UNIVERSITE MOHAMED SEDDIK BEN YAHIA - JIJEL

N° attribué par la bibliothèque

MÉMOIRE

pour obtenir le diplome de

MASTER

Spécialité : **Informatique légale et multimédia (ILM)**

Département de l'informatique

Université de Jijel

Réalisé par:

Chouaib LAIB et Abderrahim BENMESSAS

Titre:

Hachage perceptuel d'image d'empreintes digitales à base Shape Context

Encadré par : : **Wafa BIROUK**

Jury
. Rima BOUDJADJA,
. Karima ASSOUSSE,

Remerciements

Je tiens tout d'abord à remercier grandement mon encadreur Mademoiselle Wafa Birouk pour sa grande disponibilité et ses précieux conseils, idées et ses encouragement tout au lang de la rédaction de ce mémoire.

Je remerie également les membres de mon jury et tout les enseignants et les enseignantes du département de l'informatique.

Je voudrais remercier aussi toutes les personnes qui ont participé de près ou de loin à mes recherches et à l'élaboration de ce mémoire.

J'en profite pour remercier également tous ceux qui ont pu un moment ou un autre contribuer à ce mémoire, en particulier mes collègues.

Dédicace

Pour ma mère et mon père

Mes frères et mes soeurs

Mes amis

Résumé

La révolution numérique a fait fortement évoluer nos méthodes de communication et d'échange d'information avec l'explosion des réseaux de communication et les nouvelles avancées technologiques Cette évolution nous a permis d'échanger facilement et rapidement l'information sous toutes ses formes : textuelle, sonore et visuelle Ce développement nous a aussi permet d'effectuer des opérations aussi variées dans notre vie quotidienne pour cela le besoin et la nécessité de communication et de transfert sécurisé de l'information est énorme surtout dans notre immense monde numérique. Avec ces circonstances, il est devenu nécessaire de mettre en œuvre des systèmes et des outils performants adaptés aux ces menaces. En particulier, l'intégrité, la confidentialité et l'authenticité des images numériques du fait de la quantité importante de l'information visuelle qu'elles véhiculent. Les fonctions de hachage perceptuel sont fortement inspirées des fonctions de hachage cryptographique, elles se basent sur l'aspect visuel des données à hacher permettant d'établir une correspondance perceptuelle entre l'image originale et l'image à authentifier. Les manipulations acceptables (compression JPEG, bruit Gaussien,. . .) préservent l'aspect visuel de l'image à authentifier, par contre, les manipulations malicieuses (l'ajout de nouveaux objets, la suppression ou la modification majeure d'objets existants par exemple) changent le contenu sémantique de l'image. Ces dernières années ont vu beaucoup de chercheurs se pencher sur cette nouvelle approche de sécurité des données multimédia et les données d'empreintes digitales. Parmi les méthodes pour la construction de tel hache est la méthode à base Shape Context qui repose sur une décomposition circulaire de l'image d'empreinte digitale.

Abstract

The digital revolution has greatly changed our methods of communication and information exchange with the explosion of communication networks and new technological advances. This revolution allowed us to easily and quickly exchange information in all its forms : Text, sound and visual This development has also allowed us to carry out operations as varied in our daily life for that the need and the necessity of communication and secure transfer of the information is enormous especially in our immense digital world. Under these circumstances, it has become necessary to implement effective systems and tools adapted to these threats. In particular, the integrity, confidentiality and authenticity of digital images due to the large amount of visual information they convey. The perceptual hash functions are strongly inspired by the cryptographic hash functions and are based on the visual aspect of the data to be chopped making it possible to establish a perceptual correspondence between the original image and the image to be authenticated. Acceptable manipulations (JPEG compression, Gaussian noise, ...) preserve the visual aspect of the image to be authenticated, on the other hand, the malicious manipulations (the addition of new objects, the deletion or the major modification of existing objects for example) Change the semantic content of the image. In recent years have seen many researchers look into this new approach to security of multimedia data and fingerprint data. Among the methods for the construction of such an ax is the method based on Shape Context which relies on a circular decomposition of the fingerprint image.

Table des matières

Résumé iii
Abstract iv
Table des matières v

Introduction **1**

1 Généralités sur la biométrie **3**
1 Introduction 3
2 Qu'est-ce que la biométrie 3
2.1 Les différentes techniques biométriques 3
2.2 Les propriétés souhaités de la biométrie 5
2.3 La structure d'un système biométrique 6
2.3.1 Architecture d'un système biométrique 6
3 Les performances des systèmes biométrique 8
3.1 Erreurs de décision en vérification 9
3.2 Courbes de performances 9
3.3 Points de fonctionnement 10
4 Applications de la biométrie 12
5 Le marché de la biométrie 13
6 Conclusion 14

2 Système de reconnaissance d'empreinte digitale **17**
1 Introduction 17
2 L'empreinte digitale 18
2.1 Caractéristiques des empreintes digitales 18
2.2 Propriétés des images d'empreintes digitales 20
2.3 Systèmes de reconnaissance d'empreintes digitales 21
2.3.1 Principe général 21
2.3.2 Acquisition de l'empreinte 21
2.3.3 Le traitement de l'image et l'extraction de la signature 22
2.3.4 Le stockage et la phase d'appariement 23
2.4 Evaluation des systèmes d'authentication biométrique 24
2.4.1 Avantages 24
2.4.2 Inconvénients 24
3 Vulnérabilités et menaces 24
4 Les exigences d'un système de reconnaissance d'empreintes digitales . 26
5 Sécuriser le modèle biométrique des empreintes digitales 26
5.1 La cryptographie 26
5.2 Cryptosystème biométrique : Fuzzy Vault et Fuzzy commitment 28

5.3 Le tatouage 30
5.4 Hachage cryptographique 32
5.4.1 Critères de sécurité 32
6 Conclusion 33

3 Hachage perceptuel d'image d'empreintes digitales à base Shape Context **35**
1 Introduction 35
2 Hachage perceptuel des images 35
2.1 fonctions de hachage perceptuel 35
2.1.1 Les exigences attendues 36
2.2 Manipulations acceptables vs Manipulations malveillantes . . 36
2.3 Étapes d'un système de hachage perceptuel 37
2.4 Propriétés d'un système de hachage perceptuel 39
2.5 Classification des méthodes de hachage perceptuel 39
2.6 Robustesse et sécurité d'un système de hachage perceptuel . . 40
3 Hachage perceptuel d'image d'empreintes digitales à base Shape Context 40
3.1 Points caractéristiques robustes 42
3.2 Scale Invariant Feature transform(SIFT) 42
3.3 Détecteur de Harris 46
3.4 Hachage d'image d'emreintes digitales à base Shape Context . 49
3.4.1 Shape Context 49
3.4.2 Le hachage de shape context radiale (RSCH) : 50
4 Conclusion 54

4 Mise en Œuvre et Evaluation **55**
1 Introduction 55
2 Mise en Œuvre 55
2.1 Outils de développement 55
2.1.1 Matlab 55
2.1.2 Le serveur web Apache 56
2.1.3 Le SGBD MySql 56
2.2 Présentation de l'application 56
2.2.1 Menu principale 56
2.2.2 Hachage perceptuel à base shape context 57
2.2.3 Système de reconnaissance d'empreinte digitale . . . 59
2.2.4 Onglet Identification 60
2.2.5 Onglet Vérification 60
2.2.6 Onglet Paramètre 61
2.2.7 Onglet A propos 61
2.2.8 Onglet Fermeture 61
3 Tests et Evaluation 62
3.1 Présentation de la base de données 62
3.2 Mesures de performance 63
3.2.1 Protocole utilisé pour le test d'identification 63
4 Conclusion 65

Conclusion et perspectives **67**

Bibliographie 69

Table des figures 73

Liste des tableaux 75

Introduction

Aujourd'hui, les nouvelles avancées technologiques ont fait fortement évoluer nos méthodes de communication et d'échange d'information. Cette évolution explosive des nouvelles technologies d'information nous a permis d'échanger facilement et rapidement l'information sous toutes ses formes : textuelle, sonore et visuelle, sur des réseaux publics de plus en plus larges. Par exemple, les échanges sur Internet font désormais partie de notre quotidien, nous permettent d'effectuer des opérations aussi variées que l'achat auprès de boutiques en ligne, l'émission d'ordres bancaires ou encore le simple échange de données multimédia à contenu personnel. La profusion de ces applications rend alors le besoin de communications sécurisées de plus en plus pressant. Au regard des quelques exemples que nous venons dementionner, il apparaît clairement que l'information est un élément constitutif et déterminant dans tous les domaines et la nécessité de transfert sécurisé d'information est énorme surtout dans le monde actuel induit par le phénomène "Tout numérique".

Les fonctions de hachage perceptuel sont fortement inspirées des fonctions de hachage cryptographique, elles se basent sur l'aspect visuel des données à hacher permettant d'établir une correspondance perceptuelle entre l'image originale et l'image à authentifier. Les manipulations acceptables (compression JPEG, bruit Gaussien,...) préservent l'aspect visuel de l'image à authentifier, par contre, les manipulations malicieuses (l'ajout de nouveaux objets, la suppression ou la modification majeure d'objets existants par exemple) changent le contenu sémantique de l'image. Ces dernières années ont vu beaucoup de chercheurs se pencher sur cette nouvelle approche de sécurité des données multimédia et les données d'empreintes digitales. Parmi les méthodes pour la construction de tel hache est la méthode à base Shape Context qui repose sur une décomposition circulaire de l'image d'empreinte digitale.

Objectifs

- Rendre le hache perceptuel robuste contre les attaques acceptables.
- Lors de l'authentification ou l'identification, l'image modifiée par un malveillant doit avoir un hache différent de hache de l'image originale.
- résumant : La robustesse, la discriminante et la taille petite du hache.

Déroulement du travail :

- *Chapitre 1 :* Ce chapitre expose des généralités sur la biométrie et ses techniques ou modalités, ainsi que l'architecture d'un système biométrique et ses performances.
- *Chapitre 2 :* Dans ce chapitre, nous présentons le système de reconnaissance des empreintes digitales et ses vulnérabilités, en suite nous citons quelques techniques de sécurisation de tel système.
- *Chapitre 3 :* Ce chapitre présente une méthode de hachage perceptuel des images. Le but de notre approche est de satisfaire au mieux les deux propriétés

fondamentales des systèmes de hachage perceptuel, à savoir la robustesse et la sécurité. Dans cette méthode, nous proposons une transformation de l'image originale en une image plus robuste.

— *Chapitre 3 :* Ce chapitre présente la mise en Œuvre et l'évaluation de notre méthode.

Chapitre 1

Généralités sur la biométrie

1 Introduction

Depuis quelques décennies l'explosion de l'informatique et des réseaux de communication a fait augmenter de manière significative le besoin d'identification des personnes. Jusqu'à présent les méthodes usuelles d'identification sont basées sur ce que l'on possède (carte d'identité, carte à puce, badge magnétique,...) ou sur ce que l'on sait (mot de passe, code PIN,...) mais ces méthodes posent de gros problèmes de fiabilité (falsification de document, oubli de son code, décryptage du mot de passe via des logiciels spécifiques,...). Depuis les récents actes terroristes et les menaces qui pèsent sur de nombreux pays, une identification fiable des personnes est devenue un problème majeur pour des raisons de sécurité (contrôle aux frontières, accès aux lieux publics, transport,...). Tous ces problèmes ont ainsi provoqué un développement accru des techniques biométriques d'identification.

2 Qu'est-ce que la biométrie

La biométrie est une mesure des caractéristiques biologiques pour l'identification ou l'authentification d'un individu à partir de certaines de ses caractéristiques : *comportementales* (exemple de la dynamique de frappe au clavier), *physiques* ou *physiologiques* (exemple de l'ADN). Contrairement à ce que l'on sait ou ce que l'on possède la biométrie est basée sur ce que l'on est et permet ainsi d'éviter la duplication, le vol, l'oubli ou la perte. Les caractéristiques utilisées doivent être universelles (c'est-à-dire communes à tous les individus), uniques (pour pouvoir différencier deux individus) et permanentes (c'est-à-dire invariantes dans le temps pour chaque individu).

2.1 Les différentes techniques biométriques

Parmi les différentes techniques biométriques existantes on distingue trois catégories :

1. **L'analyse morphologique :** les empreintes digitales, l'iris de l'oeil, la forme de la main, les traits du visage, le réseau veineux de la rétine.
2. **L'analyse des traces biologiques :** l'ADN, le sang, la salive, l'urine, l'odeur.

3. **L'analyse comportementale :** la reconnaissance vocale, la dynamique de frappe au clavier, la dynamique de signature, la manière de marcher.

Parmi les nombreuses méthodes d'identification biométrique l'utilisation des empreintes digitales est la méthode la plus aboutie de parts du marché, nous y reviendrons plus en détail dans le chapitre suivant. Néanmoins d'autres méthodes commencent à trouver leur place sur le marché de la biométrie :

- **La forme du visage :** [4,5] arrive en deuxième position de parts de marché. Plusieurs parties du visage (joues, yeux, nez, bouche) sont extraites d'une photo ou d'une vidéo et analysées géométriquement (distances entre différents points, positions, formes, ...). Le problème de cette méthode vient des possibles perturbations pouvant transformer le visage (maquillage, faible luminosité, présence d'une barbe ou de lunettes, expression faciale inhabituelle, changement avec l'âge,...)
- **La géométrie de la main :** [6] jusqu'à 90 caractéristiques de la main sont mesurées (forme de la main et des articulations, longueur et largeur des doigts, longueurs inter articulations). Le taux d'erreurs dans la reconnaissance est assez élevé, en particulier pour des personnes appartenant à une même famille en raison d'une forte ressemblance. De plus la forme de la main évolue beaucoup avec l'âge.
- **L'iris :** [7,8] est une technique extrêmement fiable car l'iris contient une infinité de points caractéristiques (ensemble fractal), la fraude étant néanmoins possible en utilisant des lentilles. L'acquisition de l'iris est effectuée au moyen d'une caméra pour pallier aux mouvements inévitables de la pupille. Elle est très sensible (précision, reflet) et relativement désagréable pour l'utilisateur car l'oeil doit rester grand ouvert et il est éclairée par une source lumineuse pour assurer un contraste correct.
- **La reconnaissance vocale :** [9, 10] les caractéristiques du timbre de la voix et de la prononciation sont analysées. La qualité de l'enregistrement peut poser problème et il est possible de frauder avec un échantillon vocal préenregistré.
- **La dynamique du tracé de la signature :** [11] Il s'agit d'une analyse comportementale ou différents éléments (mesure de la vitesse, ordre d'écriture, pression exercée, accélérations) sont mesurés lors de la signature. La falsification est possible en passant par une phase d'apprentissage, la signature peut varier selon le stress de l'utilisateur.

