République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Larbi Tébessi –Tébessa-

Faculté des Science Exactes et des Sciences de la Nature et de la Vie

Département : Mathématiques et Informatique

MEMOIRE DE MASTER

Domaine: Mathématiques et informatique

Filière: Informatique

Option: Système et Multimédia

Thème :

Tatouage d'images avec des données biométriques pour la preuve de propriété

Présenté par :

BELALOUI Meriem

DJAFFAL Souhaila

Devant le jury

C. Djeddi	MCB	Université de Tébessa	Président
T. Nouiwa	MAA	Université de Tébessa	Examinateur
L. Laimeche	MAA	Université de Tébessa	Encadreur
A. Meraoumia	MCA	Université de Tébessa	Co-Encadreur

Date de soutenance : **XX**/05/2017

Note: Mention:

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Larbi Tébessi –Tébessa-

Faculté des Science Exactes et des Sciences de la Nature et de la Vie

Département : Mathématiques et Informatique

MEMOIRE DE MASTER

Domaine: Mathématiques et informatique

Filière: Informatique

Option: Système et Multimédia

Thème :

Tatouage d'images avec des données biométriques pour la preuve de propriété

Présenté par :

BELALOUI Meriem

DJAFFAL Souhaila

Devant le jury

C. Djeddi	MCB	Université de Tébessa	Président
T. Nouiwa	MAA	Université de Tébessa	Examinateur
L. Laimeche	MAA	Université de Tébessa	Encadreur
A. Meraoumia	MCA	Université de Tébessa	Co-Encadreur

Date de soutenance : **XX**/05/2017

Note: Mention:

REMERCIMENT

Nous remercions en premier ALLAH le tout puissant de nous avoir accordé La volonté et le courage

pour réaliser ce mémoire

Nous souhaitant adresser nos remerciements à notre encadreur de mémoire ***M. Laimech Lakhdar****, qui est toujours disponible tout au long de*

la réalisation de ce mémoire, ainsi

pour ces orientations, ses conseils et ses remarques judicieuses l'inspiration, l'aide nous

Nous adressons nos vifs remerciements à

M. Abdallah MERAOUMIA*,*

pour nous avoir diligentés

tout au long de ce travail, pour son aide, sa

patience, sa compétence, et ses remarques.

Nous remercions à tous ceux qui ont Contribué de près ou de loin à la réalisation de Ce travail.

Dédicace

Je dédie ce modeste travail à celle qui m'a donné la force, le symbole de tendresse, qui s'est Sacrifié pour ma réussite à ma mère.

A mon père, pour m'avoir toujours encouragéà aller le plus loin possible dans mes études,

à me donner l'aide et me protéger.

A mon chèr frère qui est toujours derrière moi et m'encourager.

A ma chère sœur qui me soutenu toujours pour aller avant.

A mes amies, qui nous avons passés des bons moments ensemble.

Je vous souhaite un avenir plein de joie et de bonheur

BELALOUI Meriem

Dédicace

Je dédie ce modeste travail à mes chers parents qui n'ont pas cessé de m'encourager durant toutes mes études et qui m'accompagnés dans toute ma vie, s'inquiétant énormément pour m'offrir une meilleure vie, sans oublier leur participation dans ce travail,

Que Dieu les protège

A mes chers frères qui sont toujours présents pendant la réalisation de ce projet.

A mes chères sœurs qui sont toujours avec moi.

A mes chers amis.

Tous nos collègues surtout les étudiants du master 2 SYSTEMS ET MULTIMEDIA promotion 2017.

Enfin, je dédie ce travail à tous ceux qui me connaissent de près ou de loin.

DJAFFAL Souhaila

Résumé

Le tatouage d'image a connu ces dernières années, un essor spectaculaire. L'utilisation accrue des applications multimédia pose de plus en plus des problèmes concernant la préservation des droits d'auteurs. Récemment, plusieurs travaux ont été proposés utilisant des données biométriques vu de leurs variabilités et de leurs caractéristiques.

Dans notre travail, nous proposons une méthode de tatouage biométrique pour la protection des droits d'auteur basée sur l'utilisation de deux modalités biométriques : l'empreinte palmaire (PLM) et l'empreinte de réseau veineux de la paume (PLV). La construction de la marque à insérer dans notre travail est basé l'utilisation d'une part la méthode de BioHashing basique et d'autre part une méthode de BioHashing proposée basée sur un dictionnaire d'exemplaires. L'objectif de notre travail est de remédier le problème lorsqu'on utilise la donnée biométrique seul lors de la vérification.

Mots clés : *Droits d'auteur, Tatouage des images, Biométries, Multimodalité, Empreinte palmaire, empreinte de réseau veineux de la paume.*

Abstract

Image watermarking has in recent years a spectacular development. The increased use of multimedia applications pose increasing problems concerning the preservation of copyrights. Recently, several works have been proposed using biometrics data because of their variability and characteristics.

In our work, we propose an image watermarking method for copyright protection with biometric data using two biometrics: Palm Print (PLM) and the venous network footprint of the palm (PLV). The construction of the mark to be inserted is based on the use of a part from the standard BioHashing method and on the other part a proposed BioHashing Method based on a dictionary of copies. The goal of our work is to remedy the problem when using the biometric data alone during the verification.

Keywords : *Copyright, Image watermarking, Biometrics, Multimodality, Palmar footprint, venous network footprint.*

ملخص

عرفت تقنية وشم الصور في السنوات الأخيرة تتطور مذهل. الاستعمال المفرط لتطبيقات الوسائط المتعددة تطرح أكثر فأكثر مشاكل بخصوص الحفاظ على حقوق التأليف والنشر. مؤخرا، الكثير من الدارسات المقترحة استعملت المعطيات البيومترية وفقا لمتغيراتها وخصائصها.

في عملنا هذا، اقترحنا طريقة للوشم البيومتري بهدف حماية حقوق النشر تعتمد على استعمال ترتيبتين بيومترية: بصمة راحة اليد وبصمة الشبكة الوريدية من النخيل. انشاء العلامة المدخلة في عملنا هذا تعتمد على استعمال من ناحية طريقة البيوهاشينق الأساسية ومن ناحية أخرى طريقة البيوهاشينق المقترحة التي تعتمد على قاموس النماذج.

الهدف من عملنا هذا،هو علاج المشكل عند استخدام فقط المعطيات البيومترية خلال التحقق.

الكلمات المفتاحية: حقوق النشر، وشم الصور، بيومتري، تعدد الترتيبات، بصمة راحة اليد، بصمة الشبكة الوريدية من النخيل.

REMERCIMENT

Nous remercions en premier ALLAH le tout puissant de nous avoir accordé La volonté et le courage

pour réaliser ce mémoire

Nous souhaitant adresser nos remerciements à notre encadreur de mémoire ***M. Laimech Lakhdar****, qui est toujours disponible tout au long de*

la réalisation de ce mémoire, ainsi

pour ces orientations, ses conseils et ses remarques judicieuses l'inspiration, l'aide nous

Nous adressons nos vifs remerciements à

M. Abdallah MERAOUMIA*,*

pour nous avoir diligentés

tout au long de ce travail, pour son aide, sa

patience, sa compétence, et ses remarques.

Nous remercions à tous ceux qui ont Contribué de près ou de loin à la réalisation de Ce travail.

Dédicace

Je dédie ce modeste travail à celle qui m'a donné la force, le symbole de tendresse, qui s'est Sacrifié pour ma réussite à ma mère.

A mon père, pour m'avoir toujours
Encouragé à aller le plus loin possible dans mes études,
à me donner l'aide et me protéger.

A mon chèr frère qui est toujours derrière moi et m'encourager.

A ma chère sœur qui me soutenu
toujours pour aller avant.

A mes amies, qui nous avons passés des bons moments ensemble.
Je vous souhaite un avenir
plein de joie et de bonheur

BELALOUI Meriem

Dédicace

Je dédie ce modeste travail à mes chers parents qui n'ont pas cessé de m'encourager durant toutes mes études et qui m'accompagnés dans toute ma vie, s'inquiétant énormément pour m'offrir une meilleure vie, sans oublier leur participation dans ce travail,

Que Dieu les protège

A mes chers frères qui sont toujours présents pendant la réalisation de ce projet.

A mes chères sœurs qui sont toujours avec moi.

A mes chers amis.

Tous nos collègues surtout les étudiants du master 2 SYSTEMS ET MULTIMEDIA promotion 2017.

Enfin, je dédie ce travail à tous ceux qui me connaissent de près ou de loin.

DJAFFAL Souhaila

Chapitre II La Biométrie

Chapitre III Résultats Expérimentaux

Liste des tableaux

Liste des figures

Abréviations

ACP : Analyse des Composantes Principales

ADN : Anime Digital Network

DCT : Transformée Discrète en Cosinus

DET : Detection Error Tradeoff

DWT : Discret Wavelet Transform

ERR : Equal Error Rate

FAR : False Reject Rate

FRR : False Accept Rate

JPEG : Joint Photographic Experts Group

LSB : Least Significant bit

MSE : Mean Square Error

MAE : Mean Average Error

NIST : National Institute of Standards and Technologies

PIN : Personal Identification Number

PSNR : peak signal to noise ratio

ROC : Receiver Operating Characteristic

TIFF : Tagged Image File Format

Introduction générale

Avec le développement rapide des moyens de communication, des moyens de sauvegarde, des techniques de partage et de copie, le piratage des documents numérique est devenu très simple et très facile à faire. D'autre part, le piratage des documents numérique peut avoir une conséquence économique non négligeable : artistes, chanteurs et producteurs de cinéma, se plaignent régulièrement du piratage de leurs œuvres sur le marché parallèle, réduisant leurs droits d'auteurs à leur plus simple expression. En effet, ce phénomène massif induit une forte perte des chiffres d'faires et une destruction nette de milliers d'emplois.

Les recherches sur la protection des droits d'auteur ont principalement débuté vers 1993, et aujourd'hui plus d'une centaine d'article sont annuellement consacrés à ce sujet. Les premières techniques, comme la cryptographie reste insuffisante ou d'un emploi difficile. En effet, les dispositifs de cryptographie protègent un document numérique lors d'une transmission, mais pas au-delà. Une technique complémentaire a alors été développée : le tatouage numérique, dérivé de la dissimulation d'information.

Le tatouage numérique consiste à insérer une information invisible (dans certain cas visible) appelée aussi marque ou signature dans une image ou d'autres documents numériques, pour divers buts tel que la protection des droits d'auteur, la vérification de l'intégrité, et a lutte contre les copies illégale.

Dans la littérature, plusieurs travaux ont été proposés afin d'assurer la protection des droits d'auteur. Ces travaux sont basée, généralement, sur deux étapes importante : algorithmes d'insertion/ d'extraction en fonction de l'application envisagée et le choix de la marque à insérer. Le choix de la marque à insérer est toujours limité dans ces travaux au nom d'une compagnie, un logo ou même à une information spécialisée au propriétaire du document

numérique. Ce type des marques sont des informations moins importantes, moins liées au propriétaire, faciles à falsifier et à reproduire.

Récemment, plusieurs travaux ont été proposés utilisant des données biométriques vu leurs variabilités caractéristiques. Malgré que ces travaux qui ont été appliquée à diverses modalités biométriques (visage, empreinte digitale, empreinte palmaire) remédient le problème de piratage de la marque utilisée, ils restent toujours moins performants lorsqu'on utilise seulement les données biométriques ; en d'autre terme, lorsqu'un imposter B vole la clé secrète de client A et essaie de s'authentifier comme A. Lorsque ce problème se produit, la performance de ces méthodes peut être inférieure à celles obtenues en utilisant uniquement les données biométriques.

Dans notre travail, nous avons proposés une méthode de tatouage biométrique pour la protection des droits d'auteur basée sur l'utilisation de deux modalités biométriques : l'empreinte palmaire (PLM) et l'empreinte palmaire des veines (PLV) afin de construire la marque à insérer.

La méthode de tatouage proposée se fait en deux étapes principales : d'une part l'extraction des vecteurs de caractéristique utilisant deux algorithmes : LBP (Local Binary Pattern) et l'algorithme HOG (Histogram Oriented Gradient) et d'autre part l'utilisation la méthode BioHashing basique et une nouvelle méthode de BioHashing basée sur le code book.

Hormis l'introduction générale et la conclusion-perspective, ce document est composé de chapitre chapitres qui se présentent comme suit:

- Dans le premier chapitre, nous présentons la nécessiter de la protection des droits d'auteur ainsi les différentes techniques de sécurité de l'information existantes. Ensuite, nous présentons le tatouage numérique sa position par rapport à la stéganographie, filigrane et la cryptographie. Puis nous énumérons ses propriétés, ses contraintes, ses applications et les domaines d'insertions.

- Le deuxième chapitre présente les notions de base de la biométrie, les différentes techniques biométriques et leurs applications. Ensuite, nous présentons l'architecture d'un système biométrique ainsi les différentes phases de son fonctionnement. Puis, nous présentons la biométrique multimodal par l'introduction des différents types de fusion appliquée à la biométrie, les différents niveaux de fusion ainsi la notion de

normalisation des scores. Finalement, quelques techniques de protection des droits d'auteur exploitant des données biométrique sont présentées.

- Dans le dernier chapitre, nous décrivons les deux modalités (PLP et PLV) utiliser dans le système proposé, ainsi que le processus général de notre travail ; une petite description des algorithmes d'extraction des caractéristiques (HOG et LBP) et le principe d'algorithme BioHashing. Ensuite, les résultats expérimentaux obtenus par chaque méthode en analysent leurs performances séparément, suivies d'une discussion et comparaison des résultats.

- Le dernier chapitre présente la méthode de BioHashing basique et la méthode de BioHashing proposée afin de dissimuler les vecteurs de caractéristiques obtenues par deux méthodes d'extraction : LBP et l'algorithme HOG. Ensuite, une méthode de tatouage conventionnelle basée sur la transformée en cosinus discrète est utilisée pour l'insertion et l'extraction de la marque générer. Finalement, des bases de données types sont utilisées pour évaluer les performances de la méthode proposée en particulier dans le cas de l'utilisation de la donnée biométrique seule.

Enfin, une conclusion générale avec les perspectives visées que nous envisagerons sont données à la fin de ce mémoire.

Sécurité d'information et la protection des droits d'auteur

I.1. Introduction

De nos jours, le développement des réseaux de communication est des supports numériques entraine une diffusion massive de document stockés à l'aide de formats numériques. Ces techniques, qui permettent d'emmagasiner une grande quantité d'information en peu de place, facilitent aussi l'utilisation illégale des documents, il est en effet extrêmement aisé de récupérer un document sur Internet et de copier, modifier et même de diffuser. Ces manipulations, si elles débouchent sur la commercialisation des copies ou sur toute utilisation autre que privée, sont illégales tant que les droits d'auteur n'ont pas été versés l'ayant droit du document. Dans ces conditions, il devient donc nécessaire de mettre en œuvre des systèmes permettant de faire respecter les droits d'auteur, de contrôler les copies et de protéger l'intégrité des documents. Dans ce contexte, le tatouage numérique est très rapidement apparu comme la solution pour renforcer la sécurité des documents multimédia.

Dans ce chapitre, nous allons présenter d'abord la nécessiter de la protection des droits d'auteur ainsi les différentes techniques de sécurité de l'information existantes. Ensuite, nous présenterons le tatouage numérique et sa position par rapport à la stéganographie, filigrane et la cryptographie. Puis nous énumérons ses propriétés, ses contraintes, ses applications et les domaines d'insertions.

I.2 Nécessité de la protection des droits d'auteur

Les documents numériques quel qu'ils soient sont soumis au problème de piratage. En effet, avec le développement rapide des moyens de communication, les moyens de sauvegarde, les techniques de partage et de copie, la procédure de piratage est devenue très simple et très facile à faire. Le piratage peut avoir une conséquence économique non négligeable, les artistes, chanteurs et producteurs de cinéma, se plaignent régulièrement du piratage de leurs œuvres sur le marché parallèle, réduisant leurs droits d'auteurs à leur plus simple expression. En effet, ce phénomène massif induit une forte perte des chiffres d'faires et une destruction nette de milliers d'emplois [1].

La recherche sur la protection des œuvres a principalement débuté vers 1993, et aujourd'hui plus d'une centaine d'article sont annuellement consacrés à ce sujet. Les premières techniques, comme la cryptographie reste insuffisante ou d'un emploi difficile. En effet, les dispositifs de cryptographie protègent un document numérique lors d'une transmission, mais pas au-delà. Une technique complémentaire a alors été envisagée : le tatouage numérique, dérivé de la dissimulation d'information.

I.3 Tatouage numérique : définitions et objectifs

- ***Définition 1***

Le tatouage numérique consiste à insérer une marque invisible (dans certain cas visible) appelée aussi signature, ou tatouage, dans une image ou d'autres documents numériques, pour divers buts tel que la lutte contre la fraude, le piratage informatique et la protection des droits d'auteur. La marque insérée est essentiellement une séquence aléatoire, un logo binaire ou une image de niveaux de gris : elle doit être connue uniquement par le propriétaire ou par le diffuseur. [2]

- ***Définition 2***

Le tatouage numérique est l'art d'enfouir un message binaire dans un signal représentant un contenu de manière imperceptible et robuste. La robustesse signifie qu'il est possible de détecter le tatouage même si le contenu a subi des transformations (filtrage, ajout de bruit).

