18] J. Daughman,” High Confidence Visual Recognition of Persons by a Test of Statistical Independence”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, Vol. 15, p. 1148-1161.

[19] S.Jidong, L.Xiaoming, "Fusion of Radar and AIS Data", 7th International Conference on Signal Processing-ICSP'04, Beijing, China, 2004, Vol.3, pp, 2604-2607.

[20] C.Berger ; M.Voltersen ; R.Eckardt ; Eberle, J. ; Heyer, T. ; Salepci, N. ; Hese, S. ; Schmullius,C. ; Tao, J. ; Auer, S. ; Bamler, R. ; Ewald, K. ; Gartley, M. ; Jacobson, J. ; Buswell, A. ; Du, Q. ;Pacifici, F., "Multi-Modal and Multi Temporal Data Fusion", IEEE Journal 55 of Selected Topics in Applied Earth Observations and Remote Sensing, Jun 2013, Vol.6, N.3, pp.1324-1340.

[21] Y.Lee, K.Lee, Hyung keun Jee, Youn-Hee Gil, Woo-Yong Choi, Dosung Ahn, Sung Bum Pan, "Fusion for Multimodal Biometric Identification",5th International conference on Audio and video-based biometric person authentication-AVBPA, Hilton Rye Town, N.Y.USA, July 2005, pp. 1071-1079.

[22] A.ross, A.jain, "human recognition using biometrics: an overview", Ann telecomunication, 2007, 62, (n 1-2), pp.11-35.

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

[24] L. Allano. La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles. Thèse de doctorat, Institut National des Télécommunications, Université d'Evry-Val d'Essonne, 2009.

[25] A.Bouchemha. Etude et Application des transformées géométriques à la Compression des images hautes résolutions et à la Biométrie (Authentification/Vérification de l'empreinte palmaire) (Doctoral dissertation, Université de Annaba) ,2016.

[26] M.el Abed, "évaluation de système biométrique "thèse de doctorat de l'université de caen basse-normandie ,2006.

[27] A- K.Jain, A.Ross and S.Pankanti, "Biometrics: A Tool for information Security", IEEE Trans. On information forensics and security,2006, Vol.1,(2),pp125-143.

[28] Nicolas Morizet, "Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris", Thèse présentée pour obtenir le grade de Docteur, Ecole Nationale Supérieure des Télécommunications, Paris, 18 Mars 2009.

[29] R.W.Belgeuchi, " Sécurité des systèmes biométriques : révocabilité et protection de la vie privée" thèse de Doctorat de l'Ecole Nationale Supérieure d'Informatique, 2015.

[30] Umut Uludag, Sharath Pankanti, Salil Prabhakar, and Anil K Jain. Biometric cryptosystems: issues and challenges. Proceedings of the IEEE, 2004, 92(6):948_960.

[31] Ari Juels and MartinWattenberg. A fuzzy commitment scheme. In Proceedings of the 6th ACM conference on Computer and communications security, 1999, pp 28_36. ACM.

[32] Ari Juels and Madhu Sudan. A fuzzy vault scheme. Designs, Codes and Cryptography, 2006, 38(2):237_257.

[33] https://www.biometrie-online.net/biometrie/biometrie-et-vie-privee.

[34] Andrew Teoh Beng Jin and Lim Meng Hui ,'' Cancelable biometrics' Scholarpedia, 5(1):9201,2010.

[35] Yi Cheng Feng, Pong C Yuen, and Anil K Jain. A hybrid approach for generating secure and discriminating face template. IEEE Transactions on Information Forensics and Security, 2010, 5(1) :103_117.

[36] Thian Song Ong, Andrew Teoh Beng Jin, and David Chek Ling Ngo. Application specific key release scheme from biometrics. IJ Network Security, 2008, 6(2) :127_133.

[37] Florence Jessie MacWilliams and Neil James Alexander Sloane. The theory of error correcting codes, 1977, vol 16. Elsevier.