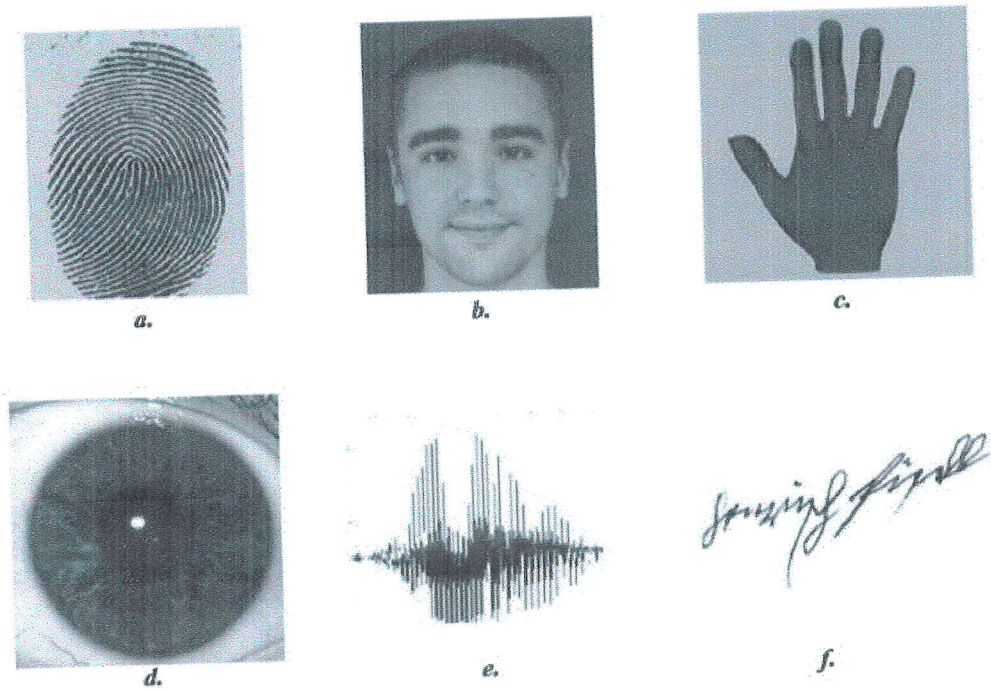


FIGURE 1.1 – Exemples de différentes caractéristiques biométriques : empreinte digitale(a), visage(b), main(c), iris(d), empreinte vocale(e), signature(f).

Certaines de ces techniques sont très prometteuses (iris) et commencent juste à émerger, d'autres sont encore au stade expérimental (analyse comportementale). Mais l'utilisation des empreintes digitales reste la méthode la plus aboutie actuellement.

2.2 Les propriétés souhaités de la biométrie

De manière générale, pour q'une caractéristique biologique, physique ou comportementale puisse être utilisée comme une modalité biométrique pertinente, elle doit répondre aux critère suivantes :

- **Universalité :** Elle doit être possédée par tous les individus.
- **Unicité :** Elle doit être différente pour chaque individu.
- **Permanence :** Elle doit être stable et invariante dans le temps.
- **Collectabilité :** Elle doit être facile à collecter et facilement quantifiable.
- **Performance :** Elle doit pouvoir assurer des bonnes performances en authentification.
- **Acceptabilité :** Elle doit pouvoir être acceptée comme modalité biométrique par les utilisateurs.

A ces critères, nous rajoutons celui de la **Sécurité**, dans le sens ou la caractéristique biologique pertinente pour la biométrie devrait etre idéalement infalsifiable (Impossible à détourner, à voler, à copier,...).

Cependant, aucune des modalités biométriques susmentionées ne répond complètement à toutes ces éxigences, un compromis est donc fait entre tous ces critères pour choisir la modalité la plus adaptée à l'application visée. En particulier, nous verrons

dans ce travail que le critère de la **sécurité** est encore un point très problématique, pouvant d'ailleurs limiter son *accéptation*

2.3 La structure d'un système biométrique

En général un système biométrique est un système automatique de mesure basé sur la reconnaissance de caractéristiques propres à un individu : physique ou comportementale. Il est basé sur l'analyse de données liées à l'individu qui peuvent être classées en trois grandes catégories : analyse basée sur la morphologie, analyse de traces biologiques, l'analyse comportementale. Il peut être représenté par quatre modules principaux :

1. Le module de capture est responsable de l'acquisition des données biométriques d'un individu (cela peut être un appareil photo, un lecteur d'empreintes digitales, une caméra de sécurité, etc.)

2. Le module d'extraction de caractéristiques prend en entrée les données biométriques acquises par le module de capture et extrait seulement l'information pertinente afin de former une nouvelle représentation des données. Généralement, cette nouvelle représentation est censée être unique pour chaque personne et relativement invariante.

3. Le module de correspondance compare l'ensemble des caractéristiques extraites avec le modèle enregistré dans la base de données du système et détermine le degré de similitude (ou de divergence) entre les deux.

4. Le module de décision vérifie l'identité affirmée par un utilisateur ou détermine l'identité d'une personne basée sur le degré de similitude entre les caractéristiques extraites et le(s) modèle(s) stocké(s).

2.3.1 Architecture d'un système biométrique

Bien que sous formes diverses et mesurant des caractéristiques différentes, les systèmes biométriques partagent tous la même architecture. Ce sont tous des systèmes de reconnaissance de formes. Ils sont composés d'un ou plusieurs systèmes d'acquisition qui vont mesurer le ou les traits physiques ou comportementaux de l'individu. Lorsque le système utilise plusieurs traits qu'il lie à un individu, on parle de systèmes multimodaux. A l'inverse, si une seule caractéristique est utilisée, on utilise le terme de systèmes uni-modaux. Un système d'information extrait, encode, stocke et compare ces données.

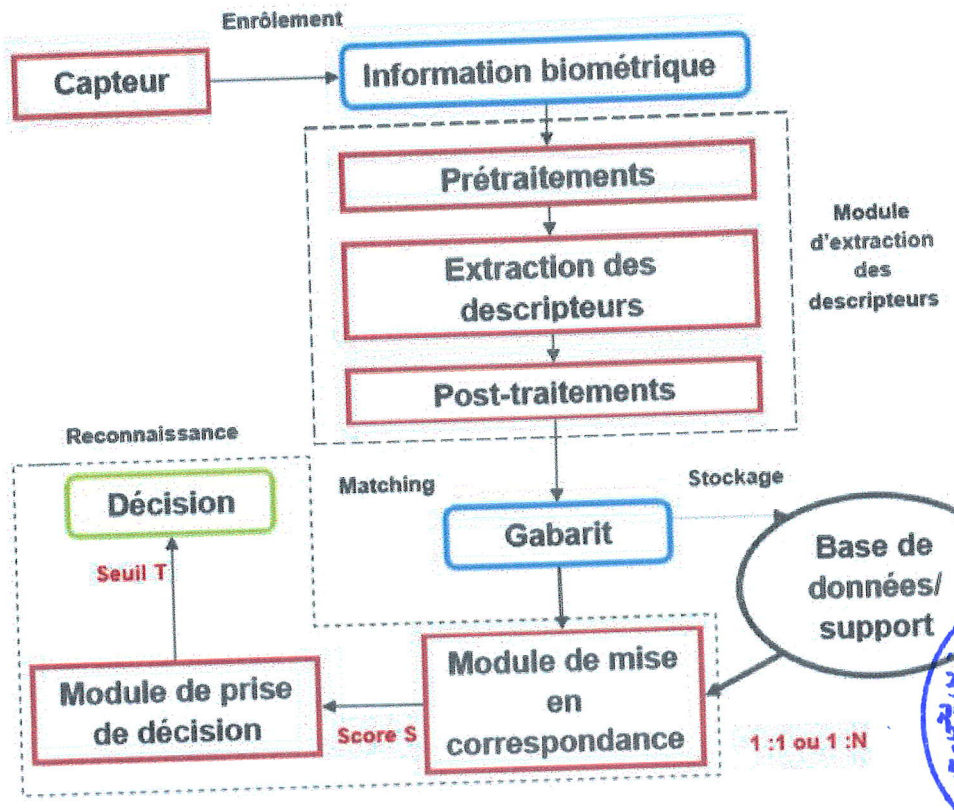


FIGURE 1.2 – Architecture d'un système biométrique

Les systèmes biométriques sont tous automatisés permettant un traitement rapide de l'information. Que le système soit utilisé en mode authentification ou identification, deux étapes sont nécessaires avant que celui-ci soit pleinement opérationnel :

- ***Apprentissage (Enrôlement)*** **:** C'est la phase initiale qui s'effectue une seule fois. Au cours de l'apprentissage, on fait une capture (acquisition) de la caractéristique biométrique. En général cette capture n'est pas directement stockée, des transformations lui sont appliquées, le modèle étant une représentation compacte du signal permettant de faciliter la phase de reconnaissance et de diminuer la quantité des données à stocker. Les tâches essentielles de cette phase sont : l'extraction de paramètres et la construction d'un modèle de représentation (appelé gabarit).
- ***Reconnaissance*** **:** Opération se déroulant à chaque fois qu'une personne se présente devant le système, elle consiste en l'extraction d'un ensemble de paramètres comme pour l'étape d'apprentissage suivie d'une autre étape de comparaison et de prise de décision selon le mode opératoire du système :

  1. ***Identification*** : (à partir de l'échantillon biométrique recherche du gabarit correspondant), elle permet d'établir l'identité d'une personne (qui suis-je ?) à partir d'une base de données, il s'agit d'une comparaison du type "un contre plusieurs".

  2. ***Vérification*** : (échantillon biométrique correspond bien au gabarit), également appelée authentification(contrôle), consiste à confirmer ou infirmer l'identité d'une personne (suis-je celui que je prétend être ?). Il s'agit d'une

comparaison du type "un contre un", les caractéristiques de l'individu sont comparées à celles présentes dans un enregistrement de référence.

3 Les performances des systèmes biométrique

Dans un système biométrique, nous avons vu que l'information biométrique obtenue à partir du capteur lors de l'enrôlement(apprentissage) est comparée à une information déjà existante, sauvegardée au préalable. Un score de similarité (ou de dissimilarité) permettant de quantifier la ressemblance (ou la dissemblance) entre les deux informations biométriques est alors calculé entre elles.

Pour évaluer les performances d'un système biométrique dans un mode de vérification, un grand nombre de comparaison est effectué sur une base de données de test. Nous nous plaçons dans le cas où les différents échantillons d'un même individu sont considérés comme indépendants. Chaque échantillon biométrique de chaque individu de la base est alors comparé à tous les autres échantillons de la base de données. Quand les deux échantillons comparés proviennent du même individu, la comparaison est appelée « comparaison client ». Dans le cas où les deux échantillons proviennent d'individus différents, la comparaison est appelée « comparaison imposteur ». Les densités de scores pour les comparaisons client et imposteur sont générées à partir de la base de données entière.

La précision du système biométrique est alors évaluée par la capacité à séparer ces deux densités. Cette séparation se fait grâce au seuil à partir duquel la décision d'acceptation ou de rejet d'identité sera prise. Quand le score est supérieur au seuil, la décision est considérée comme positive et l'identité de l'individu est acceptée. Dans le cas contraire, la décision est considérée comme négative et l'identité de l'individu est rejetée.

Pour un système biométrique idéal, les deux densités de scores ne se recouvrent pas. Cependant, pour un système biométrique réel, les deux densités se recouvrent et aucune valeur de seuil ne permet de les séparer complètement. Ce recouvrement traduit alors des erreurs de décision, en particulier des fausses acceptations (dans le cas où une comparaison imposteur retourne un score de similarité élevé) et des faux rejets (dans le cas où une comparaison client retourne un score très bas).

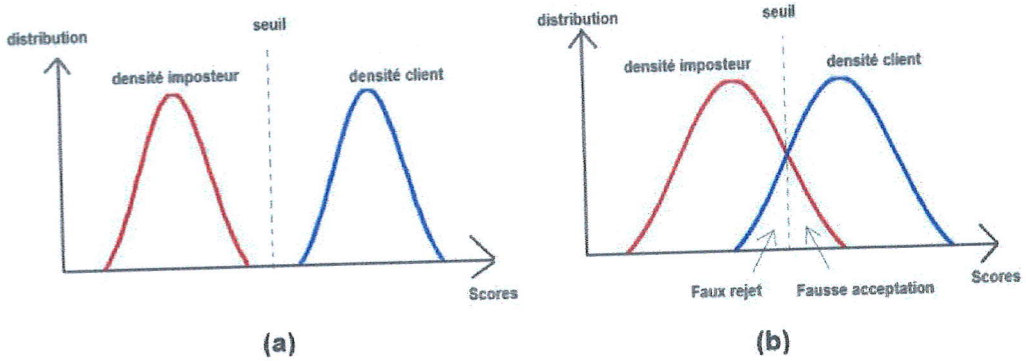


FIGURE 1.3 – (a) Densités de scores (de similarité) pour un système biométrique idéal. (b) Densités de scores pour un système biométrique réel.