Objectifs de tatouage numérique

Le tatouage numérique a comme objectif de cacher des messages en insérant des marques à des fins commerciales. Elle permet de prévenir les contournements des droits d'auteurs. Pour y

arriver, un système de tatouage est inséré dans le fichier. Cela permet de limiter les copies et les contrefaçons sur le fichier de base.

I.4 Lien de tatouage numérique avec d'autres technologies de sécurité

Le tatouage fait partie de la science de la dissimulation d'information. Cette science est en fait l'ensemble des moyens permettant de protéger tout document en assurant sa confidentialité, son intégrité et son authenticité. On distingue deux sous-classes dans la dissimulation d'information : la stéganographie et le filigrane.

Dans la suite, nous décrivons brièvement chacune de ses disciplines en mettant en évidence les différences qui existent entre elles et la cryptographie.

I.4.1 Stéganographie

La stéganographie étudie les techniques pour permettre à des partenaires de communiquer de façon caché en établissant un véritable protocole de communication secrète au-dessus d'autres protocoles anodins, c'est ce qu'on appelle canal de communication secrète (*cover Channel*), le mot caché signifie que la présence de l'information n'est pas perceptible parce qu'elle vie dans un support d'un caractère anodin qui peut être de type image, vidéo, audio, ou un texte. Le message dissimulé n'a aucun lien avec le support chargé de transport [3].

I.4.2 Filigrane

Le filigrane a pour but de limiter le nombre de copies. L'application de son système détecte les copies illégales du document original. Lorsqu'une de copie de celui-ci est réalisée, une empreinte (que l'on qualifie d'identifiant) y est inscrite. Si une copie illégale est réalisée, il possible de retrouver la source grâce à l'identifiant inscrit dans l'empreinte. Ainsi, on ne s'oriente pas comme avec le tatouage numérique sur la source du document mais sur le destinataire. De la sorte, chaque copie contient une information propre à l'utilisateur, rendant le document unique [3].

I.4.3 Cryptographie

La cryptographie est le domaine le plus proche des techniques de dissimulation d'information et est sujet à de nombreuses confusions. Son but premier est de chiffrer l'information et de la rendre illisible mais non de la cacher. Elle permet également d'échanger des données entre des correspondants sans que les personnes non-autorisées en prennent connaissance [3].

I.5 Principes des schémas de tatouage

Les schémas de tatouage s'appuient tous sur un même principe qui se traduit par deux phases importantes qui sont la phase d'insertion de la marque et la phase de détection (voir figure I.1 et figure I.2) [4]. Ces deux phases s'appliquent souvent et en général sur un seul espace choisi selon le contexte et l'objectif visé par le schéma en question.

I.5.1 Phase d'insertion

La phase d'insertion présentée dans la figure I.1 comprend les étapes suivantes :

1. Compression et chiffrement de la marque à insérer (étape optionnelle),
2. Sélection d'un support porteur de la marque,
3. Utilisation d'un algorithme d'insertion dont l'objectif est la sélection des sous parties favorable à l'insertion dans le support.
4. Insertion de la marque à l'aide d'une clé secrète.

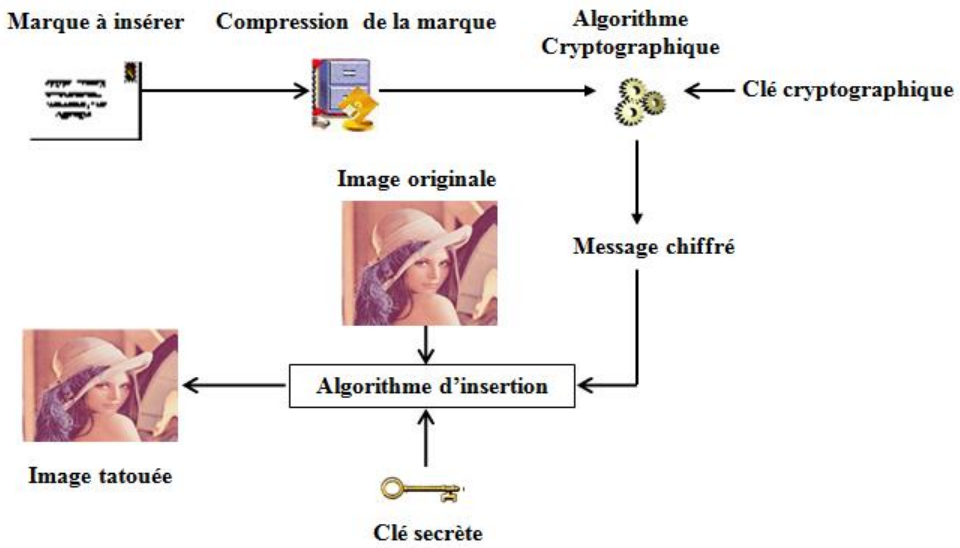


Figure I.1 : Phase d'insertion

I.5.2 Phase d'extraction

La phase d'extraction présentée dans la figure I.2 comprend les étapes suivantes :

1. Utilisation d'un algorithme d'extraction dont l'objectif est la sélection des sous parties contenant la marque dans le support.

2. Retrouver les positions de la marque chiffré dans les parties favorable à l'aide de la clé secrète utilisée.
3. Déchiffré le message à l'aide de la clé cryptographique puis le décompresser.

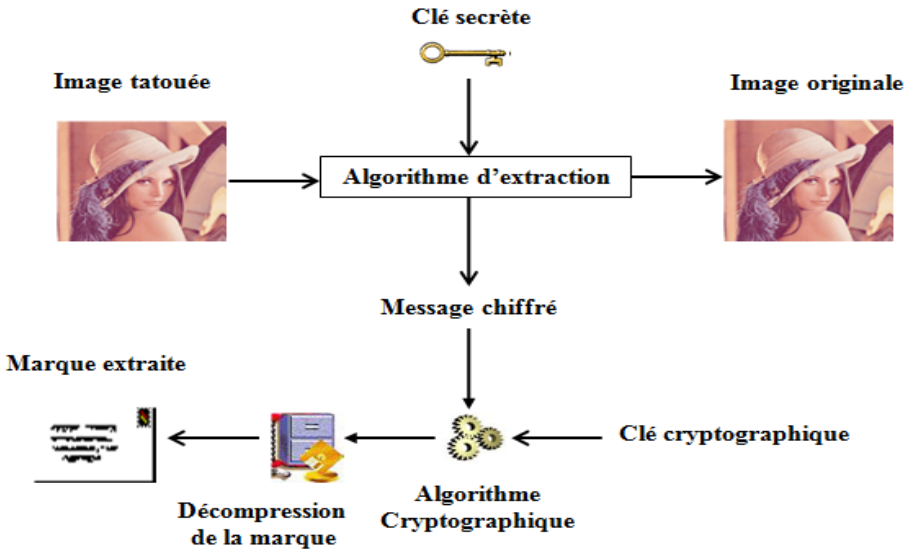


Figure I.2 : Phase d'extraction

I.6 Contraintes du tatouage numérique

Trois paramètres principaux habituellement utilisés afin de quantifier la performance d'une technique de tatouage : la capacité, l'imperceptibilité et la robustesse [5].

I.6.1 La capacité

C'est la quantité d'information que l'on désire cachée par rapport à la quantité d'information associée au support numérique (image audio, vidéo). Dans le tatouage la capacité est variable selon l'application, par exemple, la capacité se limite souvent de 16 à 64 bits pour assurer un service de droit d'auteurs à l'aide d'un identifiant, mais pas pour cacher des informations explicites comme un logo de société afin d'assurer des services d'intégrité.

I.6.2 L'imperceptibilité

Appelé aussi invisibilité, le but est de faire en sorte que le support tatoué reste fidèle au support original.

I.6.3 La robustesse

Le but de cette propriété est de récupérer les données cachées même si le support a été manipulé. On peut définir la robustesse par la résistance du marquage face à des manipulations du support. Dans le cas où le support est une image, les manipulations peuvent être de type géométrique

(rotation, zoom, découpage,…), modifier certaines caractéristiques du support numérique (histogramme des couleurs, saturation,…). Il peut aussi s'agir de tous les types de dégradations fréquentielles de l'image (compression avec pertes).

Il est facile de remarquer que ces trois critères sont contradictoires, si par exemple on augmente la taille de l'information à insérer dans ce cas le support risque d'être détecté, de la même manière si le but est de rendre la marque plus robuste, cela aura en contrepartie pour la rendre plus visible.

Donc il est nécessaire de trouver un compromis entre l'imperceptibilité, la capacité et la robustesse. Ce compromis est généralement représenté par la figure I.3.

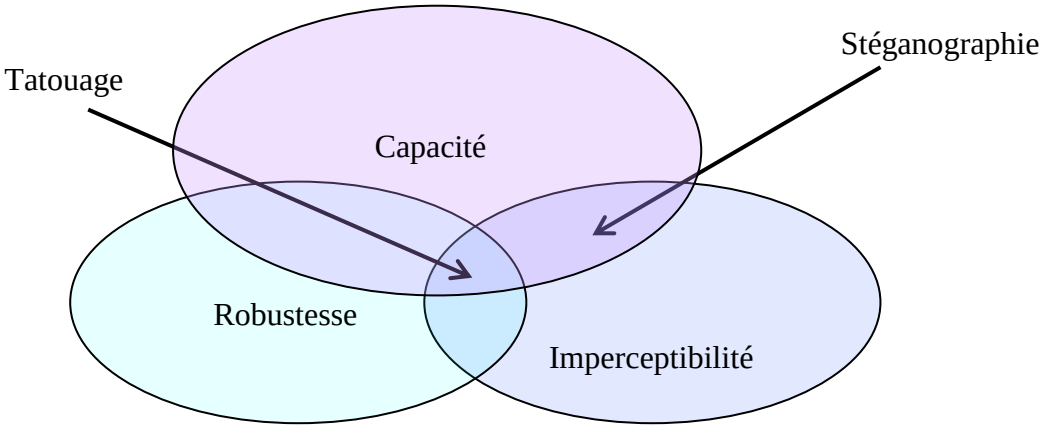


Figure I.3: Compromis entre Imperceptibilité, Capacité et la Robustesse

I.7 Types de tatouage numérique

Plusieurs formes et degrés de tatouages existent**.** Ils sont généralement répertoriés par leurs degrés de priorités : robuste ou fragile et visibles ou non visibles [6].

I.7.1 Tatouage robuste

Il s'agit ici de pouvoir récupérer la marque même si l'image tatouée a été manipulée. Il est nécessaire de distinguer plusieurs types d'attaques selon qu'elles sont considérées comme étant biens ou malveillantes, destructives ou non. Les attaques bienveillantes regroupent les manipulations effectuées par un utilisateur de bonne foi. On trouve dans cette catégorie la compression JPEG, les conversions de format en général, les changements de résolution (zoom), etc.

I.7.2 Tatouage fragile

Le tatouage fragile présente pratiquement un intérêt pour assurer un service d'intégrité de support numérique. L'idée de ce service n'est pas de prouver que oui ou non un support numérique est original ; mais plutôt qu'un document est non falsifié (voir figure I.4).

Figure I.4 : Quelle est la vraie image ?

I.7.3 Tatouage visible

Le principe fondamental du tatouage visible consiste à masquer partiellement un support numérique à l'aide d'une ou plusieurs marques visibles (voir figure I.5), qui ne peuvent être correctement effacées que si l'on possède une clé secrète adéquate.

Figure I.5 : Exemple d'un tatouage visible

I.7.4 Tatouage invisible

Le tatouage invisible peut être considéré comme une forme de stéganographie, puisque l'utilisateur final ignore la présence du tatouage et donc de l'information cachée (voir figure I.6).

Figure I.6 : Exemple d'un tatouage invisible

I.8 Classification des techniques de tatouage

Les techniques de tatouage numérique se distinguent les uns des autres essentiellement par les quatre points clés suivants [5]:

- La manière de sélectionner les blocs dans le support original qui porteront la marque.
- Le choix d'un espace de travail pour réaliser l'opération d'enfouissement (dans le domaine spatial ou transformé comme DCT, ondelettes, FFT etc.).
- La stratégie utilisée pour mettre en forme l'information à cacher avant son enfouissement : redondance, codes correcteurs, bits de resynchronisation.
- La manière de mélanger intérieurement la marque avec le support (modulation) ; l'idée de base consiste le plus souvent à imposer une relation binaire entre les bits du la marque et des caractéristiques choisies de l'image porteuse.

Il existe principalement deux grandes familles de méthodes : celles qui opèrent dans le domaine spatial et celles qui opèrent dans le domaine transformé**.**

I.8.1 Le domaine spatial

Les méthodes de tatouage dans le domaine spatial sont basées, généralement, sur l'insertion de données dans les bits de poids faible ou LSB (Least Significant bit) des pixels, c'est la technique de tatouage d'image la plus connue (voir figure I.7). Elles ont l'avantage d'être facilement implantées mais sont généralement peu robustes aux attaques, l'ajout de bruit la compression de l'image peut facilement dégrader la qualité de l'image ou même supprimer la marque.

Dans le cas d'une image non compressée codée sur 24 bits, chaque pixel est décrit par trois octets. Pour insérer la chaîne de bits 10000011, on utilise le bit de poids faible de chaque octet

(00100111 11101001 11001000)		(00100111 11101000 11001000)
(00100111 11001000 11101001)	+ **10000011**	(00100110 11001000 11101000)
(11001000 00100111 11101001)		(11001000 00100111 11101001)

Figure I.7 : Exemple d'insertion dans les bits de poids faible

Les changements des bits de poids faible (de 0 à 1 ou de 1 à 0) sont totalement imperceptibles pour l'œil puisqu'il n'est modifié que d'un point. Prenons l'exemple d'une image rouge ayant 256 possibilités de représentation. Il est évident qu'à l'œil nu, la différence entre le 254ème et le 255ème rouge n'est pas visible (voir figure I.8).

Figure I.8 : 255ème rouge (a), 254ème rouge (b)

I.8.2 Le domaine fréquentiel

En tatouage, les algorithmes d'insertion dans le domaine fréquentiel sont très couramment utilisés, car les images échangées sur Internet sont le plus souvent les images compressées au format JPEG et TIFF. Ces algorithmes, reposent principalement sur la méthode d'insertion par remplacement des LSBs appliquée aux coefficients DCT quantifiés (Discret consinus Transform) et aux coefficients DWT (Discret Wavelet Transform).

Autres techniques de tatouage numérique sont basées sur l'utilisation de la matrice de quantification utilisant les composantes ayant la même valeur afin de cacher les données (voir figure I.9) [7].

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	**22**	29	51	87	80	62
18	**22**	37	56	68	109	103	77
24	35	55	64	81	104	113	81
49	68	74	87	101	121	120	101
72	92	95	112	112	100	103	33

Figure I.9 : Matrice de quantification utilisée dans la norme JPEG

Par exemple on peut prendre les composantes (4,3) et (5,2) valant 22. Après la multiplication scalaire des deux matrices (DCT, Quantification), on obtient une matrice B. Si l'on veut coder un 1, alors on mettra la donnée dans la composante ayant la valeur la plus élevée, si l'on veut coder un 0, c'est le contraire.

Le principal désavantage de cette technique est que les algorithmes de tatouage fonctionnant avec ce principe ne sont pas très résistants aux transformations géométriques comme les translations ou les rotations.

I.9 Evaluation des méthodes de tatouage

Il est difficile d'évaluer un algorithme de tatouage vu les multiples applications envisagées et les critères qui rentrent en jeu. Il est néanmoins possible d'identifier un des éléments qui influencent l'évaluation de tatouage telle que la qualité de l'image.

Mesure de la qualité de l'image

Il n'existe aucun algorithme capable sans une image de référence de mesurer la qualité (ou le degré de dégradation) absolue d'une image. Cette mesure est basée sur la comparaison de pixels entre l'image originale et l'image Tatouée. Parmi ces mesures nous retrouvons : l'entropie relative, l'erreur quadratique moyenne, l'erreur moyenne absolue et le rapport signal sur bruit.

- **L'erreur quadratique (MSE)**

  L'erreur quadratique compare deux images pixel par pixel. Son expression est définit par:

$$MSE = \frac{1}{MN}\sum_i \sum_j (I(i,j) - I_w(i,j))^2 \quad \text{(I.1)}$$

Où I (i, j) est la valeur de la luminance du pixel (i, j) de référence et I_w (i, j) celle de l'image à tester, les deux images étant de taille MxN. Cette mesure nous donne une indication sur la dégradation introduite au niveau du pixel. Plus le MSE est grand, plus le niveau de dégradation est élevé.

- **L'erreur moyenne absolue (MAE)**

L'erreur moyenne absolue est donnée par :

$$MAE = \frac{1}{MN}\sum_i \sum_j |I(i,j) - I_w(i,j)| \quad \text{(I.2)}$$

Cette mesure quantifie les moyennes des différences absolues dans I et I_w.

- **Le rapport signal sur bruit**

La mesure de distorsion le plus utilisée afin de quantifier la distorsion entre deux images est : le rapport signal sur bruit (Peack Signal to Noise Ratio).
Le PSNR est défini par :

$$PSNR = 10log_{10}(\frac{x_{max}^2}{MSE}) \quad \text{(I.3)}$$

Ou x_{max} désigne la luminance maximale et *MSE* définit l'erreur quadratique moyenne calculée entre les pixels des deux images à comparer. Une valeur de *PSNR* égale à l'infini correspond à deux images parfaitement identiques. Elle décroît en fonction de la distorsion et relie donc l'erreur quadratique moyenne à l'énergie maximale de l'image. On considère généralement en stéganographie qu'un message est imperceptible pour un *PSNR* supérieur à 36 dB, et plus il est élevé, moins la distorsion est importante.