[38] Savvides, M.; Vijaya Kumar, B.V.K. and Khosla, P.K. Cancelable Biometrics Filters for Face Recognition. Int. Conf. of Pattern Recognition, 2004, vol.3: 922-925.

[39] Hirata, S. and Takahashi, K.Cancelable Biometrics with Perfect Secrecy for Correlation-Based Matching. Lecture Notes in Computer Science 5558: 868-878, 2009.

[40] Jeong, M.Y. et al.Changeable Biometrics for Appearance Based Face Recognition. Biometric Consortium Conference, 2006, Biometrics Symposium : 1-5.

[41] Lee, C. H.; Choi, C.Y. and Toh, K.A. (2007b). Alignment-Free Cancelable Fingerprint Templates Based on Local Minutia Information. *IEEE Transactions on Systems, Man and Cybernetics, Part B,* vol. 37, no. 4: 980-992.

[42] Sergey Tulyakov, Faisal Farooq, Praveer Mansukhani and Venu Govindaraju, "Symmetric hash functions for secure fingerprint biometric systems'', Pattern Recognition Letters 28 (2007) 2427–2436.

[43] N. K. Ratha, J. H. Connel, and R. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," IBM Syst. J, 2001, vol. 40, no. 3, pp. 614–634.

[44] R. M. Bolle, J. H. Connel, and N. K. Ratha, "Biometrics perils and patches," Pattern Recogn , 2002, vol. 35, no. 12, pp. 2727–2738.

[45] N. Ratha, S. Chikkerur, J. Connell, and R. Bolle, "Generating cancelable fingerprint templates," IEEE Trans. Pattern Anal. Mach. Intell, Apr. 2007 vol. 29, no. 4, pp. 561–572.

[46] R. M. Bolle, J. H. Connel, and N. K. Ratha, "Biometrics perils and patches," Pattern Recogn. , 2002, vol. 35, no. 12, pp. 2727–2738.

[47] Vishal M. Patel, Nalini K. Ratha, and Rama Chellappa,'' Cancelable Biometrics'', IEEE Signal processing magazine,2015.

[48] Lu G, Zhang D, Wang K. "Palmprint recognition using eigenpalms features", In: *Pattern Recogn Lett*, 2003, 24(9),pp. 14631467.

[49] Hengjian Li, Lianhai Wang. "Chaos-Based Cancelable Palmprint Authentication System", In: *Procedia Engineering*, 2012, Volume 29, pp. 1239-1245.

[50] Farsana F J, Dr.K.Gopakumar. "Private Key Encryption of Speech Signal Based on Three Dimensional Chaotic Map", In: *International Conference on Communication and Signal Processing*, 2017, pp. 2197-2201.

[51] Nada Hamad, Mizanur Rahman, Saiful Islam. "Novel remote authentication protocol using heart-signals with chaos cryptography", In: *International Conference on Informatics, Health & Technology (ICIHT)*, 2017, Riyadh, Saudi Arabia, pp. 1-7.

[52] Xiaolin Wu, Bin Zhu, Yutong Hu, Yamei Ran. "A Novel Color Image Encryption Scheme Using Rectangular Transform-Enhanced Chaotic Tent Maps", In: *EEE Access*, Vol. 5, pp. 6429-6436.

[53] Kenji Sugimoto, Michael Sebek, Didier Henrion. "Polynomial matrices and recursive QR factorization", In: European Control Conference (ECC), 2001, Porto, Portugal.

[54] Hong Kong Polytechnic University (PolyU) multispectral palmprint database, http://www.comp.polyu.edu.hk/∼biometrics (2011).

[55] Gomez-BarreroM, Rathgeb C. "Protected facial biometric templates based on local gabor patterns and adaptive bloom filters", In: ICPR, 2014, pp. 44834488.