3.1 Erreurs de décision en vérification

Nous avons vu que deux erreurs étaient possibles dans la décision de vérification : une fausse acceptation quand un imposteur est accepté et un faux rejet quand un client est rejeté alors qu'il ne le devrait pas. A partir de ces deux erreurs, on peut définir les taux d'erreurs suivants : [24]

- **False Acceptance Rate (FAR) :** C'est la probabilité qu'un imposteur soit considéré comme client. Il est égal au nombre de fausses acceptations divisé par le nombre d'imposteur dans la base :

$$FAR = \frac{FA}{NI}$$

- **False Rejection Rate (FRR) :** C'est la probabilité qu'un client soit considéré comme imposteur. Il est égale au nombre de faux rejets divisé par le nombre de client dans la base.

$$FRR = \frac{FR}{NC}$$

Dans certains cas, il est possible de les estimer en fonction de la taille de la base de données, à partir de lois empiriques par exemple. On peut également les évaluer à partir de modèles statistiques.

3.2 Courbes de performances

Nous voyons d'après la Figure 1.3 et les équations précédentes que le FAR et le FRR dépendent du seuil de décision. En particulier, nous pouvons voir que l'amélioration du FAR en déplaçant le seuil se fera au détriment du FRR, et vice versa. La valeur de ce seuil est donc un paramètre crucial des systèmes biométriques.

Ainsi, les performances sont toujours évaluées en considérant différentes valeurs de seuil, afin de couvrir le plus d'applications différentes. En fait, les systèmes biométriques sont souvent repérer sur des bases de données pour évaluer notamment le FAR espéré pour un certain niveau de FRR donné.

Pour visualiser les performances, on utilise en règle générale les courbes de performances. Elles permettent de représenter les performances pour toutes les valeurs de seuil considérées. En particulier, on peut représenter les variations du FAR et du FRR en fonction de la valeur du seuil. On peut également représenter la variation du FRR en fonction du FAR lorsque le seuil varie. On obtient ainsi une courbe ROC (Receiver Operating Characteristics) [24], comme représenté en Figure 1.4 (a). En représentant les deux axes des taux d'erreurs en échelle logarithmique, on obtient de manière équivalente une courbe DET (Detection Error Tradeoff),comme représenté en Figure 1.4 (b). souvent préféré à la courbe ROC.

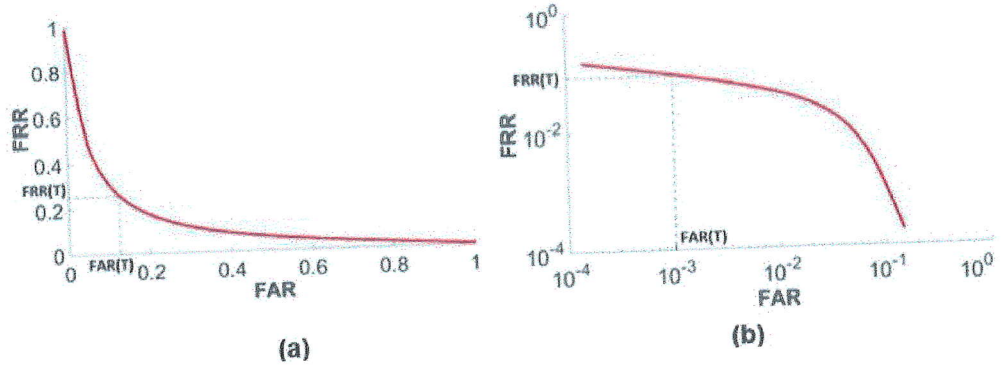


FIGURE 1.4 – (a) Courbe ROC. (b) Courbes DET.

3.3 Points de fonctionnement

Un point de fonctionnement est un indicateur de performances. En fonction de l'application considérée, plusieurs points de fonctionnement peuvent être considérés. [24] Par exemple(figure 1.5) :

1. **Equal Error Rate (EER), ou taux d'erreurs égales :** Ce point de fonctionnement correspond au seuil où FAR=FRR.

2. **FRR à FAR fixé :** Ce point de fonctionnement correspond au seuil pour lequel la valeur du FAR est égale à une certaine valeur désirée. Les performances sont exprimées par le FRR associé à ce FAR particulier. Pour les applications où la sécurité, au sens de l'application biométrique, et pas au sens de la sécurité même de la biométrie (antispoofing) est primordiale, le FAR est choisi très bas, au détriment du FRR donc.

3. **FAR à FRR fixé :** Ce point de fonctionnement correspond au seuil pour lequel la valeur du FRR est égale à une certaine valeur désirée. Les performances sont exprimées par le FAR associé à ce FRR particulier. En général on choisit un FRR bas pour les applications où l'efficacité prime sur la sécurité (applications civiles, comme par exemple l'authentification sur un ordinateur personnel), ce qui est traduit par le terme commodité.

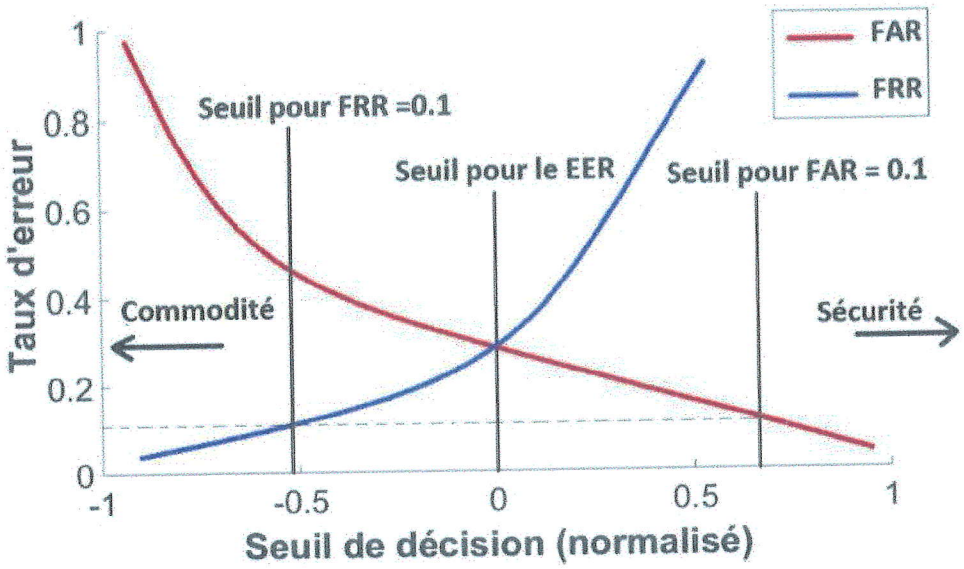


FIGURE 1.5 – Illustration des points de fonctionnement sur une courbe des taux d'erreurs en fonction du seuil de décision.

En général, les performances d'un système biométrique sont données par la mesure de deux taux d'erreurs : le FRR (False Rejet Rate ou Taux de Faux Rejet) et le FAR (False Acceptation Rate ou Taux de Fausse Acceptation). Le FRR ou le TFR (Taux de Faux Rejets) : estime le pourcentage d'utilisateurs valides qui ne seront pas reconnus par le système. Le FAR ou le TFA (taux de fausse acceptation) : estime le pourcentage d'utilisateurs non connus qui seront faussement reconnus par le système.

Le paramétrage d'un système consiste à trouver le bon équilibre entre ces deux taux, le FAR augmentant lorsque le FRR diminue, et inversement. Un contrôle d'accès très sécurisé aura un FAR très bas, pour garantir qu'aucune personne non autorisée n'accède au site, mais, en contrepartie le FRR sera élevé, ce qui signifie que des utilisateurs valides se verront refuser l'accès. Les autres mesures de performance sont les temps d'encodage de l'empreinte et de mise en correspondance. Là encore, ces valeurs peuvent varier considérablement d'une application à une autre. Un troisième paramètre FER (False Equal Rate) mesure le taux d'échec à l'enrôlement. Il traduit la probabilité d'absence d'une caractéristique biométrique pour un individu dans une population, donne un point sur lequel le TFA. est égal au TFR. La figure 1.6 illustre le FRR et le FAR à partir de distributions des scores authentiques et imposteurs.

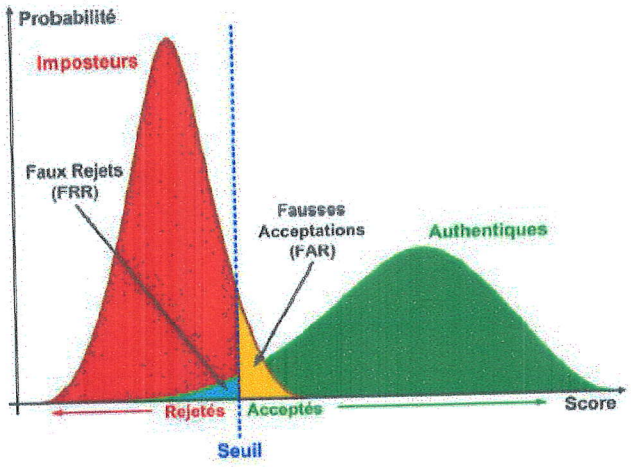


FIGURE 1.6 – Illustration du FRR et du FAR. [1]

4 Applications de la biométrie

Aujourd'hui, les principales applications sont la production de titres d'identité, le contrôle d'accès à des sites sécurisés, le contrôle des frontières, l'accès aux réseaux, systèmes d'information et stations de travail, le paiement électronique, la signature électronique et même le chiffrement de données. Cette liste n'est pas exhaustive, et de nouvelles applications vont très certainement voir rapidement le jour.

Les techniques biométriques sont appliquées dans plusieurs domaines et leur champ d'application couvre potentiellement tous les domaines de la sécurité où il est nécessaire de connaître l'identité des personnes. Les applications peuvent être divisées en trois groupes principaux :

1. ***Application commerciales :*** telles que l'accès au réseau informatique, la sécurité de données électroniques, le commerce électronique, l'accès d'internet, l'ATM, la carte de crédit, le contrôle d'accès physique, le téléphone portable, le PDA, la gestion des registres médicales, l'étude de distances, etc.. . .
2. ***Applications de gouvernement :*** telles que la carte nationale d'identifications, le permis de conduite, la sécurité sociale, le contrôle de passeport, etc.. . .
3. ***Applications juridiques :*** telles que l'identification de cadavre, la recherche criminelle, l'identification de terroriste, les enfants disparus, etc.
4. ***Les applications de la biométrie :***
   - *Contrôle d'accès aux locaux :* - Salles informatiques. - Sites sensibles (service de recherche, site nucléaire).
   - *Equipements de communication :* - Terminaux d'accès. - Téléphones portables.
   - *Systèmes d'informations :* - Lancement du système d'exploitation, - Accès au réseau. - Transaction (financière pour les banques, données entre entreprises).
   - *Machines Equipements divers :* - Distributeur automatique de billets. - Lieu sensible (club de tir, police). - Contrôle des adhérents dans les clubs privés. - Contrôle des temps de présence.

- ***Etat/Administration :*** - Fichier judiciaire. - Services sociaux (sécurisation des règlements). - Système de vote électronique.

5 Le marché de la biométrie

La biométrie connaît un engouement sans précédent. La croissance mondiale de la biométrie depuis quelques années est incontestable, tant le nombre d'intervenants est grand, même s'il existe peu d'informations publiques concernant ce marché. On peut toutefois considérer certaines données et certains chiffres sur son évolution au fil des années, tant à l'échelle mondiale, qu'américaine ou européenne.

Le marché de la sécurité informatique est encore atomisé, peu de fournisseurs peuvent prétendre offrir une gamme complète de produits. Les spécialistes estiment que ce marché est en pleine croissance et qu'il va également se concentrer. L'Internet et le commerce électronique sont des marchés porteurs pour la sécurité, mais ils ne sont pas les seuls. Le télétravail, la mise à dispositions d'informations aux clients et sous traitants sont également des facteurs de risque pour les entreprises qui ouvrent leur système d'informations.

Le besoin grandissant de sécurité sur les terminaux mobiles a été mis en exergue par une enquête récente, publiée par Toshiba. Celle-ci soutient que 90 des cadres dirigeants et chefs d'entreprise européens stockent des données sensibles, voire confidentielles sur leur outil de communication et parmi eux, 22 admettent avoir pourtant déjà perdu cet outil.

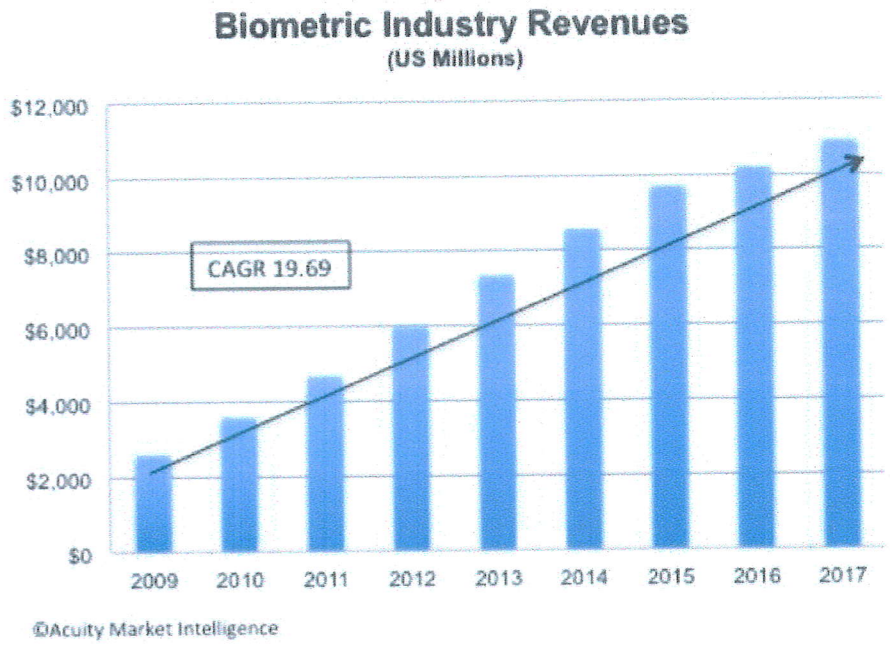


FIGURE 1.7 – Evolution du marché international de la biométrie.