I.10 Attaques de tatouage numérique

L'attaque est définie comme étant tout traitement susceptible d'altérer la marque ou provoquer une ambiguïté lors son exécution.
On distingue plusieurs types d'attaques intentionnelles ou non. Parmi ces attaques nous retrouvons :

- Les transformations géométriques (décalage, rotation, zoom…)
- La compression avec pertes, essentiellement le JPEG
- L'addition d'un bruit
- Le filtrage

I.10.1 Compression JPEG

La compression JPEG est une technique de compression avec perte qui supprime les informations redondantes des images dont le but de diminuer la taille du fichier image. L'avantage de cette méthode réside dans les taux importants de compression que l'on puisse obtenir. Plus celui-ci va être élevé, plus l'on va supprimer une gamme de fréquences importantes et plus l'image va être dégradée.

I.10.2 Ajout de bruit

Le bruit est une altération de l'image : toute l'information pertinente dans l'image n'est pas simplement accessible. Des exemples de bruit artificiel peuvent être :

- ***Le bruit gaussien*** qui consiste à un ajout successif de valeurs générées aléatoirement à chaque pixel de l'image.
- ***Le bruit Salt & Pepper*** (sel et poivre) qui transforme aléatoirement les pixels de l'image en pixels noir ou blanc.

I.10.3 Filtrage

Le bruitage d'une image ayant utilisation particulièrement limitée, voyons à présent les différents types de filtre servant justement à récupérer une certaine compréhension de l'image en y filtrant les bruits. Les filtres les plus utilisés sont : filtre médian, filtre gaussien et filtre moyen.

I.11 Applications de tatouage

I.11.1 Protection des droits d'auteur

La protection des droits d'auteur a été une des premières applications étudiée en tatouage d'image. Ce service reste cependant toujours d'actualité et concerne encore la majorité des publications. L'objectif est d'offrir, en cas de litige, la possibilité à l'auteur ou au propriétaire d'une image d'apporter la preuve qu'il est effectivement ce qu'il prétend être, et ce même si l'image concernée a subi des dégradations par rapport à l'original.

I.11.2 Vérification de l'intégrité du contenu d'une image

L'idée de base consiste à utiliser les techniques de tatouage d'image afin de cacher dans certaines zones de l'image des informations sur d'autres zones. Ces informations servent à alerter l'utilisateur face à une éventuelle modification ou découpe de l'image par une personne non autorisée et à localiser précisément les régions manipulées, voire éventuellement à les restaurer.

I.11.3 Contrôle d'accès

L'objectif est d'ôter tout intérêt commercial au support numérique en y superposant un tatouage. Seules les personnes ayant les droits d'accès sont en mesure d'inverser le processus de marquage de manière à reconstituer le support original. On peut, par exemple, y faire figurer l'adresse où commander le support en clair, le nom de la société, etc.

I.11.4 Indexation

On peut envisager l'utilisation du tatouage afin de faciliter l'accès à des banques de données. La marque n'a pas besoin d'être robuste à de nombreux types d'attaque, puisqu'il ne s'agit plus de protection mais d'identification. Par exemple, un médecin peut inclure dans une radiographie, de façon discrète afin de ne pas la dénaturer, le nom du patient traité, son diagnostic et ses observations. Ce cas est le plus simple, puisqu'une attaque visant à détruire la marque ne présente aucun intérêt et n'est donc a priori pas à craindre.

I.12 Conclusion

Dans ce chapitre, nous avons présenté brièvement, dans une première étape, les outils de la sécurité d'information afin de de montrer la position de tatouage numérique qui fait l'objet de notre projet. Ensuite, nous avons présenté le concept général du tatouage des images ainsi le schéma général de tatouage, ses différents critères, les différentes attaques, ses applications et les domaines d'insertion.

Dans le chapitre suivant, nous allons décrire la biométrie monomodale et multimodale afin de comprendre l'utilisation des données biométriques dans le tatouage numérique des images.

La Biométrie

II.1 Introduction

Les schémas de tatouage numériques, présentés dans le premier chapitre, sont basés, généralement, sur deux étapes importantes : l'algorithme d'insertion/ extraction et le choix de la marque à insérer. Dans la littérature, plusieurs travaux ont été proposés afin de développer des algorithmes d'insertion et d'extraction en fonction de l'application envisagée. Alors que le choix de la marque insérer est toujours limité au nom d'une compagnie, un logo ou même à une information spécialisée au propriétaire du document numérique. Ce type des marques sont des informations moins importantes, moins liées au propriétaire, faciles à falsifier et à reproduire. Récemment, plusieurs travaux ont été proposés utilisant des données biométriques vu de leurs caractéristiques.

Dans ce chapitre nous allons présenter les notions de base de la biométrie, les différentes techniques biométriques et leurs applications. Ensuite, nous allons présenter l'architecture d'un système biométrique ainsi les différentes phases de son fonctionnement. Ensuite, nous allons présenter la biométrique multimodal par l'introduction des différents types de fusion appliquée à la biométrie, les différents niveaux de fusion ainsi la notion de normalisation des scores. Finalement, quelques techniques de protection des droits d'auteur exploitant des données biométrique sont présentées.

II.2 C'est quoi la biométrie ?

La biométrie est une technologie d'identification et de vérification qui consiste à transformer une caractéristique biologique, morphologique ou comportementale en une empreinte numérique.

Définition1

"La biométrie recense nos caractères physiques et comportementaux (voir figure II.1) les plus uniques, qui peuvent être captés par des instruments et interprétés par des ordinateurs de façon à être utilisés comme des représentants de nos personnes physiques dans le monde numérique. Ainsi, nous pouvons associer à notre identité des données numériques permanentes, régulières et dénuées de toute ambiguïté, et récupérer ces données rapidement et automatiquement à l'aide d'un ordinateur." [8].

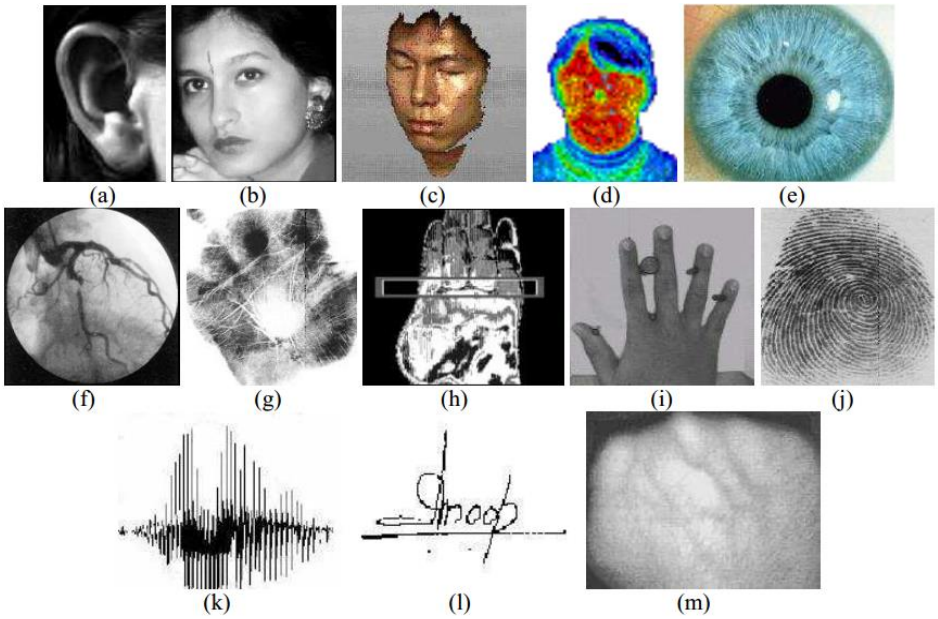


Figure II.1 : Exemple des traits biométriques utilisé pour l'identification [9].

Définition2

"La biométrie est la science qui étudie à l'aide de mathématiques, les variations biologiques à l'intérieur d'un groupe déterminé. " [10].

Définition3

"Toute caractéristique physique ou trait personnel automatiquement mesurable, robuste et distinctif qui peut être employé pour identifier un individu ou pour vérifier l'identité qu'un individu affirme. " [11].

II.3 Modalités biométriques

Il existe différents types de modalités biométriques qui peuvent être classées en trois grandes catégories (voir figure II.2) :

a. **Biométries morphologiques :** sont basée sur une partie du corps humain tel que l'empreinte palmaire, le visage, l'empreinte digitale, l'iris… .
b. **Biométries comportementales :** sont celles utilisant un trait personnel du comportement comme la signature, la dynamique de frappe, la voix… .
c. **Biométries biologiques** : elle regroupe des caractéristiques biologiques comme l'ADN, l'odeur, l'urine, la salive…etc.

Il existe également des biométries morpho-comportementales telles que la voix qui est à la fois liée à la morphologie des cordes vocales mais également au comportement par le fait que la voix peut facilement être modifiée par la personne en fonction de ses états émotionnels.

La figure II.2 présente les différentes modalités biométriques classées dans trois catégories.

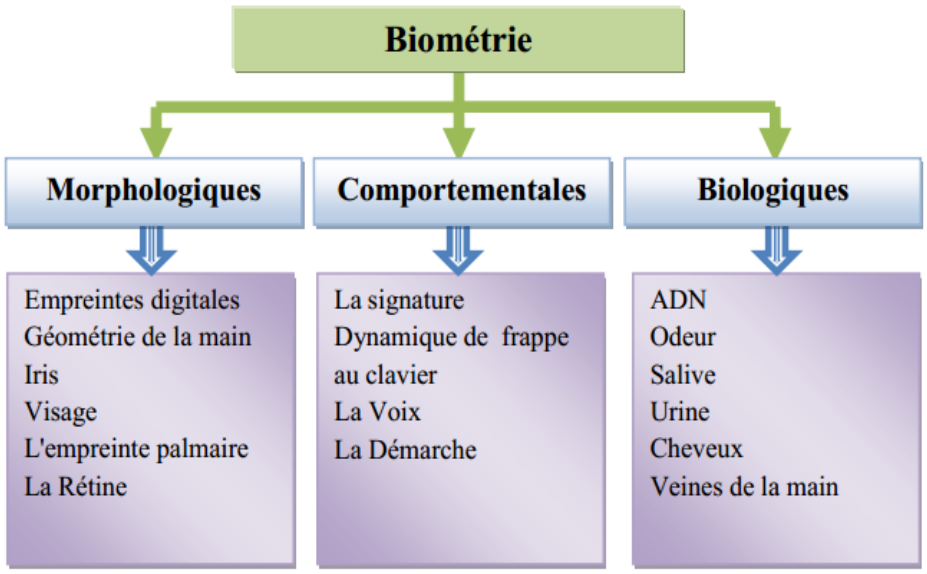


Figure II.2 : Classification d'un certain nombre de modalités biométriques [12]

La comparaison entre les différentes modalités biométriques permet de choisir une modalité en fonction des contraintes liées à l'application. En effet, chaque caractéristique (ou modalité) biométrique a ses forces et ses faiblesses, et faire correspondre un système biométrique spécifique à une application dépend du mode opérationnel de l'application et des caractéristiques biométriques choisies. En France le Club de la Sécurité des Systèmes d'Information Français [13] a proposé une comparaison (avantages / inconvénients) des principales modalités biométriques en se basant sur la facilité ou l'ergonomie d'utilisation, la vulnérabilité aux attaques, aux contournements, la fiabilité relative à la précision et à l'efficacité de la reconnaissance (voir tableau II.1).

Tableau II.1 : Avantages et inconvénients des modalités biométriques [13].

Modalité	Avantages	Inconvénients
Empreintes digitales	Coût, ergonomie moyenne, facilité de mise en place, taille du capteur	Fiabilité des appareils de mesure, acceptabilité moyenne, possibilité d'attaques (rémanence de l'empreinte,...)
Forme de la main	Très ergonomique, bonne acceptabilité	Système encombrant, coût, perturbation possible par des blessures et l'authentification des membres d'une même famille, permanence des données
Visage 2D	Coût, peu encombrant, bonne acceptabilité	Jumeaux, psychologie, déguisement, vulnérabilité aux attaques
Rétine	Fiabilité, pérennité	Coût, acceptabilité faible, installation difficile
Iris	Fiabilité	Acceptabilité très faible, contrainte d'éclairage
Voix	Fiabilité	Vulnérable aux attaques
Signature	Ergonomie	Dépendant de l'état émotionnel de la personne, fiabilité
Frappe au clavier	Ergonomie	Dépendant de l'état physique de la personne

II.4 Etude comparative entre quelques modalités biométriques

Aucune modalité biométrique n'est optimale. La correspondance entre une modalité biométrique et une application dépend du mode opérationnel de l'application et des propriétés de la modalité biométrique (voir tableau II.2).

Tableau II.2 : étude comparative entre les modalités biométriques [14]

Modalités biométriques	Universalité	Distinctif	Permanence	Mesurabilité	Acceptabilité
Empreinte digitale	Moyenne	Haute	Haute	Moyenne	Moyenne
Visage	Haute	Faible	Moyenne	Haute	Haute
Iris	Haute	Haute	Haute	Moyenne	Faible
Rétine	Haute	Haute	Moyenne	Faible	Faible
ADN	Haute	Haute	Haute	Faible	Faible
Signature	Faible	Faible	Faible	Haute	Haute
Voix	Moyenne	Faible	Faible	Moyenne	Haute
Démarche	Moyenne	Faible	Faible	Haute	Haute

Frappe clavier	Faible	Faible	Faible	Moyenne	Moyenne
Géométrie de la main	Moyenne	Moyenne	Moyenne	Haute	Haute
Veines main	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne

II.5 Système biométrique

Les systèmes biométriques s'appuient sur plusieurs processus distincts : enregistrement, capture directe, extraction de modèle et comparaison de modèle (voir figure II.3).

- L'objectif de l'enregistrement consiste à collecter des échantillons biométriques, et à générer des modèles numériques pour des comparaisons ultérieures. Nous pouvons distinguer la "capture directe" de l'enregistrement en la définissant comme le processus visant à collecter des échantillons biométriques en direct lors d'une tentative d'accès ou d'identification, puis à les comparer à une "galerie" de modèles précédemment enregistrés.
- L'extraction de modèle nécessite un traitement du signal des échantillons biométriques bruts (ex : images ou échantillons audio) afin d'obtenir un modèle numérique. Les modèles sont habituellement générés et stockés lors de l'enregistrement pour gagner du temps lors du traitement des comparaisons ultérieures. La comparaison de deux échantillons biométriques applique des calculs algorithmiques destinés à évaluer leur similarité.
- Lors de la comparaison, un score de correspondance est attribué. S'il est supérieur à un seuil donné, les modèles sont considérés comme identiques. En règle générale, les algorithmes d'extraction de modèle biométrique et de comparaison sont propriétaires (différents et secrets), aussi ne peuvent-ils pas être utilisés au sein d'un même système avec ceux d'autres fournisseurs (ex : pour comparer des modèles générés par différents produits, ou pour utiliser un algorithme de recherche de correspondance d'une société afin de comparer des modèles générés par les algorithmes d'une autre société).

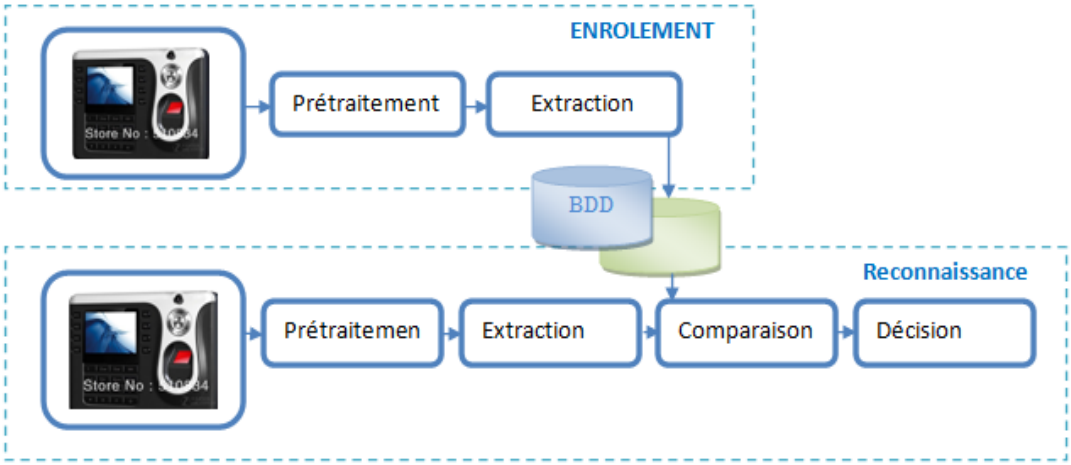


Figure II.3 : Système de reconnaissance biométrique

II.6 Modes de fonctionnement d'un système biométrique

II.6.1 Phase d'enrôlement

C'est une phase d'apprentissage qui a pour but de recueillir des informations biométriques sur les personnes à identifier. Plusieurs campagnes d'acquisitions de données peuvent être réalisées afin d'assurer une certaine robustesse au système de reconnaissance aux variations temporelles des données. Dans cette phase, les caractéristiques biométriques des individus sont saisies par un capteur biométrique, puis représentées sous forme numérique (signatures), et enfin stockées dans la base de données [15].