[56] G. Bhatnagar and Q. M. J. Wu. "Chaos-based security solution for fingerprint data during communication and transmission", In: *IEEE Trans. Instrum. Meas.*, 2012, 61(4), pp. 876887.

[57] G. Bhatnagar and Q. M. J. Wu. "Enhancing the transmission security of biometric images using chaotic encryption",In: *Multimedia Syst.*, 2014, 20(2), pp. 203214.

Acronymes

Les termes suivants, classés dans l'ordre alphabétique, sont utilisés dans le texte.

ADN Acide Désoxyribonucléique – *DeoxyriboNucleic Acid.*

ASR Taux de réussite de l'attaque – *Attack Success Rate.*

CMC Courbe des scores cumulées – *Cumulative Match Characteristics.*

C-PCANet Réseau d'analyse en composantes principales révocables – *Cancelable Principal Component Analysis Network.*

EER Taux d'erreurs égaux – *Equal Error Rate.*

FAR Taux de fausse acceptation – *False Acceptance Rate.*

FRR Taux de faux rejet – *False Rejection Rate.*

GAR Taux d'acceptation des clients – *Genuine Acceptance Rate*.

ICA Analyse en composantes indépendantes – *Independent component analysis.*

ICANet Réseau d'analyse en composantes indépendantes – *Independent component analysis Network.*

IoT Internet des Objets – *Internet of Things.*

PCA Analyse en composantes principales – *Principal Components Analysis.*

PCANet Réseau d'analyse en composantes principales – *Principal Components Analysis Network.*

PLM Empreinte palmaire – *Palmprint.*

PLV Empreinte des réseaux des veines – *Palm-Vein.*

ROI Région d'intérêt – *Region Of Interest.*

ROC Fonction d'efficacité du récepteur – *Receiver Operating Characteristic.*

ROR Premier rang de reconnaissance – *Rank-One Recognition.*

RPR Rang de la reconnaissance parfaite – *Rank of Perfect Recognition.*

SVM Machine à vecteurs support – *Support Vector Machine.*

Annexe A

Extraction de la Région d'intérêt

Les images des modalités biométriques obtenues (données brutes) n'étant pas directement exploitables par les systèmes biométriques, elles doivent subir un prétraitement au cours duquel une région d'intérêt (Region of Interest-ROI) est extraite. Le rôle de la phase de prétraitement est de configurer et de modifier l'image (modalité biométrique d'origine) afin de la préparer à l'extraction des caractéristiques.

Un ROI est une région d'intérêt dans une image et peut être utilisé comme point de départ pour de nombreux algorithmes de traitement d'image. Par conséquent, la qualité de l'algorithme utilisé pour détecter les ROIs conditionne souvent la qualité du résultat de toute la chaîne de traitement que l'on souhaite appliquer à une image. En outre, le fait que les mêmes ROI (ou plus ou moins) puissent être détectés sur deux images différentes mais représentant la même scène, est une propriété importante et généralement requise pour tous les algorithmes de détection de ROI. Dans le domaine de la biométrie, la signification des ROIs dépend du type de modalité biométrique. Il peut correspondre aux zones des yeux ou aux zones autour de la bouche dans l'image du visage, tout comme il peut correspondre à l'iris dans l'image de l'œil. La méthode d'extraction dépendra donc de la modalité biométrique.

La phase de prétraitement, pour les deux modalités biométriques, consiste à isoler le ROI (un rectangle au centre de la surface intérieure de la paume) du reste de l'image de la main acquise à partir des capteurs. Comme l'extraction de ce ROI n'est pas forcément idéale, une

tolérance de quelques pixels en translation est introduite dans les deux sens, vertical et horizontal. La méthode d'extraction de ROI appliquée dans notre système est basée sur l'algorithme décrit dans [58].