IBG (International Biometric Group) édite régulière analyse complète des chiffres d'affaires, des tendances de croissance, et des développements industriels pour le marché de la biométrie actuel et futur. La lecture de ce rapport est essenti inves-

tisseurs dans les entreprises biométriques, ou les développeurs de solutions biométriques. [3]

On s'attend à ce que le chiffre d'affaires de l'industrie biométrique incluant les applications judiciaires et celles du secteur public, se développe rapidement. Une grande partie de la croissance sera attribuable aux controle d'accès aux systèmes d'informations(ordinateur/réseau) et au commerce électronique, bien que les applications du secteur public continuent à être une partie essnetielle de l'industrie.

On prévoit que le chiffre d'affaires des marchés émergents (accès aux systèmes d'information, commerce électronique et téléphonie, accès physique, et surveillance) dépasse le chiffre d'affaires des secteurs plus matures (identification criminellle et identification des citoyens).

- On s'attend à ce que l'empreinte digitale gagne 31% du marché de biométrie, suivi de l'identification de visage à 15%.
- On projette que les revenus annuels de l'identification de l'iris excèdent $250m d'ici.
- On s'attend à ce que l'Asie et l'Amérique du Nord soient les plus grands marchés globaux pour les produits biométriques et les services.
- Les systèmes Multi biométriques émergeront pour comporter approximativement 5% de tout le marché de la biométrie.
- Les empreintes digitales continuent à être la principale technologie biométrique en termes de part de marché, près de 32% du chiffre d'affaires total (hors applications judiciaires). La reconnaissance du visage, avec 15% du marché (hors applications judiciaires), dépasse la reconnaissance de la main, qui avait avant la deuxième place en terme de source de revenus après les empreintes digitales. [3]

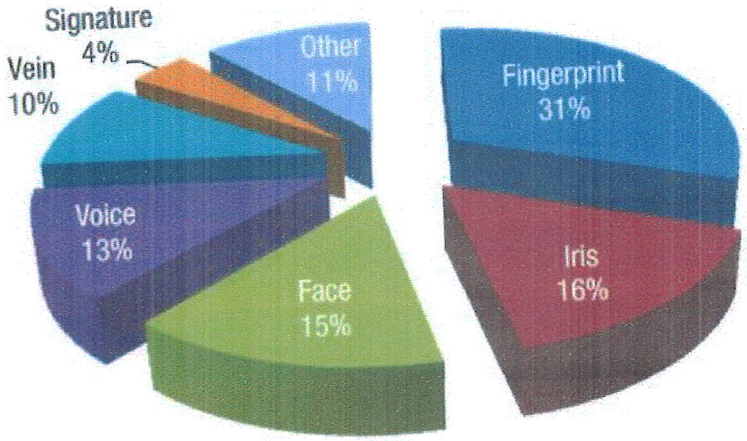


FIGURE 1.8 – Part de marché des differentes méthodes biométriques.

6 Conclusion

Dans ce chapitre nous avons décrit les technologies utilisées dans les systèmes biométriques pour l'identification des personnes , leurs architectures et leurs différentes applications, ainsi nous avons donné un aperçu sur les techniques de mesure

des performances des systèmes biométriques et montré les différentes modalités biométriques tout en soulignant les avantages et les inconvénients de chacune. Nous avons constaté aussi que les performances des systèmes biométriques dépendent de plusieurs facteurs et qu'elles varient d'un système à un autre.

Parmi les modalités utilisées dans la reconnaissance biométrique, nous avons trouvé que les minuties de l'empreinte digitale sont les traits les plus intéressants à cause de leurs précisions et leurs stabilités. De mème l'utilisation de l'empreinte digitale suscite de plus en plus l'intérêt de la communauté scientifique car elle présente plusieurs challenges et verrous technologiques. Dans le chapitre suivant nous allons présenter le système de reconnaissance d'empreintes digitales et quelques techniques de sécurisation déja existées pour ce système.

Chapitre 2

Système de reconnaissance d'empreinte digitale

1 Introduction

L'empreinte digitale est l'une des techniques les plus connues du grand public, elle est centenaire. C'est grâce aux travaux d'Alphonse Bertillon, dans les années 1880, que l'on a commencé à pouvoir identifier des récidivistes sans avoir recours au marquage ou à la mutilation. L'idée d'en faire un instrument d'identification à part entière s'est imposée avec les recherches du Britannique Galton, qui démontra la permanence du dessin de la naissance à la mort, son inaltérabilité et son individualité.

La donnée de base dans le cas des empreintes digitales est le dessin représenté par les crêtes et sillons de l'épiderme. Ce dessin est unique et différent pour chaque individu. En pratique, il est quasiment impossible d'utiliser toutes les informations fournies par ce dessin (car trop nombreuses pour chaque individu), on préférera donc en extraire les caractéristiques principales telles que les bifurcations de crêtes, les "îles", les lignes qui disparaissent, etc. Une empreinte complète contient en moyenne une centaine de ces points caractéristiques (les "minuties"). Si l'on considère la zone réellement scannée, on peut extraire environ 40 de ces points. Pourtant, là encore, les produits proposés sur le marché ne se basent que sur une quinzaine de ces points (12 au minimum vis-à-vis de la loi), voire moins pour beaucoup d'entre eux (jusqu'à 8 minimum). Pour l'histoire, le nombre 12 provient de la règle des 12 points selon laquelle il est statistiquement impossible de trouver 2 individus présentant les mêmes 12 points caractéristiques, même en considérant une population de plusieurs dizaines de millions de personnes.

Les techniques utilisées pour la mesure sont diverses : capteurs optiques (caméras CCD/CMOS), capteurs ultrasoniques, capteurs de champ électrique, de capacité, de température...

Ces capteurs sont souvent doublés d'une mesure visant à établir la validité de l'échantillon soumis (autrement dit, qu'il s'agit bien d'un doigt) : mesure de la constante diélectrique relative de l'échantillon, sa conductivité, les battements de cœur, la pression sanguine, voire une mesure de l'empreinte sous l'épiderme...

2 L'empreinte digitale

Une empreinte digitale est le dessin formé par les lignes de la peau des doigts, des paumes des mains, des orteils ou de la plante des pieds. Ce dessin se forme durant la période fœtale. Il existe deux types d'empreintes : l'empreinte directe (qui laisse une marque visible) et l'empreinte latente (saleté, sueur ou autre résidu déposé sur un objet).

Elles sont uniques et immuables, elles ne se modifient donc pas au cours du temps (sauf par accident comme une brûlure par exemple). La probabilité de trouver deux empreintes digitales similaires est de 1 sur 1024. Les jumeaux, par exemple, venant de la même cellule, auront des empreintes très proches, mais pas semblables. On classe les empreintes selon un système vieux d'une décennie : le système Henry. Dans ce système, le classement repose sur la topographie générale de l'empreinte digitale et permet de définir des familles telles que les boucles, les arches et les tourbillons.

Elles sont composées, de terminaisons en crêtes, soit le point où la crête s'arrête, et de bifurcations, soit le point où la crête se divise en deux. Le noyau est le point intérieur, situé en général au milieu de l'empreinte. Il sert souvent de point de repère pour situer les autres minuties. D'autres termes sont également rencontrés : le lac, l'île, le delta, la vallée, la fin de ligne. Ces caractéristiques peuvent être numérisées. Une empreinte complète contient en moyenne une centaine de points caractéristiques, mais les contrôles ne sont effectués qu'à partir de 12 points. Statistiquement, il est impossible de trouver 2 individus présentant 12 points caractéristiques identiques, même dans une population de plusieurs millions de personnes.

2.1 Caractéristiques des empreintes digitales

Une empreinte digitale est constituée d'un ensemble de lignes localement parallèles formant un motif unique pour chaque individu, on distingue les stries (ou crêtes, ce sont les lignes en contact avec une surface au toucher) et les sillons (ce sont les creux entre deux stries). Les stries contiennent en leur centre un ensemble de pores régulièrement espacés.

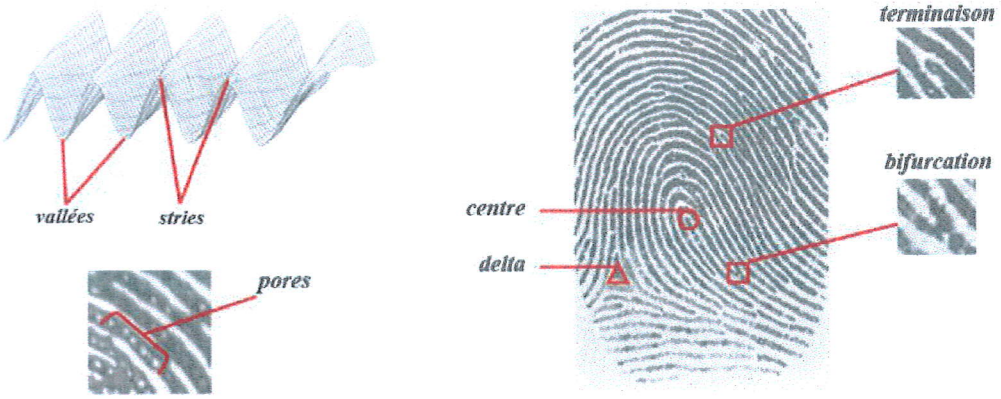
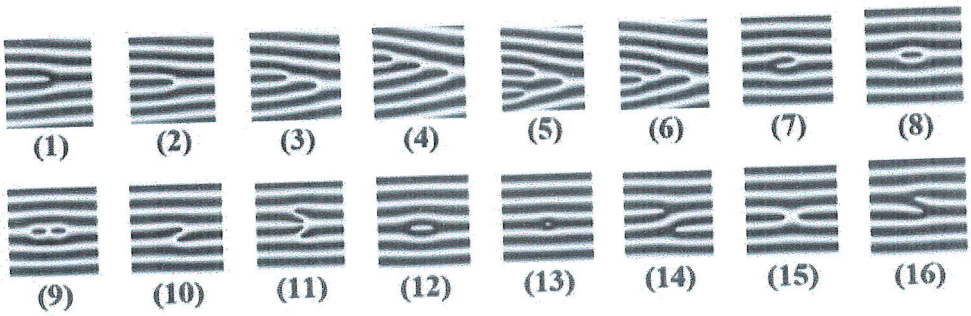


FIGURE 2.1 – Caractéristique d'une empreinte digitale

Chaque empreinte possède un ensemble de points singuliers globaux (les centres et les deltas) et locaux (les minuties). Les centres correspondent à des lieux de

convergences des stries tandis que les deltas correspondent à des lieux de divergence. Une étude a monté l'existence de seize types de minuties différentes mais en général les algorithmes ne s'intéressent qu'aux bifurcations et terminaisons qui permettent d'obtenir les autres types par combinaison.



1.	terminaison	9.	boucle double
2.	bifurcation simple	10.	pont simple
3.	bifurcation double	11.	pont jumeau
4.	bifurcation triple I	12.	intervalle
5.	bifurcation triple II	13.	point isolé
6.	bifurcation triple III	14.	traversée
7.	crochet	15.	croisement
8.	boucle simple	16.	tête bêche

FIGURE 2.2 – Les différents types de munitie

La position et le nombre de centres et de deltas permettent de classifier les empreintes en catégorie selon leur motif général, on distingue principalement trois grandes familles :

- Les boucles (loop) représentent 65% des empreintes rencontrées.
- Les spires (whorl) représentent 30% des empreintes rencontrées.
- Les arches (arch) représentent 5% des empreintes rencontrées.

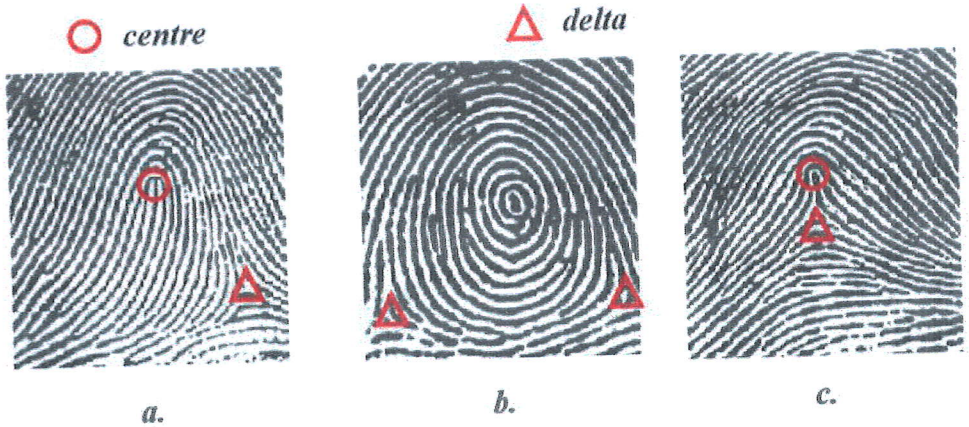


FIGURE 2.3 – Les trois principales classes d'empreinte : boucle(a), spire(b), arche(c).

L'ensemble formé par la disposition des points singuliers constitue un motif unique pour chaque individu, en effet il a été montré [12] que l'empreinte digitale se forme au cours du troisième mois de la vie, le motif général est influencé par les gènes héréditaires mais l'apparition des détails (minuties) est créée de manière accidentelle par des pressions variables aléatoires sur les surfaces tactiles. Ainsi l'empreinte est unique pour tout individu, y compris pour des vrais jumeaux et il a été montré que les méthodes de reconnaissance actuelles permettent d'identifier efficacement les jumeaux [23]. De plus les empreintes une fois formées ne changent plus au cours de la vie d'une personne, ces deux caractéristiques en font un moyen de reconnaissance très efficace.