II.6.2 Phase de reconnaissance

C'est la phase de vérification ou d'identification d'identité de la personne qui veut accéder au système, elle est primordiale dans le fonctionnement de la biométrie, Au cours de cette phase le système effectue une saisie de la donnée biométrique puis un ensemble de paramètres sera extrait comme dans la phase de l'enrôlement. Le capteur utilisé dans la phase de reconnaissance doit être aussi proche de celui utilisé dans la phase d'enrôlement.

Selon le fonctionnement du système, il existe deux modes de reconnaissance :

- **Mode de vérification :** c'est la comparaison 1-à-1, entre les données biométriques capturées (modèle de test) et les données stockées dans sa propre base (modèle d'apprentissage). Dans un tel système, un individu qui désire être identifié réclame une identité, habituellement par l'intermédiaire d'un PIN (numéro d'identification personnelle), d'un nom d'utilisateur, d'une carte d'identité, etc. Le système doit alors

répondre à la question suivante "*Suis-je réellement la personne que suis-je entrain de proclamer ?*" [15].

- **Mode d'identification :** nommée aussi mode d'authentification, le système identifie un individu en cherchant les signatures (Template) de tous les utilisateurs dans la base de données. Par conséquent, le système conduit plusieurs comparaisons 1-à-N pour établir l'identité d'un individu [16]. En résumé, un système biométrique opérant en mode identification répond à la question "*Suis-je bien connu du système ?*".

II.7 Limitations des systèmes biométriques unimodaux

Bien que les techniques de reconnaissance biométrique promettent d'être très performantes, on ne peut garantir actuellement un excellent taux de reconnaissance avec des *systèmes biométriques unimodaux*, basés sur une unique signature biométrique. De plus, ces systèmes sont souvent affectés par les problèmes suivants [17] :

- ✓ **Bruit introduit par le capteur :** du bruit peut être présent dans les données biométriques acquises, ceci étant principalement dû à un capteur défaillant ou mal entretenu.
- ✓ **Non-universalité :** Cependant, toutes les modalités biométriques ne sont pas vraiment universelles. Le *National Institute of Standards and Technologies* (NIST) a rapporté qu'il n'était pas possible d'obtenir une bonne qualité d'empreinte digitale pour environ 2% de la population (personnes avec des handicaps liés à la main, individus effectuant de nombreux travaux manuels répétés, etc.). La non-universalité entraîne des erreurs d'enrôlement dans un système biométrique,
- ✓ **Manque d'individualité :** Par exemple, une certaine partie de la population peut avoir une apparence faciale pratiquement identique due à des facteurs génétiques (père et fils, vrais jumeaux, etc.). Ce manque d'unicité augmente le taux de fausse acceptation,
- ✓ **Sensibilité aux attaques :** bien qu'il semble très difficile de voler les modalités biométriques d'une personne, il est toujours possible de contourner un système biométrique en utilisant des modalités biométriques usurpées. Les études dans [18, 19] ont montrés qu'il était possible de fabriquer de fausses empreintes digitales en gomme et de les utiliser pour contrer un système biométrique.

Ainsi, à cause de tous ces problèmes pratiques, les taux d'erreur associés à des systèmes biométriques unimodaux sont relativement élevés, ce qui les rend inacceptables pour un déploiement d'applications critiques de sécurité. Pour pallier ces inconvénients, une solution

est l'utilisation de *plusieurs modalités biométriques* au sein d'un même système, on parle alors de système biométrique multimodal.

II.8 Systèmes biométriques multimodaux

Le système biométrique multimodal consiste à combiner plusieurs modalités biométriques différentes ainsi que la consolidation d'informations présentées par les différentes modalités peut permettre une authentification précise de l'identité e améliorer les performances de reconnaissance afin de diminuer les tentatives de fraudes. Lors de l'augmentation de la quantité d'informations discriminantes de chaque personne, on souhaite augmenter le pouvoir de reconnaissance du système (vérification ou identification), et diminuer le taux d'erreur.

II.8.1 Architectures de fusion des données

Les systèmes multimodaux associent plusieurs systèmes biométriques et nécessitent donc l'acquisition et le traitement de plusieurs données. L'acquisition et le traitement peuvent se faire successivement, on parle alors d'architecture *en série*, ou simultanément, on parle alors d'architecture en *parallèle*.

Architecture en série : peut être privilégiée dans certaines applications, par exemple le cas d'un individu atteint de cataracte, il est incapable de réaliser une identification d'iris, l'architecture de multimodalité représente pour lui une solution alternative de secours comme l'empreinte digitale ou palmaire (voir figure II.4).

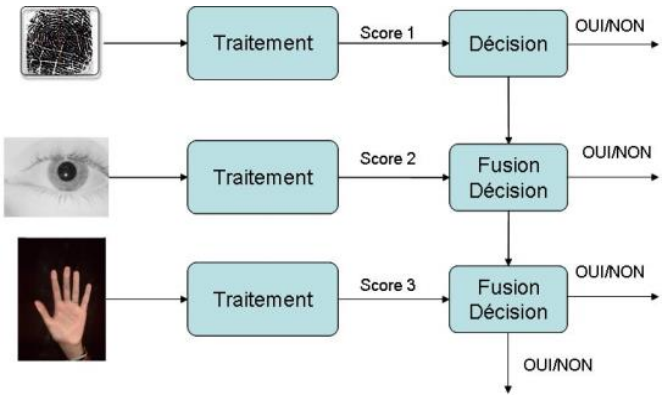


Figure II.4 : Architecture de fusion en série

Architecture en parallèle : nous permet d'utiliser toutes les informations disponibles qui nous aide à améliorer les performances du système, mais, lors de l'acquisition et le traitement

d'un grand nombre de données biométrique le temps et le matériel devient couteux, et réduit le confort d'utilisation (voir figure II.5).

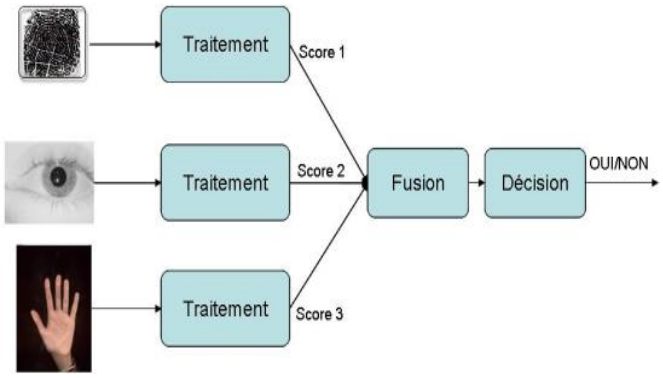


Figure II.5 : Architecture de fusion en parallèle

L'architecture en parallèle (figure II.5) est la plus utilisée car elle permet d'utiliser toutes les informations disponibles et donc d'améliorer les performances du système. En revanche, l'acquisition et le traitement d'un grand nombre de données biométriques est coûteux en temps et en matériel, et réduit le confort d'utilisation. C'est pour cela que l'architecture en série (figure II.5), peut être privilégiée dans certaines applications ; par exemple si la multimodalité est utilisée pour donner une alternative pour les personnes ne pouvant pas utiliser l'empreinte digitale. Pour la majorité des individus seule l'empreinte est acquise et traitée mais pour ceux qui ne peuvent pas être ainsi authentifiés on utilise un système à base d'iris alternativement.

II.8.2 Sources des fusions

Il existe plusieurs méthodes pour effectuer une fusion biométrique (systèmes biométriques multimodaux) indiqué dans la figure II.6.

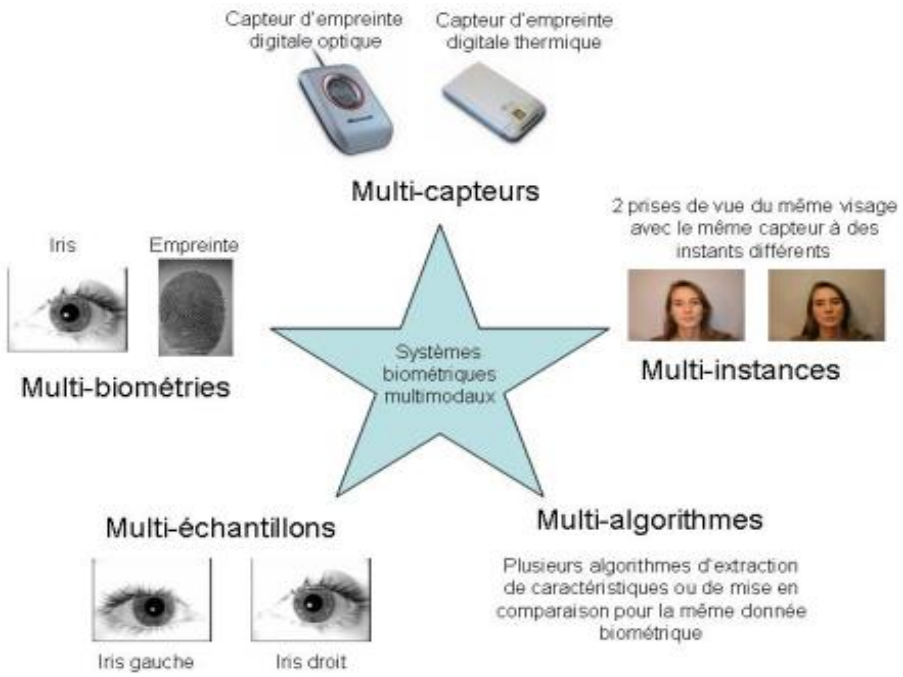


Figure II.6 : Différents systèmes multimodaux.

- ✓ **Multi-capteurs :** lorsqu'ils associent plusieurs capteurs pour acquérir la même modalité, par exemple un capteur optique et un capteur capacitif pour l'acquisition de l'empreinte digitale.
- ✓ **Multi-instances** : lorsqu'ils associent plusieurs instances de la même biométrie, par exemple l'acquisition de plusieurs images de visage avec des changements de pose, d'expression ou d'illumination.
- ✓ **Multi-algorithmes** : lorsque plusieurs algorithmes traitent la même image acquise, cette multiplicité des algorithmes peut intervenir dans le module d'extraction en considérant plusieurs ensembles de caractéristiques et/ou dans le module de comparaison en utilisant plusieurs algorithmes de comparaison.
- ✓ **Multi-échantillons** : lorsqu'ils associent plusieurs échantillons différents de la même modalité, par exemple deux empreintes digitales de doigts différents ou les deux iris. Dans ce cas les données sont traitées par le même algorithme mais nécessitent des références différentes à l'enregistrement contrairement aux systèmes multi-instances qui ne nécessitent qu'une seule référence.
- ✓ **Multi-biométries** : lorsque l'on considère plusieurs biométries différentes, par exemple visage et empreinte digitale.

Tous ces types de systèmes peuvent pallier à des problèmes différents et ont chacun leurs avantages et inconvénients. Les quatre premiers systèmes combinent des informations issues d'une seule et même modalité ce qui ne permet pas de traiter le problème de la non universalité de certaines biométries ainsi que la résistance aux fraudes, contrairement aux systèmes basés sur la multibiométrie. En effet, les systèmes combinant plusieurs informations issues de la même biométrie permettent d'améliorer les performances en reconnaissance en réduisant l'effet de la variabilité intra classe. Mais ils ne permettent pas de traiter efficacement tous les problèmes des systèmes monomodaux. C'est pour cette raison que les systèmes multi-biométries ont reçu beaucoup d'attention de la part des chercheurs.

II.8.3 Niveaux de fusion

L'application de la fusion des informations dans un système biométrique multimodale peut être faite dans n'importe quel module du système. La combinaison de plusieurs systèmes biométriques peut se faire à quatre niveaux différents : au niveau des données, au niveau des caractéristiques extraites, au niveau des scores issus du module de comparaison ou au niveau des décisions du module de décision (voir figure II.7).

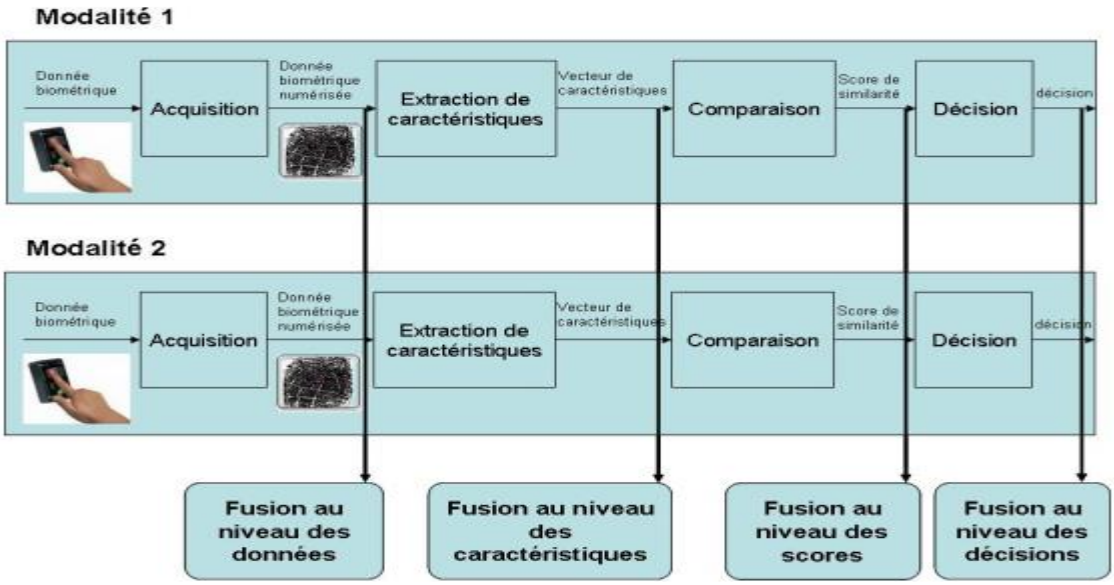


Figure II.7 : Différents niveaux de fusion.

- **Niveau capteur (*Sensor level*)**

Les données brutes provenant des capteurs sont combinées par fusion au niveau capteur [20]. La fusion au niveau capteur peut se faire uniquement si les diverses captures sont des instances du même trait biométrique obtenu à partir de plusieurs capteurs compatibles entre

eux ou plusieurs instances du même trait biométrique obtenu à partir d'un seul capteur (par exemple, il est peut-être difficile de fusionner des images de visages provenant de caméras ayant des résolutions différentes).

- **Niveau des scores (*Score level*)**

Les scores sont issus du module de comparaison qui nous donne comme résultat des scores individuels, ces derniers vont être combiné par une méthode de fusion afin d'obtenir un seul score utilisé pour prendre la décision finale.

Cette approche est la plus utilisée car elle peut être appliquée à tous types de système par des méthodes simples et efficaces.

Supposons que nous avons n scores disponibles, d_i, pour i = 1 à n, issus des n systèmes. Le score résultant D_f est alors donné par :

a. Somme des scores (*sum_score* : *SUM*)

La combinaison des scores par la somme qui consiste à calculer D_f tel que :

$$D_f = \sum_{i=1}^{n} d_i \qquad \text{(II.1)}$$

b. Somme pondérée des scores (*Sum_weighting_score* : *WHT*)

C'est une extension de la somme des scores, le score de chaque système est pondéré en se basant sur le taux d'erreur qui lui est associé, la fusion des scores est calculée comme suit :

$$D_f = \sum_{i=1}^{n} w_i \, d_i \qquad \text{(II.2)}$$

En notant l'erreur d'un système *i* comme w_i, *i*= 1, 2,…n avec n est le nombre total des systèmes. La pondération w_i associée au système *i* est donnée par :

$$w_i = \frac{1/\sum_{j=1}^{n} \frac{1}{\varepsilon_j}}{\varepsilon_j} \qquad \text{(II.3)}$$

Notons que et les $\sum_{i=1}^{n} w_i = 1$ pondérations sont inversement proportionnelles aux erreurs correspondantes et sont par conséquent plus grandes pour les systèmes les plus précis.

c. Minimum des scores (*Min_score : MIN*)

Dans cette méthode, on assigne au score final le meilleur (minimum) score calculé par les différents systèmes. Le minimum est défini comme suit :

$$D_f = min\{d_i\} \quad \text{(II.4)}$$

d. Maximum des scores (*Max_score : Max*)

Dans cette technique, on obtient le maximum des scores au score final (fusionné) de la façon suivante :

$$D_f = max\{d_i\} \quad \text{(II.5)}$$

e. Produit des scores (*Mul-score : MUL*)

Dans cette technique, on combine les scores par le produit qui consiste à multiplier tous les scores tel que :

$$D_f = \prod_{i=1}^{n} d_i \quad \text{(II.6)}$$

Pour réaliser cette méthode, il faut que tous les scores des sous-systèmes soient homogènes. Ainsi, une étape préalable de normalisation des scores est nécessaire :

Normalisation des scores

L'objectif de cette étape est de rendre les scores obtenus des sous-systèmes homogènes, avant de les combiner, ces scores peuvent être de nature différente, exemple : Certains systèmes produisent des scores de similarité**,** et d'autres produisent des distances. En effet, les sorties des systèmes individuels ne sont pas nécessairement inclues dans le même intervalle, et pour ces raisons là qu'il est nécessaire de normaliser les scores avant les combiner.

Les différentes techniques de normalisation de scores sont :

- Normalisation par la méthode Min-Max.
- Normalisation par une fonction quadratique-linéaire-quadratique (QLQ).
- Normalisation par la méthode Z-Score.
- Normalisation par la médiane et l'écart absolu médian (MAD).

- Normalisation par la méthode tangente hyperbolique "Tanh ".
- Normalisation par une fonction double sigmoïde.