☑ **Etape 1 : Réduction des bruits :** Le lissage de l'image est une opération importante, utilisée pour atténuer un bruit qui corrompt l'information, avant l'étape de binarisation. En effet, un filtre passe-bas linéaire de type gaussien permet de réduire ce bruit grâce à la localisation du bruit dans les hautes fréquences. Après avoir traversé un filtre passe-bas gaussien de taille 3 × 3 et un écart type $\sigma = 1.5$, on obtient une image lisse (voir figure A.1).

Fig. A.1 : Filtrage de l'image.

☑ **Etape 2 : Binarisation :** Binariser une image revient à segmenter l'image en deux classes : le fond (l'arrière-plan) et l'objet. Cette opération consiste à mettre le fond en noir et l'objet en blanc. Plusieurs techniques de binarisation existent, mais la technique de seuillage est la plus répandue, grâce à sa facilité de mise en œuvre et sa rapidité. Le seuillage final est effectué en comparant l'image filtrée à un seuil. Cette opération est donnée par la formule :

$$I_b(i,j) = \begin{cases} 1, & I_0(i,j) \geq T_p \\ 0, & autrement \end{cases} \qquad (A.1)$$

avec I_0 est l'image originale (après le filtrage de gaussienne) et I_b est l'image binaire obtenue. Pour générer le seuil T_p, nous avons opté pour la méthode de *Otsu* en raison de son efficacité. La figure A.2 montre l'opération de binarisation.

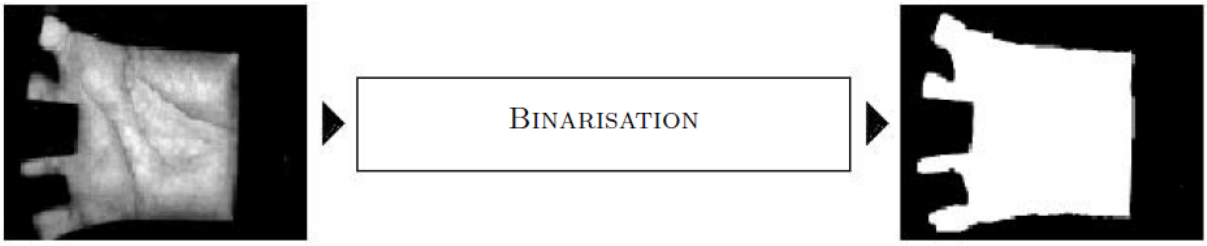


Fig. A.2 : Binarisation de l'image.

☑ **Etape 3 : Détection de contour :**

L'objectif de cette étape, réalisée à l'aide d'un algorithme classique de suivi de contour (Square-Tracing), est de déterminer le contour de la main. Sur les images binaires, les pixels sont soit noirs, soit blancs. Afin d'identifier les objets dans une image binaire, nous devons localiser les pixels blancs qui sont connectés les uns aux autres. En d'autres termes, les pixels connectés, ou voisins, forment un objet sur une image binaire qui doit être identifié avec succès. De plus, dans un pavage carré, le pixel est en contact avec 8 pixels. Ces pixels définissent leur voisinage. Ce type de connectivité est appelé connectivité de type 8 (voir Fig. A.3. *(a)*). Dans le cas où les pixels sont liés par seulement 4 pixels, la connectivité est de type 4 (voir Fig. A.3. *(b)*).

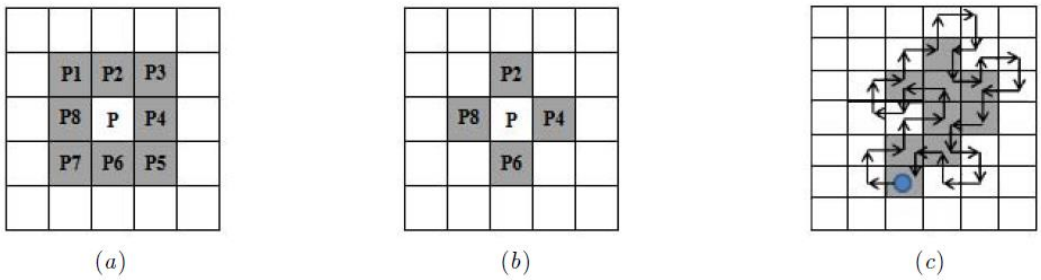


Fig A.3 : Contour dans une image binaire. *(a)* connectivité de type 8, *(b)* connectivité de type 4 et *(c)* application de l'algorithme.