2.2 Propriétés des images d'empreintes digitales

Les empreintes digitales constituent de loin la modalité biométrique la plus employée actuellement, elle représente environ 31% du part du marché de la biométrie (suivi du visage l'iris ....) cette usage prononcé des empreintes digitales peut s'expliquer par la longue histoire de leur utilisation pour les applications judiciaire. De plus, par apport à d'autres modalités, les empreintes digitales sont relativement facile à acquirir. C'est également l'une des modalités biométriques les plus performantes, et son accéptation parmi les individus est plutot bonne. Les propriétés des empreintes digitales en font une modalité biométrique particulièrement pertiente, ces propriétés sont notamment :

- Leur unicité et leur permanence, intimement liés à leur genèse.
- Leur universalité.
- La richesse de l'information qu'elle contiennent, exploitable pour la reconnaissance durant l'étape de matching.
- Leur acquisition qui reste relativement facile.

2.3 Systèmes de reconnaissance d'empreintes digitales

2.3.1 Principe général

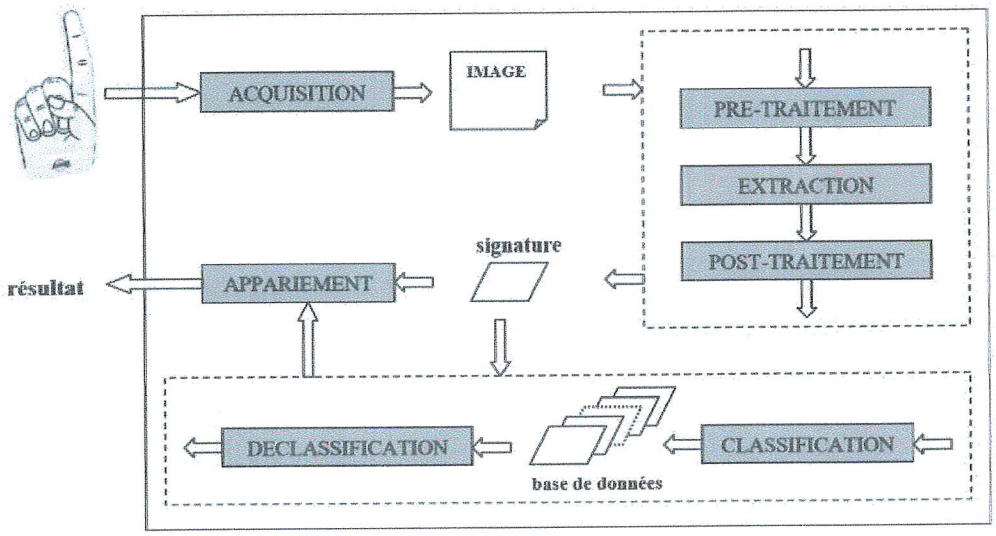


FIGURE 2.4 – Architecture d'un système biométrique basé sur l'empreinte

Un système automatique complet de reconnaissance d'empreintes digitales est une chaine de processus qui à partir du doigt d'un utilisateur en entrée renvoie un résultat en sortie, permettant ainsi à l'utilisateur d'accéder ou non à des éléments nécessitant une protection. La réalisation d'un tel système a fait l'objet de très nombreuses recherches et des méthodes très différentes de traitement ont été proposées [13]. Néanmoins ces systèmes répondent toujours à la même structure. La première phase permet d'obtenir une image de l'empreinte de l'utilisateur (acquisition), laquelle va subir un prétraitement pour extraire l'information utile de l'image (signature) suivi éventuellement d'un traitement supplémentaire permettant d'éliminer de possibles fausses informations qui se seraient glissées entre temps dans la chaine de traitement. Ensuite si l'utilisation du système consiste juste à créer une base de données (stockage) la signature est éventuellement compressée puis stockée dans la base de données au moyen d'une technique d'archivage (classification). Pour un système d'identification l'ensemble des empreintes présentes dans la base de données pouvant correspondre à celle de l'utilisateur (modèle identique) sont désarchivées et comparées (appariement) une à une avec celle de l'utilisateur, si une éventuelle correspondance est trouvée des informations personnelles concernant l'utilisateur sont renvoyées par le système. Dans le cas d'un système de vérification il n'y a qu'une seule comparaison et un résultat binaire est renvoyé, permettant l'acceptation ou le rejet de l'utilisateur.

2.3.2 Acquisition de l'empreinte

La première phase d un système de reconnaissance consiste à obtenir une image de l'empreinte du doigt. Longtemps le seul moyen existant a été l'utilisation du papier et de l'encre ce qui a rendu la têche de reconnaissance très lourde. En effet la

qualité de l'image était plutôt mauvaise (plusieurs acquisitions étaient nécessaires) et l'extraction de la signature était effectuée visuellement par un expert (processus très long et pénible). Heureusement avec le développement de l'informatique et de la microélectronique de nouveaux moyens d'acquisition ont fait leur apparition, permettant ainsi d'accélérer la chaine de traitement en l'automatisant (un capteur dédié fournit directement une image numérique).

Dans le cas d'enquêtes criminelles les empreintes sont majoritairement utilisées. En effet un doigt contaminé par une substance telle que du sang, de l'encre ou de la graisse va laisser une trace au contact d'un support solide. Un doigt propre va également laisser des traces provoquées par les sécrétions naturelles des glandes de la peau. Dans un lieu ou s'est produit un crime les enquêteurs privilégient donc la recherche d'empreintes éventuellement laissées par le malfaiteur. Ce type d'empreintes est dénommé empreinte latente. Elles sont en général détectées via l'utilisation de vapeur de superglue, puis elles sont photographiées à haute résolution pour permettre l'automatisation de l'extraction de la signature. La très mauvaise qualité de ce genre d'empreintes peut nécessiter la confirmation visuelle d'un expert.

2.3.3 Le traitement de l'image et l'extraction de la signature

Lors de l'acquisition de l'empreinte l'image obtenue contient souvent beaucoup de changement ayant des origines diverses :

- Les substances parasites présentes sur le doigt (encre, graisse, saletés, ...).
- La personne (cicatrices, métiers manuels, âge, ...).
- L'environnement ou se produit l'acquisition (température de l'air, degré d'humidité,...).
- Les caractéristiques spécifiques du moyen d'acquisition utilisé.

Pour permettre une reconnaissance fiable un pretraitement est alors necessaire pour ameliorer la qualite de l'image obtenue et eviter les erreurs. L'image est donc filtree et, pour augmenter l'efficacite du traitement, les caracteristiques locales des stries (direction et frequence) sont generalement utilisees [14].

La reconnaissance d'empreinte est basee sur l'extraction de la signature. La signature d'une empreinte digitale correspond a l'information utile necessaire a l'identification fiable de la personne ou a l'archivage dans la base de donnees. Elle permet de caracteriser de maniere unique la personne.

La tres grande majorite des techniques de reconnaissance sont basees sur la detection locale des minuties [13] et l'extraction de leurs caracteristiques (type, direction locale,...) car historiquement, c'est la technique qui a toujours ete utilisee par les experts humains. Certains algorithmes permettent d'extraire l'information des minuties directement a partir de l'image en niveaux de gris en suivant le maximum local des stries [15] neanmoins cela necessite une bonne qualite d'image a la base, c'est pourquoi la plupart des algorithmes preferent travailler sur un squelette binaire de l'image ou l'extraction est grandement facilitee. L'inconvenient de cette technique est de produire la detection d'un nombre important de fausses minuties, un post-traitement est alors necessaire pour les eliminer [16]. Il arrive egalement que les pores de l'empreinte [17] soient utilisees, mais cela necessite des images de tres hautes definitions et reste donc peu utilise.

La structure globale de l'empreinte peut aussi etre utilise [18], mais les resultats sont generalement moins precis qu'avec les caracteristiques locales. Ce genre

de methode est en general associee a l'extraction des minuties pour augmenter les performances du systeme, ou utilisee pour classifier les empreintes.

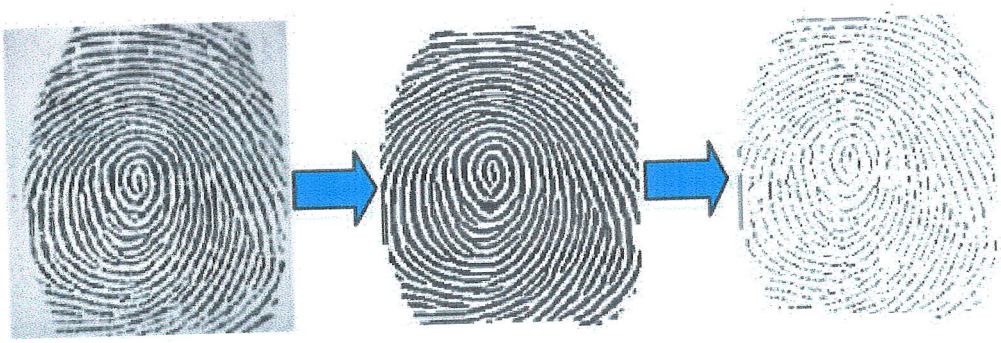


FIGURE 2.5 – synopsis du pré-traitement des images d'empreintes digitales en vue de l'extraction des minuties

2.3.4 Le stockage et la phase d'appariement

Pour les systemes disposant de grosses bases de donnees, l'identification peut poser probleme en temps de calcul si la signature d'entree doit etre comparee avec toutes les signatures presentes dans la base. C'est pourquoi un processus de classification et de declassification est necessaire pour limiter les temps de recherche.

Lorsqu'une image est stockee, un groupe specifique lui est attribue en fonction de ses caracteristiques. Lors de l'identification on desarchive l'ensemble des signatures de la base correspondant au groupe de l'empreinte necessitant l'identification. Puis chacune des images desarchivees est comparee avec celle de l'utilisateur. Ceci permet de reduire sensiblement les temps de recherche en limitant le nombre d'images a comparer, a condition que les differentes categories soient judicieusement choisies. Parmi les differentes techniques existantes [19] on distingue principalement l'approche syntaxique (l'image est decrite au moyen de regles et de symboles et une analyse grammaticale permet de lui associer une classe), l'extraction des singularites de l'image (la position des centre et delta permet de determiner la classe de l'empreinte) et l'utilisation des reseaux de neurones.

La phase d'appariement est l'etape critique du systeme, elle recoit en entree deux signatures issues de deux acquisitions differentes d'empreinte et renvoie en sortie un resultat binaire indiquant si oui ou non les deux signatures proviennent de la meme empreinte. Bien entendu deux empreintes provenant de la meme personne ne seront jamais identiques en raison de l'elasticite de la peau, de la presence de poussiere, de l'orientation du doigt lors de l'acquisition Ceci est caracteristique des systemes biometriques. La phase d'appariement va donc calculer le degre de similarite (taux d'appariement) entre les deux signatures et decider si elles peuvent etre considerees identiques en fonction d'une valeur seuil. [20]

Bien que les deux empreintes puissent etre comparees directement par correlation [21] la methode qui a suscite le plus d interet utilise les caracteristiques locales des minuties et consiste en l appariement base sur l alignement d un motif de point [24] car il est simple en theorie, efficace pour faire face a la fausse information detectee dans les phases precedentes, et rapide par rapport aux autres methodes. Cet algorithme est divise en deux processus :

- L'alignement : on evalue la transformation geometrique (orientation, translation, homothetie) entre les deux ensembles a traiter et on les aligne suivant cette transformation.
- L'appariement : on evalue le nombre d elements caracteristiques qui sont alignes (moyennant une certaine marge d erreurs car un alignement parfait est impossible) et le taux d appariement est calcule en fonction des correspondances rencontrees.

2.4 Evaluation des systèmes d'authentication biométrique

2.4.1 Avantages

- La technologie la plus éprouvée techniquement et la plus connue du grand public.
- Petite taille du lecteur facilitant son intégration dans la majorité des applications (téléphones portables, PC).
- Faible coût des lecteurs grâce aux nouveaux capteurs de type "Chip silicium".
- Traitement rapide
- Bon compromis entre le taux de faux rejet et le taux de fausse acceptation.

2.4.2 Inconvénients

- Image "policière" des empreintes digitales.
- Besoin de la coopération de l'utilisateur (pose correcte du doigt sur le lecteur).
- Certains systèmes peuvent accepter un moulage de doigt ou un doigt coupé (la détection du doigt vivant permet d'éviter ce type d'usurpation)

3 Vulnérabilités et menaces

Deux types de défaillance sont principalement identiés : le déni de service et l'intrusion. Dans le déni de service, l'utilisateur légitime est refusé par le système. Par contre, l'intrusion se réfère à un accès illégitime au système. Pour des raisons de sécurité, ils concentrent leur attention sur le risque d'intrusion dans lequel un attaquant doit d'abord obtenir les données biométriques puis essayer de les injecter dans le système biométrique. En se focalisant sur l'empreinte digitale, les auteurs exposent des procédés d'obtention des données d'empreintes digitales et des procédés pour les injecter dans le système.