Ces méthodes traitent des scores qui varient déjà tous dans le même sens, en général on considère tous les scores sous forme de similarité, et pour transformer des distances en similarité il existe deux solutions : l'inverse ou l'opposé.

- **Niveau des caractéristiques (*Feature Level*)**

Consiste à combiner différents vecteurs de caractéristiques qui sont obtenus à partir d'une des sources suivantes : plusieurs capteurs du même trait biométrique, plusieurs instances du même trait biométrique, plusieurs unités du même trait biométrique ou encore plusieurs traits biométriques.

Quand les vecteurs de caractéristiques sont homogènes (par exemple, plusieurs images d'empreinte digitale du doigt d'un utilisateur), un vecteur unique de caractéristiques résultant peut être calculé comme une somme pondérée des vecteurs de caractéristiques individuels. Lorsque les vecteurs de caractéristiques sont hétérogènes (par exemple, des vecteurs de caractéristiques de différentes modalités biométriques comme le visage et la géométrie de la main), nous pouvons les concaténer pour former un seul vecteur de caractéristiques.

Cependant, la concaténation n'est pas possible lorsque les ensembles de caractéristiques sont incompatibles. Par exemple, les minuties d'empreintes digitales et les coefficients de visages issus de l'ACP (Analyse des Composantes Principales).

- **Niveau de décision *(Decision level*)**

La fusion au niveau des décisions est l'exécution de la combinaison des décisions issue à partir des sous-systèmes unimodaux.

La personne présente son identifiant au système multimodal qui effectue les différentes captures nécessaires à la vérification d'identité, ensuite chaque sous-système (système unimodal) produit son décision binaire sous la forme de : OUI ou NON (0 ou 1 : client ou imposteur), ce qui nous donne comme résultat une série de OUI et de NON. Il existe plusieurs méthodes de fusion, mais les plus utilisées sont les méthodes à base de votes telles que le AND (si tous les systèmes ont décidé 1 alors OUI), le OR (si un système a décidé 1 alors OUI ou le vote à la majorité (si la majorité des systèmes ont systèmes ont décidé 1 alors OUI).

II.9 Mesures de performance d'un système biométrique

Les performances d'un système biométrique sont mesurées par trois critères principaux illustrés dans la figure II.8 :

1. Taux d'erreurs égales (Equal Error Rate EER) : Il est calculé à partir des deux taux : taux de faux rejet et le taux de fausse acceptation et qui correspond à l'endroit où le FRR=FAR (le meilleur compromis entre les faux rejets et les fausses acceptations).

2. Taux de faux rejet (False Reject Rate FRR) : Il représente le pourcentage des personnes censées être reconnue mais qui sont rejeté par le système.

3. Taux de fausse acceptation (False Accept Rate FAR) : Il représente le pourcentage des personnes censées ne pas être reconnue mais qui sont accepté par le système.

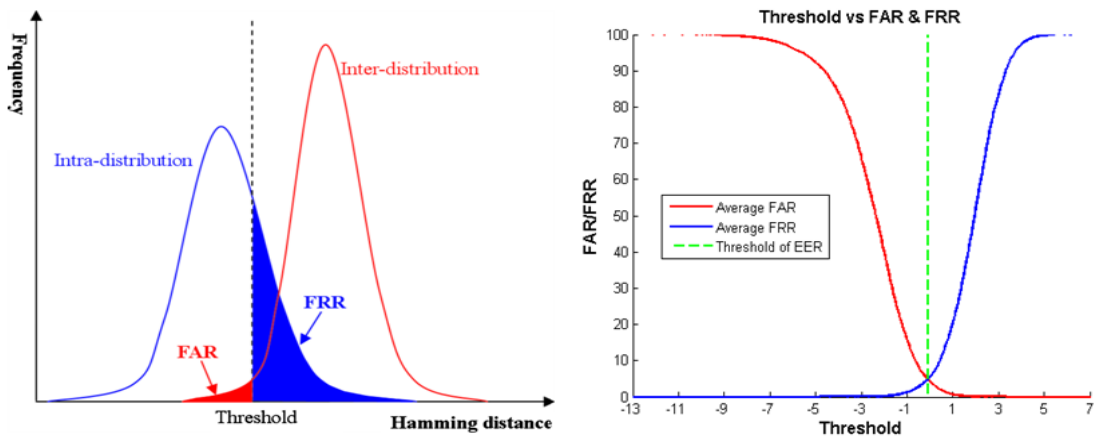


Figure II.8 : Mesures de performance d'un système biométrique : FRR, FAR et EER.

Dans les applications utilisant un système de vérification d'identité biométrique, un des paramètres importants à régler est le seuil de décision. Ce seuil va dépendre du type d'application et des performances souhaitées. En effet, certaines applications requièrent un taux de Fausses Acceptations (FAR) très faible (contrôle d'identité par exemple), alors que certaines autres ne tolèrent pas de hauts taux de Faux Rejets (FRR) (identification sur un ordinateur personnel par exemple). C'est pour cela que l'on calcule souvent les performances des systèmes à plusieurs points de fonctionnement afin de pouvoir connaître les performances du système pour chaque type d'application. Pour visualiser les performances des systèmes biométriques lorsque le seuil varie, on utilise également des courbes de performance.

➢ **Les courbes de performance**

Les courbes de performances permettent de représenter les performances pour toutes les valeurs du seuil sans fixer un seuil a priori. Par exemple on peut représenter l'évolution des deux taux d'erreurs (FAR et FRR) lorsque le seuil varie pour les distributions de scores Client et Imposteur représentés sur la Figure II.9. Sur la Figure II.10, on peut lire les valeurs des taux d'erreurs pour chaque valeur du seuil. Comme les taux d'erreurs FAR et FRR dépendent tous les deux du même seuil de décision, on peut également représenter sur une courbe la variation du FRR en fonction de FAR lorsque le seuil varie. Ces courbes s'appellent des courbes ROC (Receiver Operating Characteristic) [21], représentées sur la Figure II.9, ou des courbes DET (Detection Error Tradeoff) [21], représentées sur la Figure II.10 lorsque les échelles pour les deux taux d'erreurs sont logarithmiques.

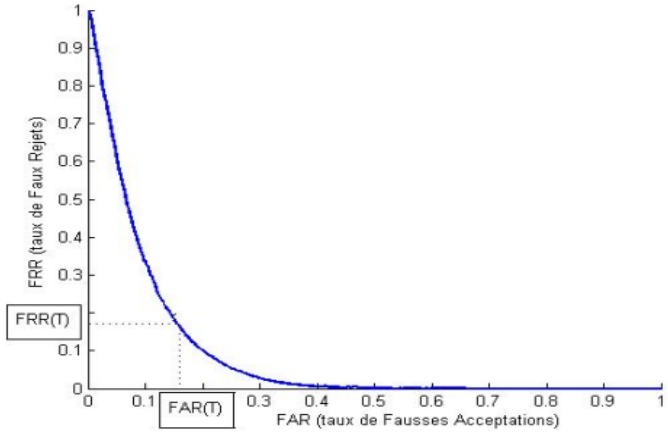


Figure II.9 : Courbe ROC : Variation du taux de Faux Rejets (FRR) en fonction du taux de Fausses Acceptations (FAR) lorsque le seuil de décision varie.

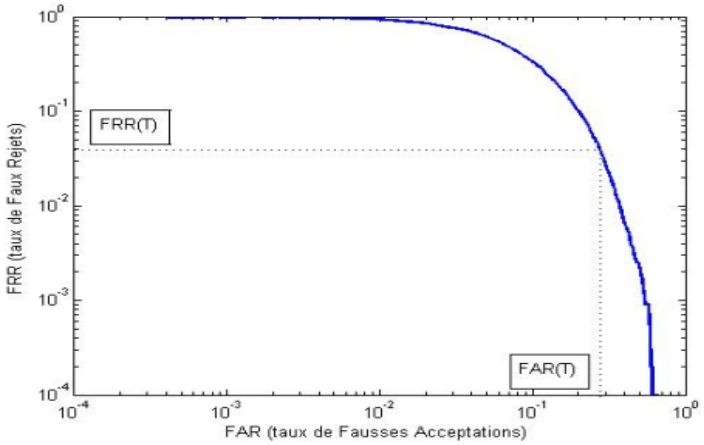


Figure II.10 : Courbe DET : Variation du taux de Faux Rejets (FRR) en fonction du taux de Fausses Acceptations (FAR) en échelle logarithmique lorsque le seuil de décision varie.

II.10 Tatouage biométriques

Récemment, différentes méthodes de tatouage ont été développées basées sur l'utilisation des modalités biométriques des individus pour améliorer l'authentification du droit d'auteur. Cette nouvelle idée est connue par le tatouage biométrique et le tatouage multibiométrique.

- Cheng et al. [22] ont proposés une nouvelle approche de tatouage biométrique afin d'insérer la signature manuscrite d'un individu comme une marque d'authentification dans l'image originale. Après la construction de la marque qui est définit par un vecteur de caractéristique, la technique de remplacement des LSBs (Least Significant Bits) est utilisée dans les coefficients d'ondelettes discrète (DWT). Les résultats expérimentaux ont montrés la robustesse de cette méthode contre les attaques géométriques, en particulier la compression JPG, filtrage passe-bas, filtrage médian, addition de bruit et la rotation.

- Une nouvelle approche est proposée dans [23] basée sur l'utilisation l'empreinte digitale dans une image à des fins de preuve de propriété par une autorité. Le principe de cette approche consiste à calculée un vecteur de caractéristique à partir d'une empreinte digitale. Ensuite, l'algorithme de BioHashing est utilisé pour transformer un vecteur de caractéristique de taille *m* à un vecteur binaire de taille *n* nommé Hash code. Les résultats expérimentaux montrent que cette approche atteint un ERR nul basé sur la l'empreinte digitale et le Hash code.

  L'inconvénient principal de cette approche, qui a été appliquée à diverses modalités biométriques (visage, empreinte digitale, empreinte palmaire) [24-29], est la faible performance lorsqu'un imposter B vole le Hash code A et essaie de s'authentifier comme A. Lorsque ce problème se produit, le La performance de BioHashing peut être inférieure à celle obtenue en utilisant uniquement les données biométriques.

- Dans un travail récent [30], les auteurs mettent en évidence les anomalies des approches basées sur la méthode BioHashing et conclut que la vérification basé sur la modalité biométrique est insuffisante. Plusieurs modalités biométriques sont utilisées dans ce travail, le visage, l'empreinte digitale et la signature manuscrite, afin d'étudier les points faible de la méthode BioHashing. Après une étude expérimental des paramètres utilisés dans l'algorithme de BioHashing, ils concluent qu'un vecteur de

caractéristique taille élevé permet une meilleur vérification mais ceci devient en contradiction avec les contraintes d'insertion dans système de tatouage (capacité et robustesse). Dans une deuxième étape, ils ont étudiés l'influence d'un paramètre de quantification par la variation de ce dernier entre l'intervalle [-300,300].

II.11 Conclusion

De nos jours la protection des droits d'auteur est réaliser à l'aide des données biométriques car elle est le moyen de sécurité le plus utilisé grâce à la variabilité des données biométriques est ses avantages.

Dans ce chapitre, nous avons, dans une première étape, présenté les différentes modalités biométriques, leurs caractéristiques et une comparaison entre elles. Ensuite, nous avons présenté les systèmes biométriques unimodaux et multimodaux avec les divers types de combinaisons des modalités possibles, les architectures et les niveaux de fusion qui peuvent être utilisés. Finalement, nous avons terminé ce chapitre par la présentation de quelques techniques de tatouage numérique basées sur les données biométrique pour la protection des droits d'auteur.

Résultats Expérimentaux

III.1 Introduction

Récemment, dans le domaine de la vérification des droits d'auteurs, les technologies biométriques devient plus en plus très utilisées à cause de leurs impacts sur le degré de sécurité et de fiabilité sur le système de sécurité d'information. Cependant, le vecteur des caractéristiques biométriques doit être respecté deux contraintes très essentiels, à savoir leur discrimination, exigé par le système biométrique, et leur petite taille exigé par l'opération de tatouage. Dans ce chapitre, nous nous intéressons tout d'abord à présenter une méthode de tatouage biométrique basique basée sur l'empreinte palmaire (PLM) et l'empreinte des veines de la paume (PLV). D'autre part, certaines techniques d'extractions des caractéristiques biométriques doivent être examinées afin d'améliorer la performance de système. Ensuite, pour ajouter une réelle plus, nous présenterons une nouvelle méthode pour cacher les traces biométriques (dissimilation ou BioHashing). Les expérimentations menées dans ce chapitre, réalisées sur deux bases des données récentes, mettront en évidence que notre méthode est plus efficace que plusieurs méthodes existantes.

III.2 Tatouage biométrique

Dans cette dernière décennie, beaucoup de travaux de recherche ont proposé différentes méthodes pour gérer les droits d'auteurs des images en utilisant, par exemple, des protocoles cryptographiques. Cependant, ces schémas réalisent une vérification peu sécurisée du propriétaire de la donnée. En effet, ces systèmes peuvent difficilement lier l'identité de l'individu à ces droits d'utilisation. Afin de pallier ce problème, des chercheurs ont pensé à utiliser de la biométrie. Le tatouage biométrique consiste à insérer une marque dans une image comme une preuve de propriété. La marque insérée doit être calculée à partir d'une

modalité biométrique liée à l'identité de l'individu afin de permettre une vérification d'identité sécurisée et respectueuse de la vie privée.

III.2.1 Description de système

Le système de vérification des droits d'auteurs consiste à insérer de manière visible ou invisible des informations dans une image pour protéger les droits d'auteur. Ensuite, cette marque est extraite afin que le propriétaire d'une image puisse prouver qu'il est bien l'auteur ou le propriétaire de celle-ci. Ce type de système est basé essentiellement sur le tatouage. En effet, dans le système basé sur le tatouage biométriques, la marque à insérée est sous forme d'un vecteur des caractéristiques biométriques qui représente le propriétaire de l'image. Ici, la biométrie peut servir à augmenter la performance de reconnaissance de l'identité comme il peut servir à améliorer la sécurité de système. Bien sûr, le vecteur des caractéristiques biométriques de propriétaire de l'image doit être discriminatif et possède une taille réduite.

III.2.2 Phases de tatouage

Le tatouage biométrique consiste à insérer une marque ou vecteur biométrique dans une image comme une preuve de propriété. Ce processus contient deux phases principales :

1) Insertion de la marque : Pendant la phase d'insertion, le système va acquérir une ou plusieurs mesures biométriques qui serviront à construire un vecteur des caractéristiques de l'identité de propriétaire ou simplement une marque. Cette marque servira de point de comparaison lors de la phase de vérification. Comme montre la figure III.1, cette phase se compose de plusieurs modules :

Figure III.1 : Processus de l'insertion de la marque biométrique

- **Capture :** Appelée l'interface utilisateurs, ce module fournit les mécanismes pour qu'un utilisateur indique son identité et entré ses modalités biométriques dans le système.
- **Prétraitement** : éviter les informations intitule qui existent dans l'image des modalités biométriques acquise.

- **Extraction des caractéristiques** : pendant cette phase un vecteur des caractéristiques biométriques est créé. Ce vecteur se comporte comme une marque.
- **Dissimulation d'informations :** Appelée BioHashing, ce module sert à cacher la marque ou le vecteur de caractéristiques.
- **Insertion** : Ce le module qu'effectue l'insertion de la marque dans l'image. Il est à noter que l'insertion peut se faire suivant une clé secrète qui sera alors nécessaire à l'extraction.

2) Extraction et vérification de la marque : Au cours de la deuxième phase ou la phase de vérification (voir figure III.2), le vecteur des caractéristiques ou la marque d'une telle personne est extrait et cacher comme lors de l'enrôlement.

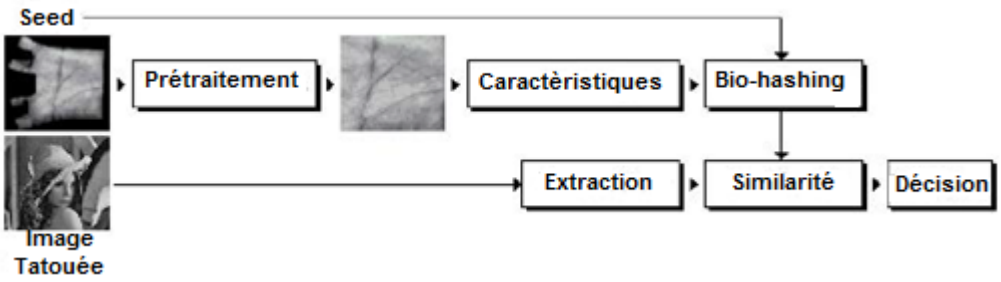


Figure III.2 : Processus d'extraction et de vérification de la marque biométrique

Ensuite, après l'extraction ou la récupération de la marque pré-insérée dans l'image, une opération de comparaison ou une mesure de similarité entre celle-ci et la marque extraite est exécutée. Cette comparaison est basée sur la distance de *Hamming* parce que les deux marques sont binaires. Enfin, le dernier module de décide si l'utilisateur est bien l'auteur ou le propriétaire de l'image tatouée.

III.3 Génération de la marque

Généralement, les applications du tatouage biométrique nécessitent l'ajout de relativement peu de données. Dans ce cas, le mécanisme d'insertion doit ajouter une marque dont la taille représente une part non-significative de celle du l'image. Malheureusement, la diminution de la taille de la marque influe sur le taux de vérification de système biométrique. Donc, la précision de ce type de système dépendra alors grandement de l'efficacité de la méthode d'extraction des caractéristiques à représenter l'utilisateur par une marque de taille réduite vis-à-vis l'image originale et en même temps aux distinctif.