On considère dans une image que la main correspond à un ensemble de pixels blancs sur fond de pixels noirs. En d'autres termes, les pixels qui nous intéressent sont les blancs (pixels utiles) et les pixels noirs représentent l'arrière-plan (le fond). On commence en bas à gauche de l'image et on balaie les colonnes de bas en haut, jusqu'à rencontrer un pixel blanc (pixel qui nous intéresse). Ce pixel est stocké comme pixel de départ. Une fois le pixel de départ détecté, si nous sommes sur un pixel utile, alors nous nous tournons vers la gauche et si nous sommes sur un pixel d'arrière-plan, nous nous tournons vers la droite. L'algorithme s'arrête une fois que nous retombons à nouveau sur le pixel de départ. A chaque fois qu'un pixel blanc est détecté, les coordonnées de ce dernier sont mémorisées afin de connaître le contour de la main. Le fonctionnement de cet algorithme est illustré à la Fig. A.3. *(c)*. La Fig. A.4 montre le contour de la main après l'application de cet algorithme.

☑ **Etape 4 : Détection des points d'intérêts :** Les points caractéristiques, représentés par le bout des doigts et les vallées entre les doigts, sont calculés avec le même algorithme de suivi de contour. Ces points coïncident avec les maxima et minima de l'abscisse des points du contour. De plus, le bout des doigts est initialisé aux points des abscisses les plus faibles de chaque doigt, et les vallées aux points des plus grandes abscisses entre les doigts.

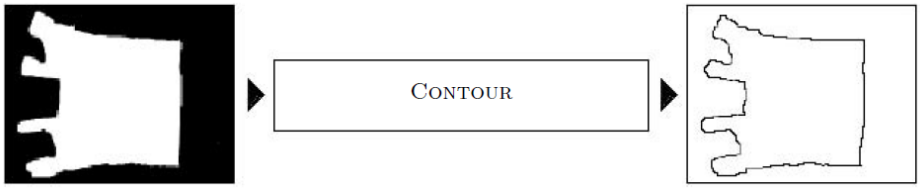


Fig A.4 : Contour de la main.

La fenêtre de paume (ROI) est construite selon l'un des modèles classiques (généralement carré). Son emplacement ne dépend que de la position des vallées F_1 et F_2, ainsi que de la largeur de la paume. La Fig. A.5 représente le résultat de l'extraction des deux points (F_1 et F_2).

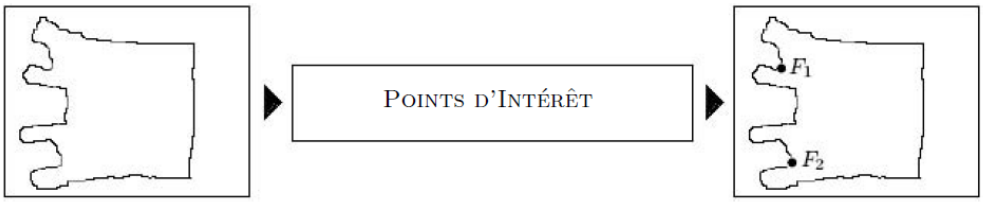


Fig A.5 : Localisation des points d'intérêt.

☑ **Etape 5 : Rotation de l'image :** Avec l'extraction de ces points, il ne reste plus qu'une étape à normaliser la paume : sa rotation. L'angle de rotation a été calculé en fonction de la ligne tracée entre les deux points F1 et F2 et l'axe des ordonnées. Une fois l'angle déterminé, nous faisons pivoter l'image d'un angle θ, afin d'aligner les axes F_1 et F_2 avec l'axe y de l'image (voir Fig. A.6).