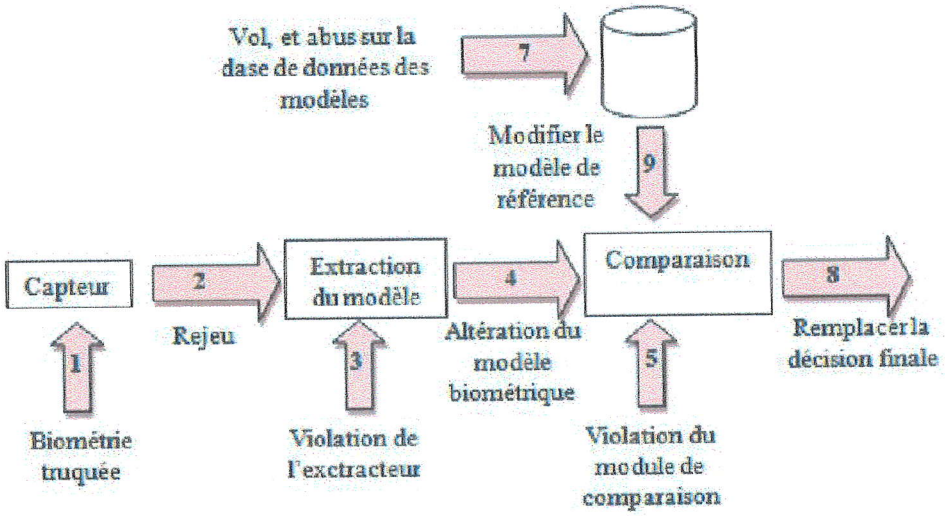


FIGURE 2.6 – Les points de vulnérabilités d'un système d'empreintes digitales [25]

Pour leur généralité, nous détaillons les points d'attaques de Ratha et al [25] :

- ***Point 1 :*** L'attaque à ce niveau consiste à présenter une fausse empreinte sur le capteur. Cette attaque est appelée spoong si la modalité utilisée est de nature physiologique et mimicry si elle est comportementales.
- ***Point 2 :*** A ce niveau, des données biométriques interceptées sont soumises au module de l'extraction de caractéristiques en passant le capteur. Le canal, entre le capteur et le module de l'extraction de caractéristiques, peut être intercepté pour voler une image (selon la modalité biométrique utilisée) d'un utilisateur légitime prise par le capteur. Cette image peut être rejouée ultérieurement au module d'extraction de caractéristiques en contournant le capteur.
- ***Point 3 :*** Le module de l'extraction de caractéristique est remplacé par un programme Cheval de Troie (Trojan-horse) qui fonctionne selon les spécications de son concepteur
- ***Point 4 :*** Le canal entre le module de l'extraction de caractéristiques et de classieur peut être espionné par un adversaire pour enregistrer un modèle biométrique d'un utilisateur légitime. Ce modèle peut être rejoué ultérieurement sur le même canal.
- ***Point 5 :*** Un programme de type cheval de Troie peut se déguiser en un classieur, en contournant le vrai module de comparaison, pour soumettre un score decorrespondance qui permet de prendre une décision qui est en faveur de l'adversaire.
- ***Point 6 :*** L'attaque à ce niveau est contre les modèles biométriques stockés (voler, remplacer, supprimer ou modier les modèles). Cette attaque pourrait être lancée au cours du temps d'enrôlement, pendant la période de vérication, ou à tout moment directement sur la base de données. Dans une application de carte à puce, où le modèle est stocké dans la carte qui est portée par l'utilisateur, si la carte est perdue ou volée, et si elle n'est pas protégée d'une manière adéquate, le modèle biométrique peut être récupéré facilement.

- ***Point 7 :*** A ce niveau, les modèles biométriques sont traqués dans les upport de transmission entre la base de données et le classieur et ils peuvent être rejoués ultérieurement sur le même canal.
- ***Point 8 :*** Le canal entre le classieur et l'application qui a envoyé une requête de vérication, peut être espionné pour accéder à la réponse d'une vérication précédente et l'enregistrer. Cette réponse peut être rejouée ultérieurement dans le même canal.

4 Les exigences d'un système de reconnaissance d'empreintes digitales

Pour bien protéger un modèle biométrique de reconnaisse d'empreintes digitales les algorithmes doivent être conçus de façon à garantir les exigences suivantes :

- **Irréversibilité (irreversibility) :** il devrait être impossible d'obtenir la référence biométrique originale à partir du modèle protégé. Cette propriété assure la condentialité de la donnée biométrique ce qui a un impact directe sur la préservation de la vie privée
- **Intraçabilité/Diversité (unlinkability/diversity) :** il devrait être possible de produire un très grand nombre de modèles protégés (à utiliser dans des applications diérentes) à partir du même modèle non protégé. Cela permettra d'éviter la poursuite et la surveillance des utilisateurs à travers diérentes bases de données.
- **- Révocabilité et renouvellement (revocability and renewability) :** en cas de compromission du modèle de référence comme son vol, il devrait être possible de le révoquer et de générer une nouvelle référence, diérente de la précédente, à partir du même échantillon biométrique.

5 Sécuriser le modèle biométrique des empreintes digitales

Dans cette section, nous parlons de quelques approches existantes pour protéger le modèle biométrique des empreintes digitales. Nous nous intéressons au niveau algorithmique afin de sécuriser le template du système de reconnaissance d'empreintes digitales.

5.1 La cryptographie

En se basant sur des mécanismes de la cryptographie, le chiffrement des caractéristiques d'empreintes digitales est considéré comme moyen de gérer l'information biométrique. Cela désignées pour maintenir la confidentialité et l'intégrité du modèle biométrique.

L'évolution de la technologie et la quantité importante d'information visuelle que les images d'empreintes digitales contiennent. certains problèmes de la sécurité sont apparus. Pour cela une image numérique d'empreintes digitales qui est composée d'une matrice de pixels a été protégée par l'approche classique de l'application des

techniques cryptographiques pour les données visuelles. Cette technique vise le chiffrement de chaque pixel le l'image ou le convertir en une valeur chiffrée. Le récepteur de l'image chiffrée qui a été transmette à travers un canal, pas forcement sécurisé, doit appliquer un algorithme de déchiffrement pour rendre l'image reçue à sa forme originale.

La sécurité de l'image transmise chiffrée dépend de la sécurité de l'algorithme de chiffrement utilisé qui peut être eectué dans le domaine spatial ou fréquentiel :

1. ***Dans le domaine spatial :*** le chiffrement est effectué directement sur les valeurs des pixels de l'image à chiffrer. L'avantage du chiffrement direct de ces valeurs est la faible complexité de calcul.
2. ***Dans le domaine fréquentiel :*** le chiffrement est effectué sur les coefficients de l'image et non pas sur ses valeurs de pixels. Les coefficients réels sont calculés à partir des transformations fréquentielles comme FFT (Fast Fourier Transform), DCT (Discrete Cosine Transform) ou DWT(DiscreteWaveletTransform).avant le processus de chiffrement une quantication est appliquée sut les coefficients.

Le niveau de sécurité de chiffrement des images numériques d'empreintes peut être élevé en chiffrant le contenu totale de l'image (chiffrement complet), et peut être aussi faible en chiffrant q'une partie précise du contenu de toute l'image(chiffrement sélectif). Dans le chiffrement complet on applique des algorithmes de chiffrement symétriques ou asymétriques précédemment présentés, qui sont plus coûteux en termes de temps de calcul et de mémoires. mais quand un niveau plus faible de sécurité est envisageable,il est préférable de faire un chiffrement sélectif pour par exemple les applications à temps réel comme les visualisations en temps réel et les caméras de surveillance, ces images doivent être rapidement transmises et le cryptage total n'est pas nécessaire. le chiffrement sélectif est toujours inférieure au niveau de sécurité à un chiffrement complet, mais il diminue la quantité de données à chiffrer, et par conséquent le temps de calcul et de mémoire. [26] ces données à chiffrer doivent être choisit d'une manière intelligente an de pouvoir afficher une image correctement après le déchiffrement. ce choix doit vérifier les propriétés suivantes [27] :

- *Acceptation visuelle :* une partie de l'information peut être visible mais l'image cryptée devrait apparaitre bruitée.
- *Chiffrement sélectif :* le chiffrement sélectif doit se produire pendant la compression de l'image.
- *Débit binaire constant :* le chiffrement sélectif ne doit pas utiliser un algorithme de chiffrement qui entraîne une augmentation de la taille des données.
- *Conformité du flux binaire :* l'algorithme de chiffrement doit fournir un ux binaire conforme, basé sur de la dénition du format choisi.

Différentes techniques de chiffrement utilisant des algorithmes standards comme le DES et l'AES ont été utilisés pour le chiffrement sélectif des images et des vidéos dans la littérature.

Il existe plusieurs algorithmes qui se résument sous le terme de la cryptographie. Cependant, même si la cryptographie a prouvé son efficacité pour sécuriser le stockage et la transmission de l'information, elle devient inadéquate lorsqu'il s'agit de biométrie. En eet, à cause de la variabilité du signal biométrique, la comparaison devrait se faire dans l'espace en clair ce qui implique qu'un attaquant puisse toujours essayer d'avoir le contrôle sur la donnée biométrique. La plupart des risques de violation de vie privée demeure problématique.

5.2 Cryptosystème biométrique : Fuzzy Vault et Fuzzy commitment

Le fuzzy commitment est une primitive cryptographique, qui peut être vu comme un secure sketch sur l'espace $\{0,1\}^n$ muni de la distance de Hamming d_H utilisant un code algébrique linéaire.

Un code algébrique $[n,k,d]$ est un sous-espace vectoriel de $\{0,1\}^n$ de dimension k et composé des vecteurs x muni du poids de Hamming $w_H(x) > d$ avec $w_H(x)$ le nombre d'éléments différents de 0 dans x. La capacité de correction du code est $t = (d-1)/2$. Le principe du fuzzy commitment est alors décrit comme suit :

- **Durant l'enrôlement**, un mot de code $C \in \{0,1\}n^n$ est calculé à partir de lclé S . Le choix de ce code dépend de la quantité d'erreur à traiter. On sauvegardera sur la base de données uniquement le couple : $(C \bigotimes X, H(C))$.
- **Durant la vérification**, la valeur $(C \bigotimes X \bigotimes Y)$ est calculée et corrigée pour dériver le secret $\grave{C}$. La comparaison réussie si $H(C) = H(\grave{C})$.

L'un des points faibles du fuzzy commitment est qu'il devient impraticable lorsque le taux d'erreur est assez élevé (un code correcteur $[n,k,d]$ peut corriger un maximum de $(d-1)/2$ erreurs). Il n'est donc possible de retrouver C que si $d_H(X,Y) <= t$ avec $t = (d-1)/2$.

Parmi les applications de ce protocole sur les empreintes digitales, nous citons celle de Tuyls et al [28]. qui utilise le code **BCH** et une méthode de sélection des éléments les plus fiables dans le vecteur X. Une autre application est celle d'Arakala et al [29]. toujours en utilisant le code **BCH** mais en impliquant des descripteurs locaux sur le modèle des minuties. Malheureusement, en pratique, ces applications engendrent des taux d'erreurs assez importants. Par exemple, dans Arakala et al nous trouvons un taux d'erreur $EER = 15\%$ sur la base de données publique FVC2000.

En 2002, Juels et Sudan [30] modifient cette approche afin qu'elle soit utilisée pour des représentations partielles sous l'appellation du fuzzy vault où le principe d'interpolation polynomiale a été utilisé comme illustré sur la figure 2.7

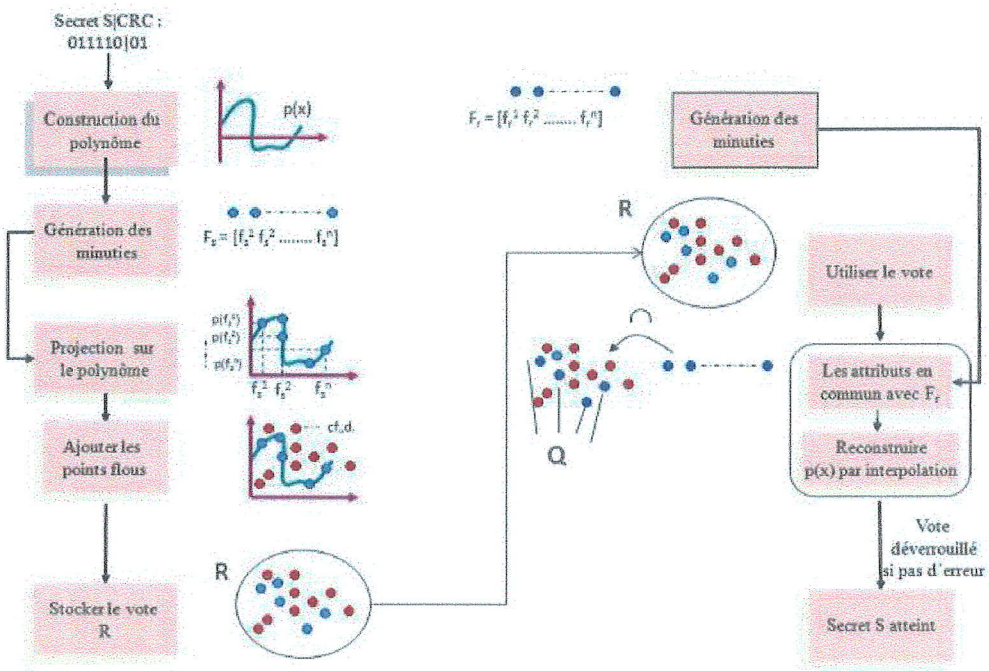


FIGURE 2.7 – Principe du fuzzy vault

Le fuzzy vault est aussi un secure sketch mais opéré sur une métrique différente qui est relative à la différence entre deux ensembles (set difference) au lieu d'une distance de Hamming. Ce schéma est potentiellement applicable aux minuties d'empreintes digitales contrairement au précédent qui ne considère en entrée que des séquences stables en ordre et en taille.

Les applications du fuzzy vault aux minuties ont été initiées dans plusieurs travaux. parmi les implémentations les plus performantes nous citons celle de Nandakumar et al. Dans ce travail, plusieurs données supplémentaires ont été introduites afin d'améliorer le taux de faux rejet du système, comme : les points de courbure maximale pour faciliter l'alignement des minuties ou leur index de qualité pour permettre de sélectionner les plus pertinentes. Ils obtiennent enfin un $FRR = 10\%$ (pour un FAR à 0%) sur une version partielle de la FVC2002

Plusieurs travaux remettent en question la sécurité et la préservation de la vie privée impliquées dans les différentes constructions du secure sketch :

- La donnée auxiliaire W du fuzzy commitment occasionne une divulgation d'information sur le modèle d'origine comme il est montré dans les travaux initié par Zhou dans montrent que la correlation existante dans les modèles biométriques binaires intraclasse peut aider à estimer à partir de W, la clé secrète S (la corrélation implique que la prédiction d'un élément x_j à partir d'un élément x_i dans le vecteur X est possible).
- Dans Simoene et al. prouvent formellement que le fuzzy commitment ne peut pas être réutilisé dans différentes bases de données à partir de la même biométrie : le fuzzy commitment est dans ce cas non révocable. En effet, ils estiment une probabilité proche de 1 pour qu'un attaquant puisse lier les sketches entre eux. Le secure sketch n'a donc pas les propriétés de préservation de vie privée puisque sa réutilisation n'est pas possible.