III.3.1 Extraction des caractéristiques

Dans le système de tatouage biométrique, après avoir fait l'acquisition d'une modalité on réalise l'extraction des caractéristiques (*Feature extraction*) dont le processus de vérification a

besoin. Diverses techniques ont déjà été proposées pour l'extraction des caractéristiques biométriques liées à l'empreinte ou à la modalité biométrique. Généralement, les techniques basées sur la texture de l'image sont très rependus.

1) Motifs binaires locaux : Les motifs binaires locaux, en anglais Local Binary Pattern (LBP), ont initialement été proposés par *Ojala* [31] afin de caractériser les textures présentes dans des images en niveaux de gris. Le concept du LBP est simple, il propose d'assigner un code binaire à un pixel en fonction de son voisinage. Ce code décrivant la texture locale d'une région est calculé par seuillage d'un voisinage avec le niveau de gris du pixel central.

Afin de générer un motif binaire, tous les pixels voisins prendront alors une valeur "1" si leur valeur est supérieure ou égale au pixel courant et "0" autrement (voir figure III.3). Les pixels de ce motif binaire sont alors multipliés par des poids et sommés afin d'obtenir un code LBP du pixel courant. On obtient donc pour toute l'image, des pixels dont l'intensité se situe entre 0 et 255 comme dans une image à 8 bits ordinaire. Plutôt que de décrire l'image par la séquence des motifs LBP, elle est décrite par un histogramme de dimension 255.

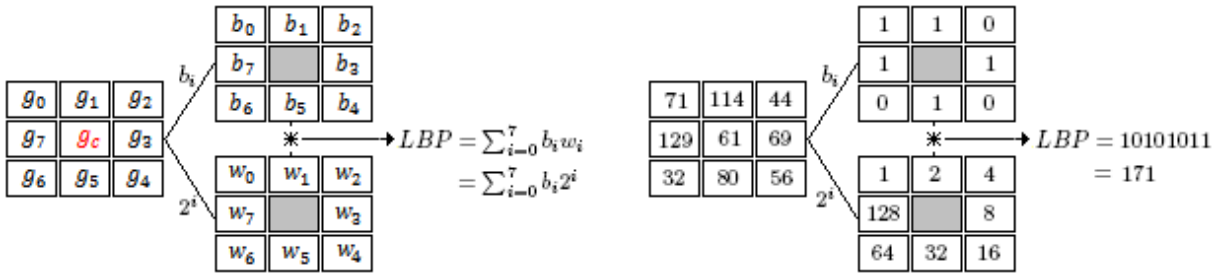


Figure III.3 : Construction d'un motif binaire et calcul du code LBP

Pour calculer un code LBP dans un voisinage de *P* pixels, on compte simplement les occurrences de niveaux de gris g_p plus grands ou égaux la valeur centrale.

$$LBP(x_c, y_c) = \sum_{n=0}^{p-1} u(g_i - g_c) * 2^n \qquad \text{(III.5)}$$

Ou x_c et y_c les coordonnées du pixel central, et g_i et g_c sont respectivement les niveaux de gris d'un pixel voisin et du pixel central. La fonction *u(x)* est définit comme suit:

$$u(x) = \begin{cases} 1 & si \quad x \geq 0 \\ 0 & autrement \end{cases}$$

Malgré la performance et l'efficacité de cette méthode mais il reste toujours un inconvénient major c'est que la méthode LBP est sensible à la rotation et au bruit ce que pousse les chercheurs de proposer plusieurs variantes basées sur LBP de base.

2) Histogramme des gradients orientés : L'histogramme des gradients orientés, en anglais Histogram of Oriented Gradients (HOG) est une nouvelle méthode utilisée en vision par ordinateur pour la détection d'objet et la détection des régions d'intérêts [32]. Le principe de base de cette méthode est de calculer les histogrammes locaux de l'orientation du gradient sur une grille dense, c'est-à-dire sur des zones régulièrement réparties sur l'image.

L'objectif de la méthode HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. Le principe de cette méthode consiste à diviser une image à des régions adjacentes de petite taille, appelées cellules, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule. La combinaison des histogrammes forme alors le descripteur HOG. Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, en calculant une mesure de l'intensité sur des zones plus larges que les cellules, appelées des blocs, et en utilisant cette valeur pour normaliser toutes les cellules du bloc. Cette normalisation permet une meilleure résistance aux changements d'illuminations et aux ombres.

III.3.2 Dissimulation (BioHashing)

L'algorithme de BioHashing sert à cacher les traces biométriques, qui sont sous forme d'un vecteur des caractéristiques, par une transformation non inversible de celui-ci vers un autre vecteur généralement binaire, ce dernier représente la marque à insérée dans l'image.

1) BioHashing basique : Dans cette méthode, la dissimilation est basée sur la projection des vecteurs biométriques dans un autre espace ou domaine (voir figure III.4).

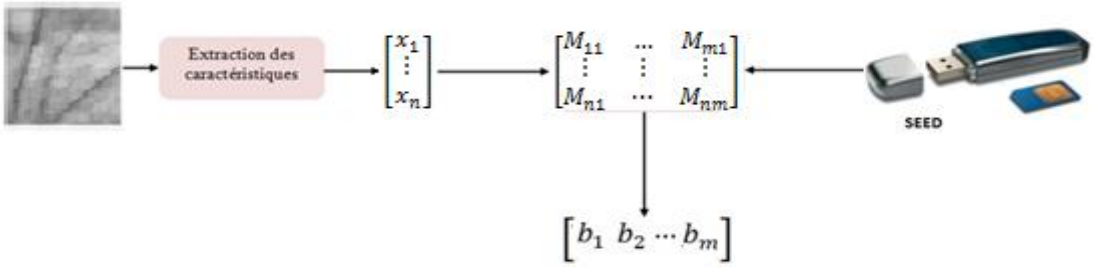


Figure III.4 : Protection des vecteurs biométriques en utilisant le BioHashing basique.

Ensuite, le résultat obtenu est binariser afin d'obtenir des vecteurs binaires qui représentent les marques. Cette projection (basée sur des matrices aléatoires et orthogonales) a été proposée afin de répondre à la propriété de révocabilité. Les étapes de BioHashing sont :

- Basé sur un mot de passe (appelé *seed)*, une matrice aléatoire *M*, de taille *n*x*m*, est générée.

La hauteur *n* de la matrice doit être le même que la taille de vecteur des caractéristiques,

elle est donc imposée par la méthode d'extraction des caractéristiques biométriques. La largeur *m* de la matrice permet de contrôler la taille de la marque, elle est donc imposée par la méthode d'insertion de la marque dans l'image (l'algorithme de tatouage).

- Transformer la matrice générée, *M*, en une matrice avec des vecteurs orthogonaux, Δ. Cette transformation est appliquée par un algorithme dit Gram-Schmidt (voir annexe A.2).

$$\Delta = F_{GS}(M)$$

- Projeter le vecteur des caractéristiques biométriques, *z*, en utilisant la matrice Δ afin de le transformer en un autre vecteur, z.

$$y = \Delta z$$

- Finalement, coder le vecteur transformé *y* comme suit :

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

τ est le seuil de binairisation. En pratique, ce seuil est choisi égal à 0 car les résultats de la projection ont la même probabilité d'être négatifs ou positifs. Ainsi, chaque bit b_i du *y* aura la même probabilité d'apparition ce qui a comme effet d'augmenter le contenu de l'information réellement présente dans *B* et ainsi sa robustesse.

2) BioHashing proposée: Plusieurs travaux montrent que le vecteur résultant de la méthode de BioHashing basique à l'inconvénient d'être non-descriptif, ce que nous donnent des taux de vérification très faible. Cependant, l'augmentation de la taille de vecteur des caractéristiques n'influe vraiment pas sur la discrimination de vecteur à cause de l'effet de la quantification (binarisation). En outre, l'augmentation de la largeur de la matrice de projection permet de donner des grands vecteurs, qui ne sont pas approprié pour l'insertion. Afin de résoudre ce problème, une nouvelle méthode, basée sur un dictionnaire des exemplaires, est proposée.

Le dictionnaire des exemplaires est une technique utilisée surtout dans le domaine de la compression d'images, il est basé sur le principe de quantification vectorielle. Dans notre méthode, le vecteur des caractéristiques biométrique est en deux dimensions. Le codage de ce vecteur consiste à remplacer chaque colonne de vecteur par un entier qui désigne leur cordonnée dans le dictionnaire. Ensuite, chaque composent de vecteur résultant est codé en binaire sur *k* bits. Les nombre des bits nécessaire pour le codage, *k*, est fixé en fonction de la taille de dictionnaire (voir figure III.5).

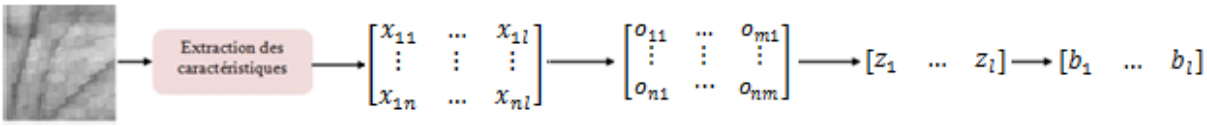


Figure III.5 : Protection des vecteurs biométriques en utilisant le BioHashing proposée.

Plusieurs méthodes de formulation de dictionnaire ont été proposée dans la littérature, dans notre travail, nous avons proposé aussi une méthode de formulation de dictionnaire. Avant de présenter cette méthode, nous avons obligés de changer un peu la méthode d'extraction des caractéristiques pour que celui-ci donne un vecteur en deux dimensions. Cette changement consiste à appliquer la méthode HOG sur des blocs, des tailles *bxb*, au lieu de l'appliquée sur l'image entière. Tous les vecteurs des blocs sont concaténés afin d'obtenir un vecteur deux dimensions. La taille de bloc est un paramètre critique. La variation de ce paramètre permet de donner des vecteurs 2D avec plusieurs tailles.

La formulation de dictionnaire nécessite une base de vecteurs d'apprentissage, pour ce faire, des modalités biométrique des plusieurs personnes ont été utilisées pour construire cette base. Cette formulation repose sur la classification supervisée des vecteurs.

- Calculer les valeurs moyennes de tous les vecteurs de la base d'apprentissage.

$$Vect_m = [m_1 \quad \dots \quad m_N]$$

- Calculer la valeur max et min des valeurs moyennes obtenues

$$m_{max} = \max(Vect_m) \qquad m_{min} = \min(Vect_m)$$

- Calculer les valeurs moyennes des tous les centroides des classes. Généralement, le nombre des classes *L* est choisi supérieur à la taille de dictionnaire.

$$C_m = \frac{m_{max} - m_{min}}{L}$$

- Moyenner tous les vecteurs qui ont des valeurs moyennes proches au même centroides.
- Supprimer toutes les classes qui sont vides.
- Finalement, choisissez les premier *K* classes, où *K* est la taille de dictionnaire.

Il est à noter que dans la phase de vérification, après l'extraction de la marque, une opération de décodage entropique ainsi qu'une opération de quantification sont appliquées sur cette marque afin d'arrivé au vecteur des caractéristiques. En outre, le module de mesure des similarités utilise une distance euclidienne.

III.4 Evaluation des performances

Deux modalités biométriques, à savoir l'empreinte palmaire et l'empreinte de réseau veineux de la paume sont utilisés afin de testé notre système. En effet, les résultats

expérimentaux que nous allons présenter sont divisés en deux catégories. Nous donnerons tout d'abord les résultats qui ont été obtenus concernant la méthode de BioHashing basique, nous donnerons également des résultats concernant une étude comparative de deux techniques d'extraction des caractéristiques (LBP-MAT et HOG-MAT). La deuxième partie des résultats se concentre tout particulièrement sur la méthode BioHashing proposée (HOG-CDB) en exploitant la technique d'extraction des caractéristiques améliorée (HOG).

III.4.1 Bases d'images

Pour l'étude de notre système, nous avons utilisé deux bases des données. Ces deux bases ont été créées on utilisant une base des données des empreintes palmaires multispectrales (MSP) [33]. Cependant, les trois couleurs (rouge, vert et bleu) sont utilisés pour obtenir une image niveau de gris de l'empreinte palmaire (PLP), tandis que, la bande proche infrarouge donne l'empreinte de réseau veineux de la paume (PLV). La base MSP, crée par l'Université de Polytechnique de Hong Kong (PolyU), a été obtenu par la collection d'images d'empreinte palmaire multispectrale de 300 individus à l'aide d'un dispositif de capture d'empreinte palmaire multispectrale, ces individus sont des étudient et des travailleurs dans l'Université Polytechnique de Hong Kong. Dans ce jeu de données, 195 personnes sont des mâles et les restes sont des femelles, et la distribution de l'âge entre 20 et 60 ans. Les gens ont été invité pour fournir d'environ de 12 images. Les images ont été recueillies dans deux occasions, où les six premières images ont été capturées lors de la première occasion et les six autres ont été capturées dans la deuxième occasion. L'intervalle moyen entre la première et la deuxième occasion est de 9 jours. Les images d'empreinte palmaire collectées dans la seconde occasion, ont été fournies sous différentes conditions d'éclairage. Cette base de données (PolyU) contient 3600 images en quatre bandes (totale des bandes égale à 14400). Toutes les images originales ont une taille de 352 x 288 pixels et une résolution < 100 dpi.

III.4.2 Protocole des tests

Dans notre travail, nous nous intéressons plus spécifiquement à la tâche de vérification. Dans ce problème, la tache de système est de vérifier si la marque entrante accompagnée d'une identité de la personne, correspond réellement à l'identité de lui. Dans le cadre de l'expérimentation réalisée dans cette phase de vérification, nous avons utilisé une base des données contenant 100 personnes (12 images pour chaque personne). Une seule image pour chaque personne, soit 100 images, sont utilisées pour construire les marques à insérées dans l'image, et les 1100 images restantes pour tester les performances des systèmes proposés. Ainsi, la distribution des imposteurs et des clients est généré par 110000 comparaisons ou

distances. En outre, une méthode conventionnelle de tatouage basée sur la DCT (voir chapitre II) est utilisée. De plus, l'image à tatouée est une image niveau de gris de taille 256x256 pixels, ce que me permet d'inséré au maximum une marque de taille 1024 bits.

III.4.3 Résultats des tests

1) Systèmes unimodaux : Dans cette partie, notre système utilise à chaque fois une seule modalité biométrique (PLP ou PLV). En outre, cette partie est subdivisée en deux sous-parties. La première sous-partie traite la BioHashing basique utilisant les deux méthodes d'extractions des caractéristiques (LBP-MAT et HOG-MAT). La deuxième sous-partie focalise sur la BioHashing proposée utilisant la méthode d'extraction des caractéristiques HOG-CDB. Dans les deux sous-parties, des tests concernant une attaque (bruit blanc et bruit salt & pepper) sont exécutés. Dans tous les tests, le choix de la taille de marque est effectué en fonction de taux d'erreurs égaux (Equal Error Rate-EER).

✍ **BioHashing basique:** Les résultats des tests de la méthode LBP-MAT sont illustrés dans le tableau III.1. La taille de la marque est variée on fait varié largeur de la matrice.

Tableau III.1 : Résultats de LBP-MAT pour les deux modalités.

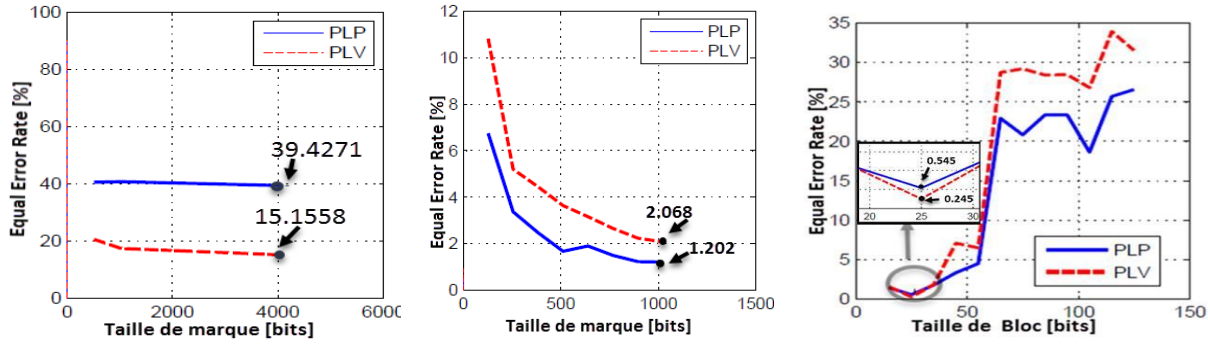
Taille de la marque (bits)	512	1024	4096
PLP	40.60	40.77	39.43
PLV	20.63	17.44	15.16

La première vue de ce tableau montre la pauvreté de ces résultats. Une grande EER est obtenu avec les deux modalités et avec des tailles de la marque très grandes. De plus, il est clair d'après ces résultats que la modalité PLV est très performant vis-à-vis la modalité PLP. Une erreur de 17.44 % est obtenue avec le PLV au lieu de 40.77% pour le PLP pour une taille de 1024 bits. La mauvaise performance de système de vérification nous a poussés d'utiliser une autre méthode d'extraction des caractéristiques (HOG-MAT). Les résultats de cette méthode sont présentés dans le tableau III.2.