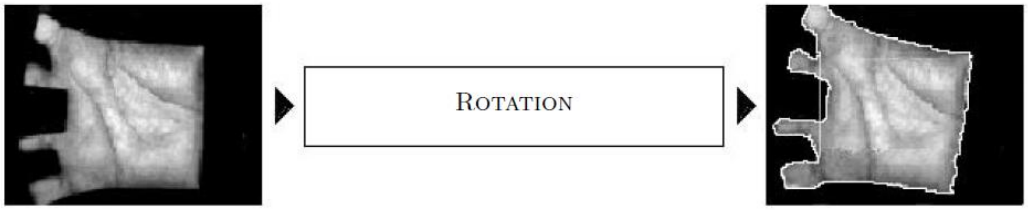


Fig A.6 : Rotation de l'image.

Dans beaucoup d'études, la rotation a lieu avant la définition de la fenêtre, et cela afin de repérer plus facilement le carre. En effet, si la rotation est effectuée sur l'image entière, les bords de la fenêtre d'extraction sont horizontaux et verticaux, le carre (correspond au ROI) est donc très facile a localiser.

☑ **Etape 6 : Localisation de ROI :** La largeur de la zone d'étude (ROI), correspond à la distance d entre le segment de référence $\overline{F1F2}$ et le carré de la paume, est fixé à quelques pixels (dans notre travail, d = 20 pixels). La largeur du carré, W, est aussi fixée à quelques

pixels (dans notre travail, W = 128 pixels). Ces deux distances sont fixées de façon à ce que la ROI soit centré sur la main. Cette région (ROI) est mise en évidence sur le schéma suivant (voir Fig. A.7).

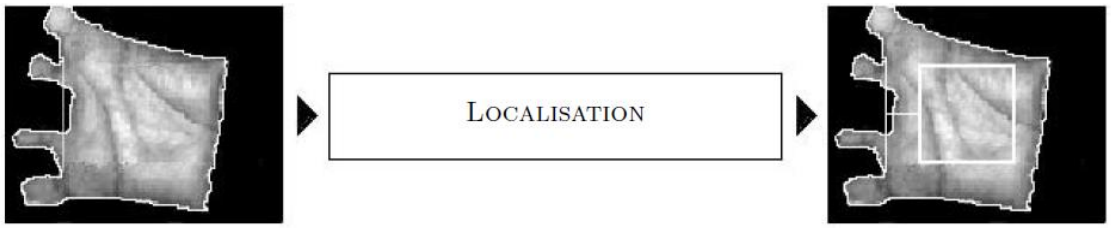


Fig A.7 : Localisation du ROI dans la paume.

☑ Etape 7 : Extraction du ROI

Une région carrée, qui correspond au ROI, a une dimension fixe (128 × 128 pixels), de sorte que toutes les régions soient conformes à une même dimension, sont ensuite extraites. La Fig. A.8 montre le résultat obtenu après l'opération d'extraction de ROI.

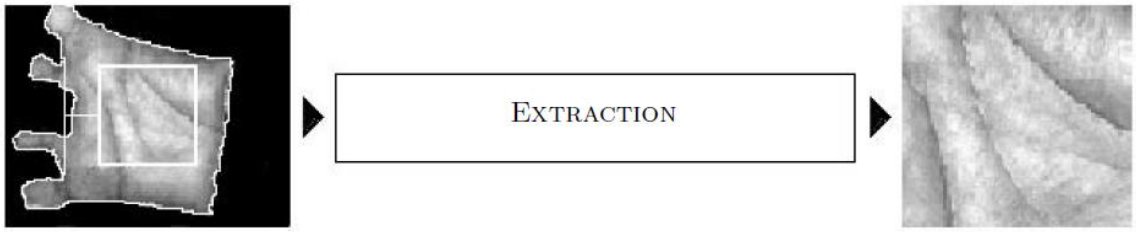


Fig A.8 : Extraction de ROI à partir de la paume.