— Dans quatre différentes attaques, par rapport au fuzzy vault ont été distinguées. Ces attaques permettent de retrouver le modèle d'origine lorsque celui-ci est protégé par le fuzzy vault sur différentes bases de données. Le fuzzy vault lui aussi n'est pas réutilisable (ou révocable). Une tentative d'amélioration de cet aspect a été proposée dans

Afin d'améliorer la sécurité du fuzzy commitment qui est en relation avec la protection de la donnée auxiliaire W, Bringer et Chabanne proposent d'utiliser la cryptographier hommomorphique. Ils combinent le fuzzy commitment avec la primitive de Goldwasser-Micali. Dans ce schéma de cryptographie asymétrique, une clé publique p_k et sa clé secrète s_k sont générées. La propriété hommomorphique de cette primitive est la suivante :

$$Enc(m, p_k) \times Enc(\acute{m}, p_k) = Enc(m \bigotimes \acute{m}, p_k)$$

ou bien,

$$Dec(Enc(m, p_k) \times Enc(\acute{m}, p_k), s_k) = m \bigotimes \acute{m}$$

Le protocole de vérification biométrique est maintenant décrit comme suit :

— **Durant l'enrôlement**, l'utilisateur U enregistre sa biométrie X auprès du serveur d'authentification. La serveur génère aléatoirement le mot de code C. En utilisant Goldwasser-Micali, le serveur chiffre $C \bigotimes X$ avec la clé publique p_k . Il sauvegarde sur la base de données $H(C)$ et $Enc(C \bigotimes X, p_k)$.

— **Durant la vérification**, l'utilisateur chiffre sa biométrie Y avec sa clé publique p_k et envoie $Enc(Y, p_k)$ au serveur. Celui-ci récupère $Enc(C \bigotimes X, p_k)$ et $H(C)$ de la base de données et envoie le produit $Enc(C \bigotimes X, p_k)$ $\times$Enc(Y , p_k) au gestionnaire de clé (key manager). Le key manager utilise la clé privée s_k pour calculer $Dec(Enc(C \bigotimes X, p_k)$ $\times$Enc(Y , p_k), $s_k) = C \bigotimes X \bigotimes Y$ et envoie le résultat au serveur d'authentification pour le décoder vers le mot de code $\grave{C}$. Il vérifie ensuite si $H(C) = H(\grave{C})$.

La propriété hommomorphique de Goldwasser-Micali assure que ni le modèle Y , ni la donnée W ne soient révélés en clair au niveau de la base de données. D'autres implémentations de la cryptographie hommomorphique pour assurer des protocoles privés de vérification biométrique existent. Pour les empreintes, nous citons le travail intéressant de Upmanyu et al. Néanmois, il reste à noter que le cryptographier hommomorphique est complexe à implémenter en terme de temps de calcul pour l'authentification biométrique.

5.3 Le tatouage

Le tatouage numérique ou "digital watermarking" en anglais est une technique d'insertion de données cachées comme la stéganographie mais l'ojectif est différent de celle de la stéganographie. Le tatouage numérique consiste à insérer un message invisible (dans certain cas visible) appelé marque 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. Pour le tatouage des images, l'insertion de la marque est effectuée en général dans le domaine spatial ou le domaine fréquentiel. Les principales contraintes techniques à prendre en considération pour concevoir un algorithme de tatouage performant sont les suivantes :

- ***La capacité :*** C'est la quantité d'information que l'on désire cachée par rapport à la quantité d'information associée au support image audio, vidéo. Dans le cas du tatouage 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, assurer des services d'intégrité.
- ***L'imperceptibilité :*** Appelé aussi invisibilité, le but est de faire en sorte que le stego- medium reste fidèle au medium original. Les données ne doivent pas être «perceptibles» dans le stégo-médium. Pour le tatouage, l'objectif est de ne pas détériorer le stégo-médium protégé. Cependant, la contrainte est plus forte en stéganographie où il s'agit plutôt d'une indétectabilité statistique.
- ***La robustesse :*** Le but de cette propriété est de récupérer les données cachées même si le stego-medium a été manipulé. On peut définir la robustesse par la résistance du marquage face à des manipulations du stégo-medium. Dans le cas où le support est une image, les manipulations peuvent être de type géométrique (rotation, zoom, découpage,. . .), elles peuvent 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 dissimulé dans ce cas le stégo-medium risque d'être détecté, de la même manière si le but est de rendre le message (marque) plus robuste, cela aura en contrepartie pour rendre ce dernier 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

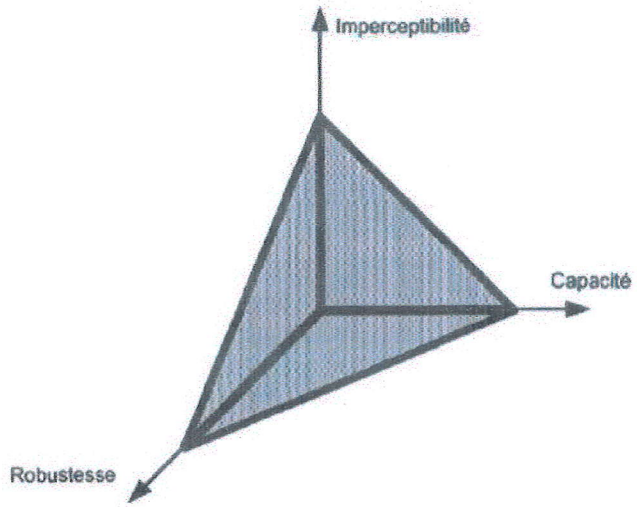


FIGURE 2.8 – Problèmatique du tatouage des images.

Pour ces raisons, les techniques de tatouage sont groupées selon différentes classifications [31] : conformément au type de la clef appliquée (asymétrique et symétrique) ; selon l'information nécessaire à l'extraction (aveugle, semi-aveugle et non aveugle) ; conformément à la robustesse (fragile, semi-fragile et robuste) ; quant à la perception du système visuel humain (visible et invisible) ; selon la préservation

de l'image originale (inversible et non-inversible) et conformément à la technique d'insertion (additive et substitutive).

5.4 Hachage cryptographique

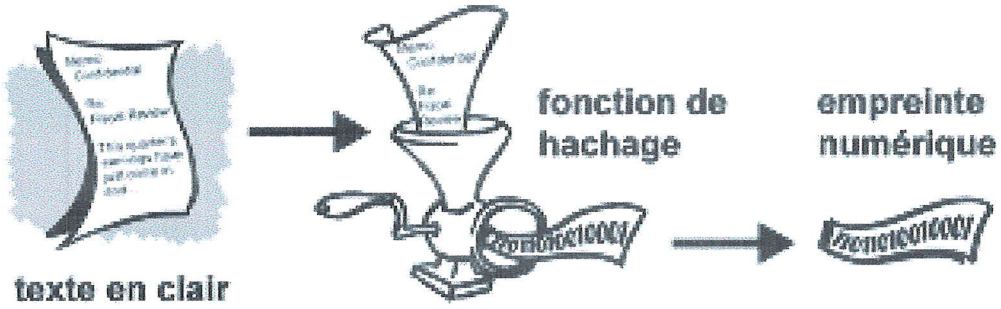


FIGURE 2.9 – Fonction de hachage cryptographique

Une fonction de hachage cryptographique est une fonction de hachage qui permet de générer à partir d'une donnée de taille arbitraire et potentiellement grande, une image de taille fixe appelée valeur de hachage,empreinte,ou encore empreinte numérique. Cette empreinte est utilisé pour identifier la données originale d'une manière univoque. les fonctions de hachage cryptographique ont été créées pour faciliter la gestion de base de données. Plutôt que de manipuler des données de taille variable et grande, des comparaisons rapides sont effectuées sur les empreintes qui sont de petites tailles.

Les fonctions de hachage cryptographiques ne sont pas considérées comme des systèmes de chiffrement du fait de :

1. *L'absence de clef :* la définition des fonctions de hachage cryptographique ne fait pas intervenir de clef. Pour certaines applications, la donnée hachée est entièrement publique, et un attaquant éventuel peut effectuer lui-même les calculs de l'empreinte
2. *L'absence d'algorithme d'inversion :* contrairement aux primitives de chiffrement par blocs ou de chiffrement asymétrique, les fonctions de hachage cryptographique n'ont pas besoin de pouvoir être inversées (fonctions à sens unique).L'impossibilité de calculer un antécédent par une fonction de hachage cryptographique est même une propriété de sécurité fondamentale de ces primitives.

5.4.1 Critères de sécurité

Pour qu'une fonction de hachage cryptographique soit idéale elle doit vérier les propriétés suivantes :

- ***La compression :*** pour tout message de n'importe quelle taille, l'empreinte associée est de taille fixe (petite taille).
- ***L'efficacité :*** la valeur de hachage d'un message se calcule très rapidement.

- ***La résistance aux collisions :*** il est pratiquement impossible de trouver deux messages différents ayant la même valeur de hachage.
- ***La résistance au calcul de préimages :*** il est pratiquement impossible, pour une valeur de hachage donnée, de construire un message ayant cette valeur de hachage.
- ***La résistance au calcul de secondes préimages :*** il est pratiquement impossible de modier un message sans changer sa valeur de hachage

Les fonctions de hachage cryptographique ont donc plusieurs rôles en cryptographie grâce aux nombreuses propriétés qu'elles vérifient. Ces fonction peuvent être utilisées pour l'authentication des images d'empreintes digitales . Le principal inconvénient de ces techniques cryptographiques est que la valeur de la signature numérique(hash) va se changer complètement même avec le changement d'un seul bit dans l'image, en raison d'une opération de compression, de filtrage ou de transformation géométrique. Pour cela les fonctions de hachage cryptographique représentent une forte source d'inspiration des techniques de hachage perceptuelles des images en visant l'aspect visuel. Ces techniques seront détailler dans le chapitre suivant.

6 Conclusion

Nous avons détaillé dans ce chapirte le système de reconnaissance d'empreintes digitales, les vulnérabilités et les menaces de ce système, ansi que les exigences de sécurité. et nous avons vu aussi quelques techniques de sécurisation déja existées pour ce système. Il existe plusieurs approches de comparaison d'empreintes digitales ; la majorité d'entre elles se base sur la comparaison de minuties de l'empreinte à identifier aux minuties des Template (c'est la comparaison de signatures des empreintes). Il convient donc d'élaborer des méthodes efficaces de détection de minuties, afin de mettre en place un système de reconnaissance fiable.

Dans le chapitre suivant nous présenterons les fonctions de hachge perceptuel des images, ensuite la méthode proposée de reconnaissance des empreintes digitales et les détecteurs utilisés pour l'extraction les caractéristiques des empreintes digitales.

Chapitre 4

Mise en Œuvre et Evaluation

1 Introduction

Après avoir présenté dans le chapitre précédent les différentes étapes de notre méthode pour sécuriser les images d'empreintes digitales, nous entamerons la partie mise en oeuvre et évaluation. Dans cette partie, nous mettrons en évidence les raisons de nos choix technique (langage de programmation et outils utilisés) ainsi que des interfaces permettant de représenter les fonctionnalités du système élaboré dans le cadre de ce présent travail. Et nous finirons par une évaluation générale de notre système.

2 Mise en Œuvre

Dans cette section, nous présonterons les outils de dévloppement qui nous ont servi à la mise en oeuvre de notre système, ainsi que les différentes interfaces de notre application.

2.1 Outils de développement

Le choix de bon environnement de programmation est très important pour le dévloppement des projets. Cela se fait suivant plusieurs facteurs : la puissance de la compilation, la facilité d'utilisation, la disponibilité de plusieurs fonctionnalités, la communication avec d'autres environnements, etc.

Afin de réaliser notre système, nous avons eu recours aux outils suivants : Matlab, MySQL comme système de gestion de base de donnés et un serveur Web apache.

2.1.1 Matlab

Matlab [2] est un logiciel de calcul numérique produit par MathWorks.C'est un longage simple et très efficace, optimisé pour le traitement des matrices, d'où son nom. Pour le calcul numérique, il est beaucoup plus concis que les "vieux" langages (C, Pascal, Fortran, Basic). Matlab contient également une interface graphique puissante,ainsi qu'une grande variété d'algorithmes scientifiques.

On peut enrichir Matlab en ajoutant des "boites à outils" (toolbox) qui sont des ensembles de fonctions, profilées pour des applications particulières (traitement de

signaux, analyses statistiques, optimisation,etc.).

2.1.2 Le serveur web Apache

Les serveurs web les plus populaires aujourd'hui sont : Apache, Microsoft IIS, Zeus et Sun One, Nous avons choisit Apache pour les avantages suivants :

- Apache est gratuit.
- Le code source d'Apache est libre ce qui permet sa personnalisation pour une applacation particulière. Cette disponibilité du code source est la raison principale pour sa popularité.
- La configuration d'Apache s'effectue en modifiant ses fichier de configuration au sein des quels des directives permettent de définir son comportement. Cette méthode de configuration lui procure une souplesse permettant à l'administrateur du serveur un controle sur les fonctionnalits et la sécurité offerte par Apache.
- Apache a une structure modulaire. L'administrateur est lible de déterminer les modules nécessaires seulemnt.

2.1.3 Le SGBD MySql

Les SGBD libres et gratuits sont nombreux : MYSQL, mSQL et Postgres en sont des exemples. Si nous avons choisit MYSQL, C'est plus pour des raisons de performances et fonctionnalités offertes. Nous citons dans la suite ses principaux avantages :

- MYSQL est beaucoup moins complexes à installer est à administrer que d'autres systèmes.
- MYSQL supporte le langage de requete SQL, comme on peut y accéder via ODBC.
- MYSQL permet des connexions multiples en meme temps et utilise différentes bases de données sumultanément.
- MYSQL dispose d'un système de controle intégré qui interdit la consultation de données à ceux qui n'en ont pas l'autorisation.