Tableau III.2 : Résultats de HOG-MAT avec les deux modalités

Taille marque (bits)	PLP		PLV	
	EER	PSNR	EER	PSNR
128	6.75	Inf	10.82	inf
256	3.36	8.18e+03	5.20	8.25e+03
384	2.47	5.14e+03	4.44	5.27e+03
512	1.67	3.91e+03	3.63	3.93e+03
640	1.90	3.05e+03	3.15	3.00e+03
768	1.50	2.49e+03	2.65	2.45+03
896	1.23	2.11e+03	2.23	2.13e+03
1024	**1.20**	**1.86e+03**	**2.07**	**1.84e+03**

Le tableau III.2 montre que la méthode HOG-MAT donne des résultats très acceptables pour une taille de marque égale à 1024 bits. Cependant, une EER égale à 1.20 % pour un seuil T_o égal 0.1977 est obtenue pour la modalité PLV. Cette erreur devient 2.07% pour un T_o égal 0.1973 pour la modalité PLP. En outre, des faibles erreurs sont obtenues pour une taille de la marque égale à 896 bits, dans ce cas, l'EERs sont égales à 1.23% et 2.23%, respectivement, la PLV et la PLP. Dans la phase de tatouage, la taille de marque influe sur la qualité de l'image tatouée. Ainsi, en utilisant le rapport signal-bruit, en anglais Peak Signal to Noise Ratio (PSNR), pour mesurer distorsion dans l'image. L'image originale. Toujours, les résultats de PSNR sont représentés dans le même tableau, il est logique que l'augmentation de la taille de la marque permet d'augmenté la distorsion dans l'image tatoué. Les figures III.6 (a) et III.6 (b) montrent une comparaison entre les deux modalités avec les deux méthodes d'extractions des caractéristiques.



(a) *(b)* *(c)*

Fig III.6 : Performances des sytèmes avec les différentes méthodes en utilisant les deux empreintes. *(a)* Méthode LBP-MAT, *(b)* Méthode HOG-MAT et *(c)* Méthode HOG-CDB.

Un type d'attaque, correspond au l'ajout de bruits, est testé. Les deux bruits, salt & pepper et le bruit blanc (*gaussien*), avec des variances entre 0.02 et 0.3 sont donc évalués. La figure III.7 illustre l'effet de ces deux bruits sur l'image.

(a) ***(b)***

Figure III.7 : l'effet de bruit sur l'image. *(a)* bruit salt & pepper et *(b)* bruit blanc.

Les résultats dans le tableau III.3 sont basés sur les meilleurs cas sélectionnés dans les premiers tests.

Tableau III.3 : Performance de système sous la présence des bruits.

Type de bruit	Taille de marque	Variation	PLV	PLM
Salt & Pepper	1024	0.02	2.140	1.182
		0.04	2.355	1.722
		0.06	3.075	2.093
		0.08	3.989	1.801
		0.10	4.124	2.498
		0.30	16.292	11.712
	896	0.02	2.4553	1.385
		0.04	3.014	1.674
		0.06	4.502	1.866
		0.08	4.764	2.586
		0.10	4.693	3.711
		0.30	17.6391	12.428
Bruit blanc	1024	0.02	3.9037	2.2789
		0.04	4.4982	2.6224
		0.06	9.1765	4.7543
		0.08	8.9409	5.9444
		0.10	12.1323	7.2090
		0.30	24.8209	15.0817
	896	0.02	3.688	2.512
		0.04	6.101	4.454
		0.06	7.732	5.172
		0.08	10.377	6.738
		0.10	11.956	8.573
		0.30	22.574	17.002

D'après ce tableau, le bruit *salt & pepper* donne un taux d'erreur pour l'empreinte PLP égale à 1.18% avec une variance de 0.02 et 2.1401% pour l'empreinte PLV avec la même variance. L'augmentation de la variance permet d'augmenter aussi le taux d'erreur jusqu'à 11.7123% avec une variance de 0.3 pour la PLP et 16.30% pour la PLV. Il est à noter, que l'influence de bruits sur les résultats de vérification due au l'algorithme de tatouage. Cependant, une méthode de tatouage robuste au bruit permet de réduire leur effet sur la phase de vérification.

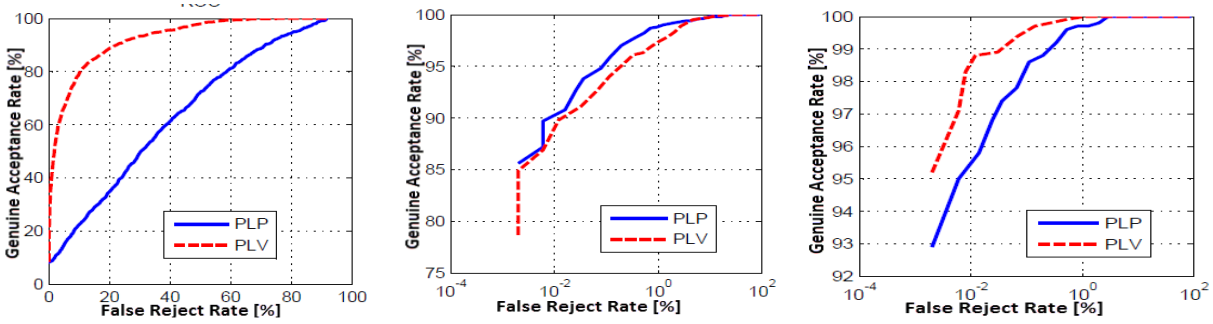
✍ **BioHashing proposée :** Avant de présenter de manière détaillée la performance du système, nous allons détermine dans cette section le meilleur type de bloc et la taille de dictionnaire qui donnent une meilleur performance dans l'opératoire de vérification. L'expérience consiste à comparer les différents types de bloc (15x15 jusqu'à 85x85), afin de sélectionner le meilleur type utilisé. Pour pouvoir comparer les différents types entre elles, il faut pouvoir déterminer quelles sont celles qui permettent la meilleure. Les résultats de comparaison présentés dans le tableau III.4 ont été obtenus sur les deux modalités. Ce tableau doit permettre de comparer les performances de chacune des types de bloc avec une variation des tailles de dictionnaire.

Tableau III.4 : Résultats de HOG-CDB avec les deux modalités

Taille Bloc	Taille marque	Taille Dictionnaire	EER-PLV	PSNR-PLV	EER-PLP	PSNR-PLP
15	2112	2048	1.545	2.7515e+3	1.408	2.7455e+3
25	**750**	**1024**	**0.254**	**8.5108e+3**	**0.545**	**8.5108e+3**
35	243	512	1.805	Inf	1.745	Inf
45	96	256	7.085	Inf	3.336	Inf
55	18	256	6.475	Inf	4.472	Inf
65	18	64	28.735	Inf	22.908	Inf
75	18	64	29.195	Inf	20.805	Inf
85	18	64	28.415	Inf	23.355	Inf
95	18	64	28.455	Inf	23.345	Inf
105	18	64	26.815	Inf	18.615	Inf
115	18	64	33.925	Inf	25.674	Inf
125	18	64	31.535	Inf	26.568	Inf

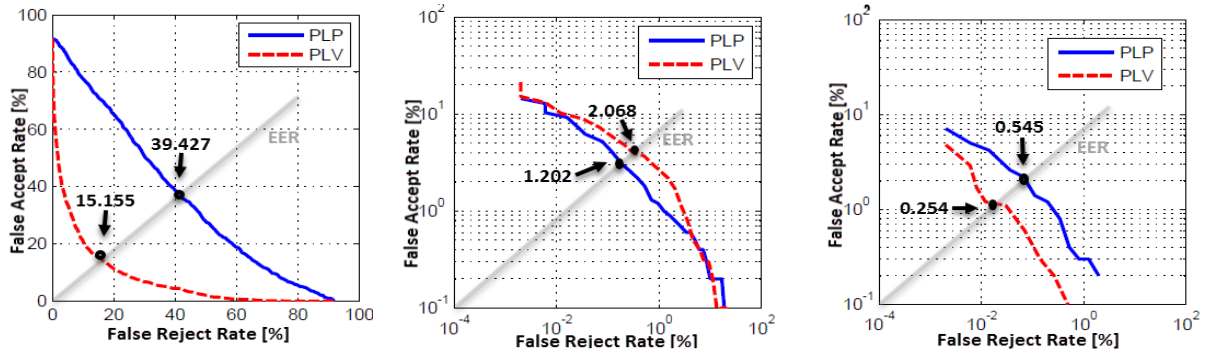
Ces derniers graphiques montre bien la haute performance dans le cas de type de bloc 25x25. Un excellent taux d'erreur, égal à 0.254% pour un seuil *To* égal à 0.6771, avec une taille réduite de la marque égale à 750 bits est obtenu. Avec le même type de bloc, une erreur égale à 0.545% (To= 0.6863) est obtenue. Il est très remarquable que notre méthode peut

donner une acceptable erreur (EER = 6.475% et 4.472%) pour une petite taille de la marque égale à 18 bits pour une taille de bloc égale à 55x55 pixels. Les figures ci-contre montrent les performances des systèmes sous les différentes méthodes d'extractions des caractéristiques.



(a) *(b)* *(c)*

Figure III.8 : Courbes ROCs de de syetems sous les differentes methodes.*(a)* LBP avec une marque de taille 4096 bits, *(b)* HOG-MAT avec une marque de taille 1024 bits et *(c)* HOG-CDB avec une marque de taille 750 bits.



(a) *(b)* *(c)*

Fig III.9 : Courbes DETs de de syetems sous les differentes methodes.*(a)* LBP avec une marque de taille 4096 bits, *(b)* HOG-MAT avec une marque de taille 1024 bits et *(c)* HOG-CDB avec une marque de taille 750 bits.

La même attaque avec les mêmes variances des bruits a été appliquée en utilisant pour tester la robustesse de système de vérification. Après avoir sélectionné la meilleure taille de bloc pour les deux modalités, l'image tatouée est bruitée avec les deux types des bruits (salt & pepper et bruit blanc). Les résultats des tests montrent l'efficacité de La méthode HOG-CDB vis-à-vis les méthodes précédentes (LBP-MAT et HOG-MAT) (voir tableau III.5). Pour le bruit blanc par exemple et avec la taille de marque 750 bits (bloc de taille 25x25 pixels) et pour presque toutes les valeurs de variance (0.02 à 0.30), le taux d'erreur est entre 2.948% et 3.374% pour la PLV et entre 2.145% et 2.339% pour la PLP.

Tableau III.5 : Performance de système (HOG-CDB) sous la présence des bruits.

Type de bruit	Taille de bloc	Variation de bruit	PLV	PLP
Salt & Pepper	25	0.02	1.039	1.425
		0.04	3.728	3.035
		0.06	11.675	6.905
		0.08	18.385	14.775
		0.10	26.802	24.105
		0.30	45.845	46.354
	35	0.02	5.294	3.455
		0.04	14.795	9.815
		0.06	21.645	18.804
		0.08	28.655	27.375
		0.10	36.102	27.983
		0.30	47.688	46.176
Bruit blanc	25	0.02	2.948	2.145
		0.04	2.470	2.195
		0.06	2.069	2.401
		0.08	2.154	2.812
		0.10	2.854	1.805
		0.30	3.374	2.339
	35	0.02	11.455	7.557
		0.04	12.179	6.875
		0.06	11.752	7.914
		0.08	11.052	6.740
		0.10	12.040	7.312
		0.30	11.064	7.803

2) Systèmes multimodaux : Dans cette série d'expériences, nous avons testés la performance de système de vérification dans le cas où la marque est construite à partir des deux modalités biométriques. Deux scenarios de fusion ont été testés, à savoir le système multi-biométrique (fusion de deux modalités biométriques, PLP et PLV) et le système multi-algorithmiques (fusion de deux marques obtenues par deux méthodes d'extraction des caractéristiques pour la même empreinte biométrique).

✍ **Système multi-biométrique:** Les résultats des évaluations dans le cas de la méthode de BioHashing basique (HOG-MAT) sont représentés dans le Tableau III.6. On voit ici que la combinaison des deux empreintes améliore efficacement la précession de système vis-à-vis l'utilisation d'une seule modalité, presque pour toutes les règles de fusion. Cependant, la fusion de PLP et PLV donne la meilleure performance, avec des erreurs *EER* égales à 0.775 % (seuil, T_o = 0.3770) et 0.960% (seuil, T_o = 0.3226) pour respectivement une taille de la

marque égale à 2048 et 1792 bits avec la règle de fusion SUM. Dans le cas de présence des bruits, l'erreur devient 0.750 % et 1.058% pour le bruit blanc et salt & pepper, respectivement. Vis-à-vis les systèmes unimodaux, une amélioration considérable est apportée sur la précession de système, ce que me prouve expérimentalement l'intérêt de la fusion pour le système de classification.

Tableau III.6 : Performance de système multi-biométrique (HOG-MAT)

Taille de marque	2048	1792
SUM	**0.7745**	**0.9595**
MAX	1.1029	1.1772
MIN	1.1595	1.3284
MULTI	0.8887	1.0402
SUM-WHT	0.9167	0.9661
Présence des bruits (variance = 0.02) Salt & Pepper		
Taille marque	2048	1792
SUM	0.7500	1.0583
MAX	1.0000	1.3588
MIN	1.3333	1.4951
MULTI	0.8436	1.0161
SUM-WHT	0.8212	1.000
Présence des bruits (variance = 0.02) bruit blanc		
Taille marque	2048	1792
SUM	1.5380	1.7090
MAX	1.8845	2.0979
MIN	2.1278	2.2792
MULTI	1.6793	1.7640
SUM-WHT	1.5743	1.7793

L'utilisation de la méthode proposée (HOG-CDB) est aussi testée, et les résultats sont présentés dans le tableau III.9. Il est clair que les différentes règles débouchent sur une amélioration remarquable des performances par comparaison à celles offertes par les modalités prise individuellement. Cependant, cette fusion conduit à une erreur faible égale à 0.105 % pour un seuil T_o égal à 0.6859 dans le cas de la règle SUM-WHT. Dans ce cas une taille réduite de la marque (égal à 1500 bits), vis-à-vis le premier cas, est obtenue. Aussi, le système peut fonctionner avec une ERR égale à 0.686% pour une minimum taille de marque qu'égale à 486 bits. Ce résultat permet de confirmer encore la supériorité de la fusion de données. D'après le Tableau, la fusion peut aussi améliore la performance de système dans de présence de bruits,

Tableau III.7 : Performance de système multi-biométrique (HOG-CDB).

Taille marque	1500	486
SUM	0.111	0.705
MAX	0.333	0.958
MIN	0.117	1.451
MULTI	0.113	1.482
SUM-WHT	**0.105**	**0.686**
Présence des bruits (variance = 0.02) Salt & Pepper		
Taille marque	1500	486
SUM	1.018	5.046
MAX	1.035	5.294
MIN	1.038	5.294
MULTI	1.046	6.455
SUM-WHT	0.232	2.028
Présence des bruits (variance = 0.02) bruit blanc		
Taille marque	1500	486
SUM	0.955	3.697
MAX	1.723	6.5594
MIN	1.530	6.951
MULTI	0.675	8.810
SUM-WHT	0.841	5.061

✍ **Systèmes multi-algorithmiques:** Dans ce scenario de fusion, la même modalité biométrique set analysée avec deux méthodes d'extraction des caractéristiques et les résultants sont fusionnés. En effet, dans le but de diminuer le taux d'erreur, nous allons mettre en commun les deux méthodes HOG-MAT et HOG-CDB pour concevoir un système multi-algorithmiques. Toutes les règles de fusion (SUM, MAX, MIN, MULTI et SUM-WHT) sont testées. Aussi, les meilleurs paramètres, minimisant l'EER, dans deux méthodes, HOG-CDB et HOG-MAT, sont sélectionnés. D'après le tableau III.8, nous pouvons observons que toutes les règles de fusion améliorent les performances du système, cela confirme l'avantage de la fusion. La règle de fusion SUM-WHT offre la meilleure performance par rapport aux autres règles, avec un taux d'erreur égal à 0.198% pour la modalité PLV et un seuil T_o égal à 0.377 et 0.400% pour la modalité PLP avec un seuil T_o est égal à 0.654.

Tableau III.8 : Performance de système multi-algorithmique (PLP et PLV).