Résumé : Ces dernières années, la disponibilité de dispositifs d'acquisition de données biométriques à faible coût et le développement impressionnant de la technologie numérique, en plus de la haute sécurité requise pour les données sensibles, ont conduit à une croissance significative de l'utilisation des technologies biométriques pour la reconnaissance automatique de l'identité humaine. Malheureusement, les données biométriques d'un individu sont très sensibles du fait de leur association permanente avec l'utilisateur, ce qui justifie l'inquiétude croissante concernant la confidentialité et l'anonymat des individus face à toute tentative de piratage. Par conséquent, de nombreuses études se sont concentrées sur la recherche de moyens d'extraire des caractéristiques biométriques qui peuvent être révoquées et remplacées à tout moment lorsqu'elles sont compromises. Dans ce mémoire, une nouvelle méthode d'extraction de caractéristiques profondes et révocables (C-PCANet) utilisant des cartes chaotiques est proposée. Notre système peut fournir efficacement des caractéristiques biométriques profondes légères et révocables qui peuvent être utilisées dans de nombreuses applications de haute sécurité.

.***Mots clés :*** Biométrique révocable, Caractéristique profonde, PCANet, Cartes chaotiques, Empreinte palmaire, Empreinte du réseau veineux de la paume.

Abstract: In recent years, the availability of low-cost biometric data acquisition devices and the impressive development of digital technology, in addition to the high security required for sensitive data, have led to significant growth in the use of biometric technologies for automatic recognition of human identity. Unfortunately, an individual's biometric data are very sensitive due to their permanent association with the user, which justifies the growing concern regarding the confidentiality and anonymity of individuals in the face of any hacking attempt. Therefore, many studies have focused on finding ways to extract biometric features that can be canceled and replaced at any time when they are compromised. In this paper, a new cancelable deep feature extraction method (C-PCANet) using chaotic maps is proposed. Our scheme can effectively provide lightweight and cancelable deep biometric features that can be used in many high-security applications.

Keywords: Cancelable Biometric, Deep feature, PCANet, Chaotic, Palmprint, Palm-vein.

ملخـــص : في السنوات الأخيرة ، أدى توافر أجهزة الحصول على البيانات البيومترية منخفضة التكلفة والتطور المثير للإعجاب للتكنولوجيا الرقمية ، بالإضافة إلى الأمان العالي المطلوب للبيانات الحساسة ، إلى نمو كبير في استخدام التقنيات البيومترية للاعتراف التلقائي بهوية الافراد. . لسوء الحظ ، فإن البيانات البيومترية للفرد حساسة للغاية نظرًا لارتباطها الدائم بالمستخدم ، مما يبرر القلق المتزايد بشأن الخصوصية وعدم الكشف عن هوية الأفراد في مواجهة أي محاولة اختراق. لذلك، ركزت العديد من الدراسات على إيجاد طرق لاستخراج الخصائص البيومترية التي يمكن إلغاؤها واستبدالها في أي وقت عند اختراقها. في هذه الورقة ، تم اقتراح طريقة جديدة لاستخراج الميزات العميقة القابلة للإلغاء(C-PCANet) باستخدام الخرائط الفوضوية. يمكن لنظامنا توفير وظائف البيومترية عميقة و خفيفة وقابلة للإلغاء بكفاءة والتي يمكن استخدامها في العديد من التطبيقات التي تتطلب مستو ى عالي من الأمان..

الكلمـات المفتاحيـة : السمات الحيوية القابلة للإلغاء، ميزة عميقة، خرائط الفوضى، بصمة الكف، بصمة اوردة الكف .