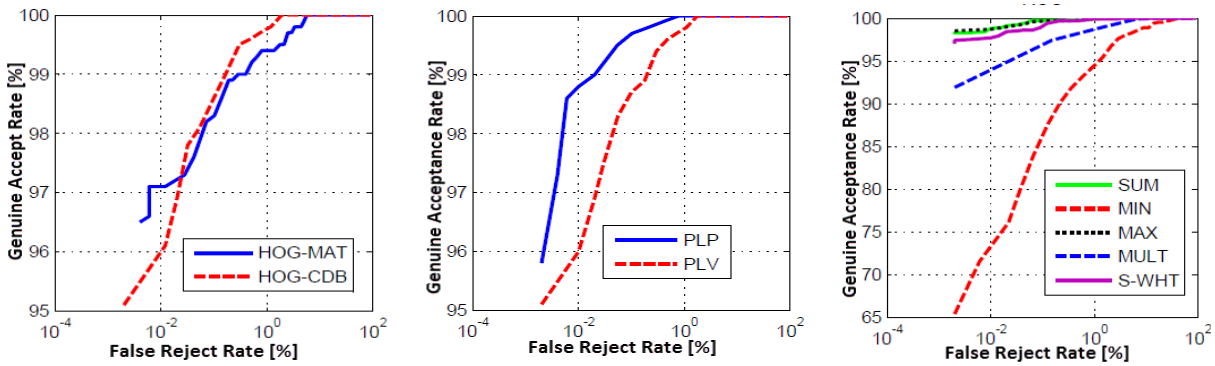
FUSION multi-algorithmique (PLV)		
Taille marque	750+256=1006	768+243=1011
SUM	0.7498	1.3007
MAX	0.2296	1.2773
MIN	4.9558	2.381
MULTI	3.0494	1.5801
SUM-WHT	**0.198**	**1.1403**
FUSION multi-algorithmique (PLP)		
Taille marque	750+256=1006	768+243=1011
SUM	1.6493	0.8708
MAX	0.4886	1.4193
MIN	2.8643	1.1683
MULTI	1.9818	0.9219
SUM-WHT	**0.4000**	**0.8168**

✍ **Systèmes hybrides :** Les systèmes hybrides peuvent ses composés de plusieurs scénarios de fusion. Dans notre travail, les systèmes multi-biométriques et les systèmes multi-algorithmiques sont combinés au niveau score afin de diminuer le taux d'erreur. Le Tableau III.9 présente les taux d'erreur obtenus avec les différentes règles de fusion (SUM, MAX, MIN, MULTI et SUM-WHT). Nous remarquons que la règle SUM-WHT offre le meilleur résultat ; soit avec la taille de marque égal à 2012 bits (0.2857% avec un seuil de To=0.6859), ou avec la taille 2022 bits (0.4926% avec un seuil To=0.2437), par contre la règle MIN donne le mauvais résultat (2.6669%). Noter que les marques des grandes tailles peuvent être insérés facilement dans des images à grandes résolutions ainsi que dans des images couleurs ou multi/hyper spectrale (ces images contiennent plusieurs bandes spectrale.

Tableau III.9 : Performance de systèmes hybrides

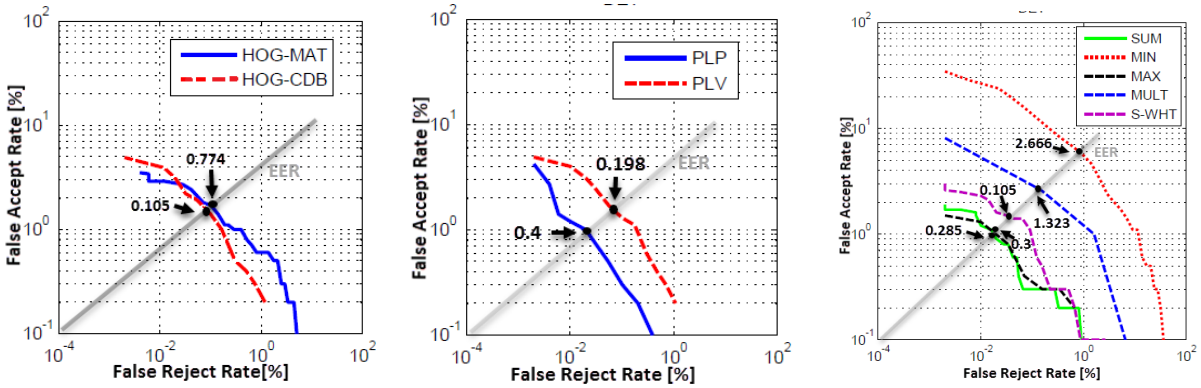
Taille de marque	2012	2022
SUM	0.2857	0.4926
MAX	0.3000	0.6224
MIN	2.6669	1.3936
MULTI	1.3232	1.4810
SUM –WHT	**0.105**	**0.7785**
Taille de marque	3548	2278
SUM	0.1982	0.5835
MAX	0.334	0.764
MIN	1.234	1.288
MULTI	0.4333	1.481
SUM – WHT	0.255	1.806

Les différentes figures ci-dessous présentent des comparaisons entre les différent cas.



(a) *(b)* *(c)*

Figure III.10 :Résultats des systèmes multimodaux (courbe ROC). (a) Multi-biomtriques des deux méthodes. (b) Multi-algorithmique avec les deux modalités (c) Système hybride avec les cinq régles de fusion.



(a) *(b)* *(c)*

Figure III.11 : Résultats des systèmes multimodaux (courbe DET). (a)Multi-biomtriques des deux méthodes. (b) Multi-algorithmiques avec les deux modalités (c) Système hybride avec les cinq régles de fusion.

III.5 Comparaison

Après avoir effectué les tests de la performance de système, nous allons faire une étude comparative entre les deux méthodes HOG-MAT et HOG-CDB dans les différentes architectures (unimodal et multimodal). L'objectif c'est de faire le bon choix entre les architectures et les méthodes proposées pour concevoir un système de tatouage biométrique efficace. D'après les résultats obtenus, nous remarquons que la combinaison des plusieurs systèmes unimodaux offre une performance très élevée ; il montre une bonne séparabilité des classes clients et imposteurs, avec l'utilisation de la multi-algorithmiques et la multi-biométrique. D'autre part, la méthode proposée (HOG-CDB) montre la forte efficacité vis-à-vis la méthode basique (HOG-MAT). Une différence très remarquable entre les résultats de

HOG-MAT et de HOG-CDB, apparaitre, que ce soit dans le taux d'erreur ou dans la taille de marque à insérée. Comme conclusion, nous pouvons dire que notre méthode proposée offre la possibilité de produire des marques distinctives avec des petites tailles. En outre, l'ensemble des tests effectués montrent que l'utilisation des systèmes multimodaux améliore efficacement le taux d'erreur.

Tableau III.10 : Synthèse des résultats obtenus

	EER (%)	Taille Marque (bits)	Seuil	Image
HOG-CDB	0.202	750	0.6771	Image de 256x256
Multi algorithmique (SUM-WHT)	0.198	1006	0.3770	Image de 256x256
Multi biométrie (SUM-WHT)	**0.105**	1500	0.6859	Image couleur /multi spectrale
Système hybrides (SUM-WHT)	**0.105**	2012	0.6859	Image couleur /multi spectrale

III.6 Conclusion

Dans ce chapitre, nous avons proposé une méthode qui permet d'extrait une marque contient des données biométries, cette marque est insérée dans une image dans le but de construire un système de tatouage biométrique d'images. A travers les différents résultats obtenus, on a constaté que la méthode proposée (HOG-CDB) montre une meilleure performance pour les deux modalités (PLP et PLV). Ainsi, l'utilisation des systèmes multimodaux donne une efficacité à notre système et améliore des résultats.

Conclusion Générale

Le travail présenté dans ce mémoire s'inscrit dans le contexte de la vérification des droits d'auteur. Nous avons utilisés, séparément, la méthode de BioHashing basique et une autre méthode que nous avons proposée basée sur un dictionnaire des exemplaires afin de générer la marque à insérer à partir de deux modalités biométriques, à savoir l'empreinte palmaire (PLP) et l'empreinte de réseau veineux de la paume (PLV).

Après avoir introduit les concepts généraux de tatouage numérique des images et la biométrie, nous avons présenté les deux méthodes d'extraction de caractéristiques utilisées: HOG (Histogram Oriented Gradient) et LBP (Local Binary Pattern). Ensuite, nous avons décrit la méthode de BioHashing basique ainsi notre méthode de dissimulation proposée. Une étude comparative de deux techniques d'extraction des caractéristiques (LBP-MAT et HOG-MAT) est donnée utilisant la méthode de BioHashing basique. D'autre part, nous avons testé la méthode de BioHashing proposée (HOG-CDB) en utilisant la technique d'extraction des caractéristiques améliorée (HOG). Outre les systèmes unimodaux, nous avons testé quelques systèmes multimodaux. Ces différents systèmes sont testés dans le but d'améliorer les performances de système.

En validant ces systèmes sur une base de données type de 100 personnes, nous avons dégagé que notre méthode offre la possibilité de produire des marques distinctives avec des petites tailles. En outre, l'ensemble des tests effectués montrent que l'utilisation des systèmes multimodaux améliore efficacement le taux d'erreur.

Perspective

Le travail n'est cependant fini, nous avons proposés une méthode permettant de réaliser une preuve de propriété d'un média en utilisant une vérification biométrique. D'une manière plus générale, nous envisageons de poursuivre le thème de recherche en utilisant un schéma de tatouage d'image plus robuste.

Cependant, l'ambition est de poursuivre le thème de recherche dans le domaine de la protection de copie (filigrane). Cet élargissement passera par l'étude d'autres applications que la protection des droits d'auteurs, donc d'autres application impliquent d'autres outils théoriques (la présence d'une marque) et d'autres architectures systèmes pour protéger les nouvelles valeurs.

Annexe

1. Algorithme BioHashing

L'algorithme BioHashing sert à transformer un vecteur de taille *n* à un vecteur binaire de taille *m* telle que *m*<*n*. Le principe de cette méthode est définit comme suit :

- Étant donné une clé secrète *K* appelé seed et un vecteur de caractéristiques *v* représentant une donnée biométrique.
- générez une séquence de nombres pseudo aléatoires afin de construire un ensemble de vecteurs pseudo-aléatoires V_i . Nous notons qu'une variété d'algorithmes de génération de nombres pseudo aléatoires sont publiquement disponibles. Dans la méthode proposée, nous avons utilisé la méthode Blum-Blum-Shub.
- Transformer les vecteur V_i en un ensemble de vecteur orthonormé en utilisant la procédure de Gram-Schmidt afin d'obtenir une matrice *M* de vecteurs linéairement indépendants.
- Calculer le produit entre le vecteur de caractéristiques biométriques *v* et la matrice *M*.

- Binariser le vecteur obtenu, à partir d'un seuil τ_b afin d'obtenir un vecteur binaire $B=\{b_1,b_2,\dots,b_m\}$ tel que :

$$b_i = \begin{cases} 0 & si\ b_i \leq \tau_b \\ 1 & si\ b_i > \tau_b \end{cases}$$

2. Algorithme de Blum Blum Shub

L'algorithme Blum Blum Shub est un algorithme qui génère des nombres pseudo-aléatoire. Le principe de cet algorithme est définit comme suit :

Entrées :

- P, Q : deux nombres premiers initialiser par P= 5651 ; Q=5623.
- H : la taille des nombres générer.

Traitement :

1. calculer N =p× q.
2. Choisissez un nombre aléatoirement appelé *Seed* tel que :
   0 < *Seed* < N & PGCD (Seed, N) =1.
3. Initializer X_0 à $(Seed)^2$
4. Choisissez un nombre aléatoirement appelé *Seed* tel que :
   0 < *Seed* < N & PGCD (Seed, N) =1.
5. Calculer $X_n=X_{n-1}$ mod N
6. Aller à l'étape 4 *H-1* fois

1. Orthogonalisation de Gram-Schmidt

Le procédé d'ortho-normalisation de Gram-Schmidt est un algorithme qui permet de créer une matrice orthonormée à partir d'une matrice libre. On utilise le procédé de Schmidt dans le but de calculer des projetés orthogonaux et avoir une matrice orthogonale et indépendante.

On construit une base orthogonale *K* par récurrence, selon un procédé appelé le procédé d'orthogonalisation de *Gram-Schmidt*.

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

$$u_1 = v_1 \; et \; u_1 \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - proj_{v1} \; et \; u_2 \frac{u_2}{\|u_2\|}$$

$$u_i = v_i - \sum_{k=1}^{i} -1 proj_{uk} v_i \; et \; u_i \frac{u_i}{\|u_i\|}$$

- *proj* est l'opérateur de projection orthogonale.
- $proj_u v = \langle u,v \rangle u$.
- $\langle u,v \rangle$ est le produit scalaire.
- ||.|| la norme d'un vecteur.

Références

[1] Vidyasagar M. P., Song H., Elizabeth C. “A Survey of Digital Image watermarking Techniques”. School of Information Systems, Curtin University of technology, Perth, Western Australie. 2009.

[2] Dugelay J.L. & Roche S. “Introduction au tatouage d’images”. Annales des Télécommunications, 54, no 9-10, pp. 427-437, 1999.

[3] Petitcolas, F.A.P., Anderson, R.J. and G.K.M. “Information hiding: A survey”, Proceedings of the IEEE, special issue on protection of multimedia content, vol. 87, N°7, pp. 1062-1078, July 1999.

[4] Johann B., “Analyse de Canaux de Communication dans un Contexte non Coopératif”, Thèse pour obtenir le grade de docteur, ESAT - Laboratoire de Virologie et Cryptologie, B.P. 18, 35 998 Rennes Cedex, 2007.

[5] Frédéric R. “Etude d’Outils pour la Dissimulation d’Information”, Thèse de doctorat, Université paris XI, 2002.

[6] Laimeche L., Merouani F.H., “Détection des informations cachées dans les images numériques”, Thèse de magistère, Université Badji Mokhtar, Annaba, 2009.

[7] Hartung F. and Girod B., “Watermarking of uncompressed and compressed video”, Signal Processing, vol. 66, N° 3, pp. 283-233, 1998.

[8] www.aware.com/ WP_WhatareBiometrics_0514_French_v01.

[9] Maltoni, D. Maio, A.K. Jain, S. Prabhakar, “Handbook of Fingerprint Recognition”, Springer, 340 pages, 2005.

[10] Wodward J.D. & al., *Biometrics, “*A Look at Facial Recognition”, Documented Briefing prepared for the Virginia State Crime Commission.

[11] Vidyasagar M. Potdar, Song Han, Elizabeth Chang. “A Survey of Digital Image watermarking Techniques”. School of Information Systems, Curtin University of technology, Perth, Western Australie. 2009.

[12] Faisal T., "Reconnaissance de la paume de la main", Ecole nationale Supérieure d’Informatique (ESI) Oued-Smar, Alger, 2010. "

[13] http://www.clusif.asso.fr "Technique de Contrôle d'Accès par Biométrie", dossier technique, CLUSIF, 2004.

[14] Dang H. V., "Biométrie pour l’Identification", Rapport final, Institut de la Francophonie pour l'Informatique, Hanoï, Vietnam, 07 – 2005.

[15] Fedias M., "Combinaisons de données d’espaces couleurs et de méthodes de vérification d’identité pour l’authentification de visages", Université Mohamed Khider, Biskra.

[16] Dang H. V., "Biométrie pour l'Identification", Rapport final, Institut de la Francophonie pour l'Informatique, Hanoï, Vietnam, 07 – 2005.

[17] Jain A. K. and Ross A. "Multibiometric systems". Communications of the ACM, special issue on multimodal interfaces, Vol. 47, No. 1, pp. 34–40, January 2004.

[18] Matsumoto T., Matsumoto H., Yamada K., and Hoshino S. "Impact of Artificial "Gummy" Fingers on Fingerprint Systems". In: Proceedings of SPIE: Optical Security and Counterfeit Deterrence Techniques IV, pp. 275–289, January 2002.

[19] Putte T. and Keuning J. "Don't Get Your Fingers Burned". In: Proceedings of IFIP TC8/WG8.8 Fourth Working Conference on Smart Card Research and Advanced Applications, pp. 289–303, 2000.

[20] Lyengar S., Prasad L., and Min H. *Advances in Distributed Sensor Technology*.1995.

[21] Chen K., Wang L., and Chi H. "Methods of combining multiple classifiers with different features and their applications to text-independent speaker identification". International Journal of Pattern Recognition and Artificial Intelligence, Vol. 11, No. 3, pp. 417–445, 1997.

[22] Low C.Y., Andrew B.J., Tee C. "Fusion of LSB and DWT Biometric Watermarking for Offline Handwritten Signature", Congress on Image and Signal Processing, CISP '08, Volume: 5, Page(s): 702 – 708, 2008.

[23] Jin A.T.B., Ling D.N.C., Goh A., "BioHashing: two factor authentication featuring fingerprint data and tokenized random number", Pattern Recognition 37 (11) (2004) 2245–2255.

[24] Jin A.T.B., Ling D.N.C., "Cancelable biometrics featuring with tokenized random number", pattern Recognition Lett. 26 (10) (2005) 1454–1460.

[25] Jin A.T.B., Ling D.N.C., Goh A., "Personalized cryptographic key generation based on Face Hashing", Computer Security. J. 23 (7) (2004) 606–614.

[26] Connie T., Teoh A., Goh M., Ngo D., "Palm Hashing: a novel approach for dual factor authentication", Pattern Anal. Appl. 7 (3) (2004) 255–268.

[27] Pang Y.H., Jin A.T.B., Ling D.N.C., "Cancelable palmprint authentication system", Int. J. Signal Process. 1 (2) (2005) 98–104.

[28] Kong B., Cheung K., Zhang D, Kamel M., You J., "An analysis of BioHashing and its variants", Pattern Recognition, to appear.

[29] Lumini A. and Nanni L. "Empirical tests on BioHashing. Neuro Computing", 69(16) :2390–2395, October 2006.

[30] Andrew B.J. Teoh, Yip W.K., and Sangyoun L. "Cancellable biometrics and annotations on BioHash". Pattern recognition, 41:2034–2044, 2007.

[31] X. Tan, B. Triggs, \Fusing Gabor and LBP Feature Sets for Kernel-Based Face Recognition",Analysis and Modeling of Faces and Gestures, pp. 235-249, 2007.

[32] Datal N. Triggers B., "Histograms of oriented gradients for human detection", IEEE computer society on computer vision and pattern recognition, 2005.

[33] The Hong Kong Polytechnic University, PolyU MSP Database, http ://www.comp.polyu.edu.hk/sbiometrics/MultispectralPalmprint/MSP.htm.