2.2 Présentation de l'application

Dans cette section nous donnerons un aperçu général de la phase pratique de notre application qui consiste à décrire les différentes étapes et fonctionnalités disposées de la téchnique proposée et nous présenterons les interfaces principales.

2.2.1 Menu principale

La première rencontre avec notre applacation sera via la page d'accueil, celle-ci contient un accés permettant de voir toutes les fonctionnalités et les étapes utilisées durant le hashage perceptuel à base shape context, et un autre accés qui mène à un système de reconnaissance d'empreinte digitale en basant sur la méthode proposée, comme montré dans la figure.

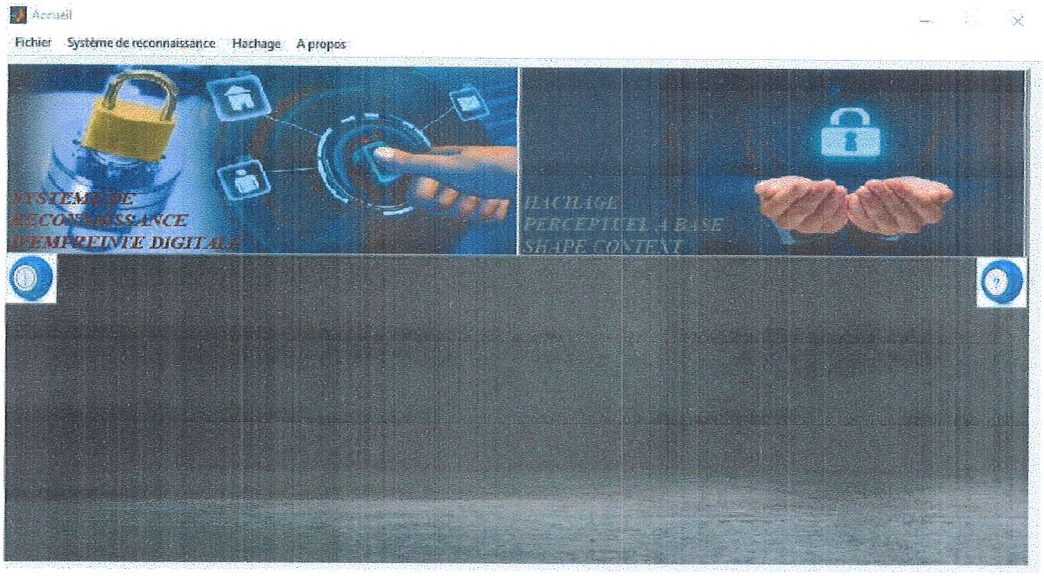


FIGURE 4.1 – Interface principale.

2.2.2 Hachage perceptuel à base shape context

Dans ce Système nous présenterons les processus utilisés et les mésures de performance de notre méthode. comme illustré dans la figure suivante :

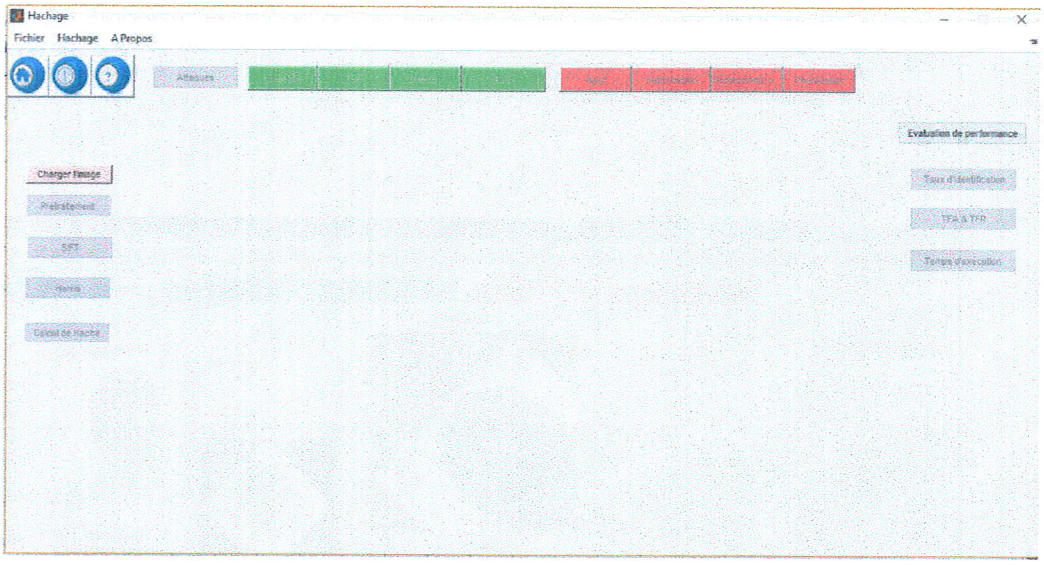


FIGURE 4.2 – Hachage perceptuel à base shape context.

cette onglet décrit les démarches suivées à partir du chargement de l'image jusqu'à la création de la signature numérique (hash) en passant par le détecteur de SIFT et Harris et la téchnique Shape context.

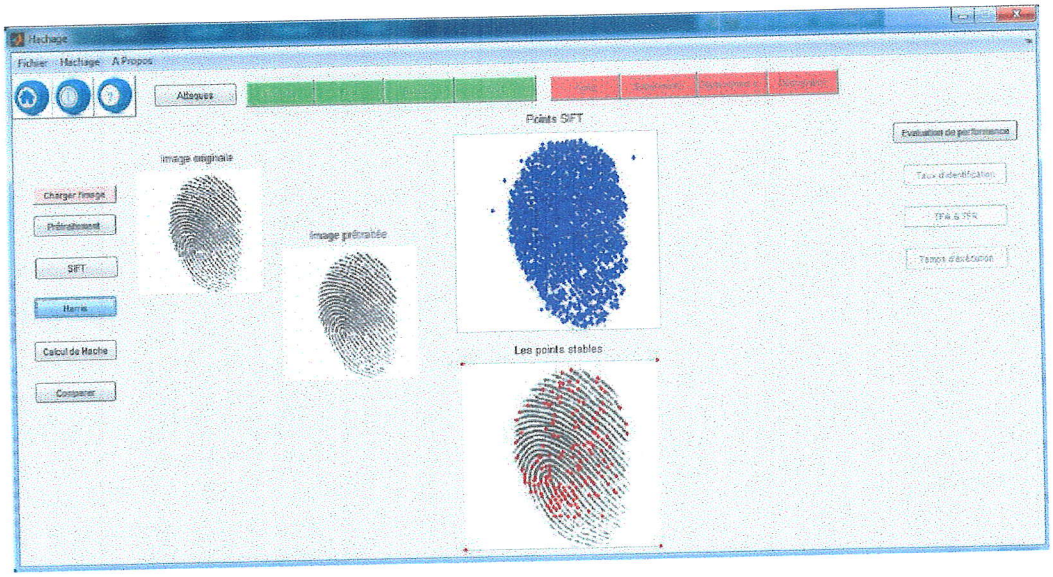


FIGURE 4.3 – Présentation

Ainsi que l'application de quelques attaques acceptables et d'autres malveillantes sur les images d'empreinte digitale et faire la comparaison entre eux.

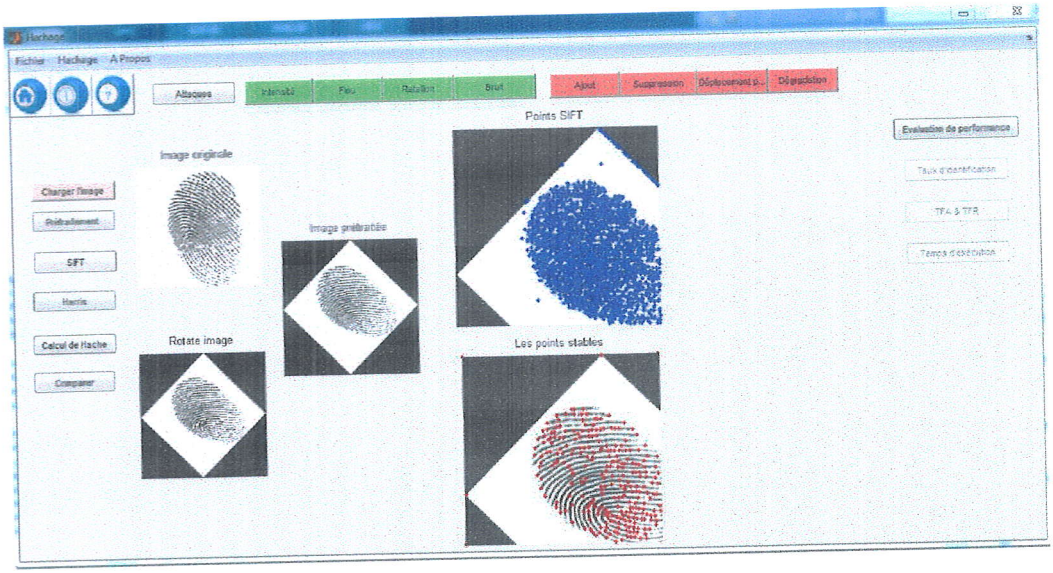


FIGURE 4.4 – Rotation

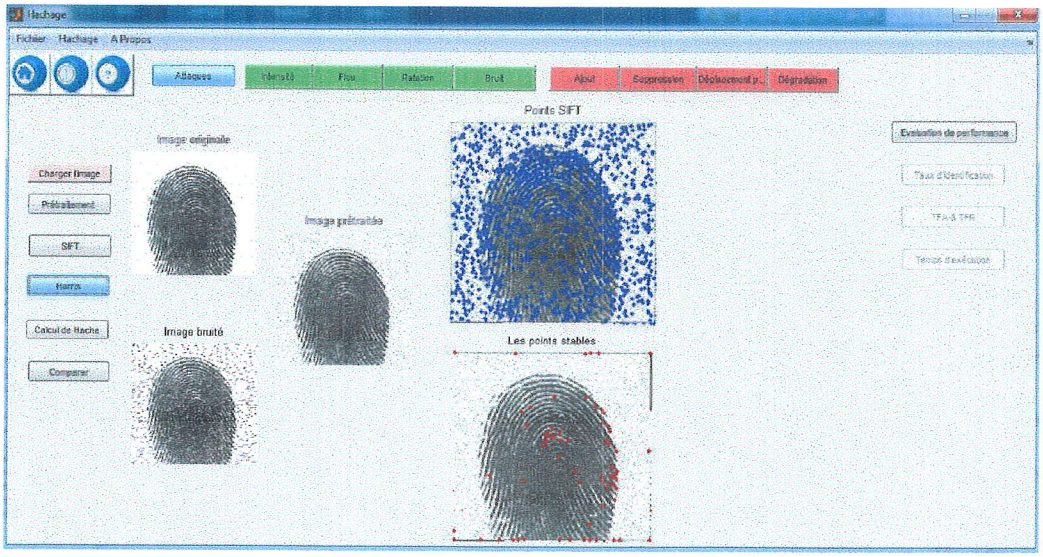


FIGURE 4.5 – Bruit

Et puis nous comparons les images et leurs hashs (l'image attaquée et l'image originale).

Nous trouverons aussi une partie de l'évaluation des performances se résume dans le taux d'identification et d'authentification, les courbes de TFA et TFR, temps d'éxécution, etc...

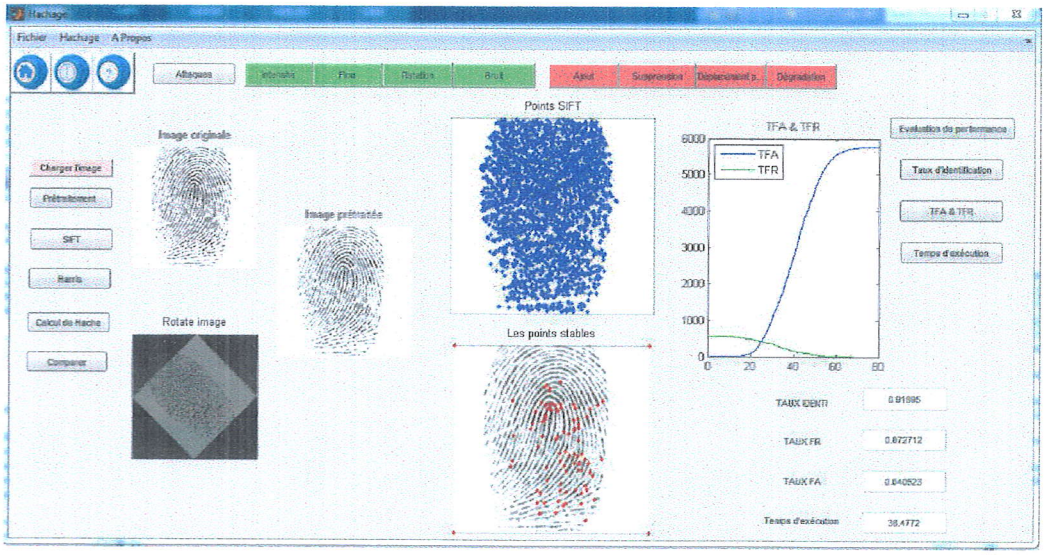


FIGURE 4.6 – performance

2.2.3 Système de reconnaissance d'empreinte digitale

Dans ce système nous verrons trois onglets ***(Enrollement, Identification et Vérification)*** comme montré dans la figure suivante :

Pour enregistrer une nouvelle personne nous avons le bouton *Enrollement* en entrant les informations personnelles de chaque personne (Nom, Prenom,Photos, ...), ansi que son image d'empreinte digitale et sa signatures numériques (Hache). qui seront stockées dans la base de données.

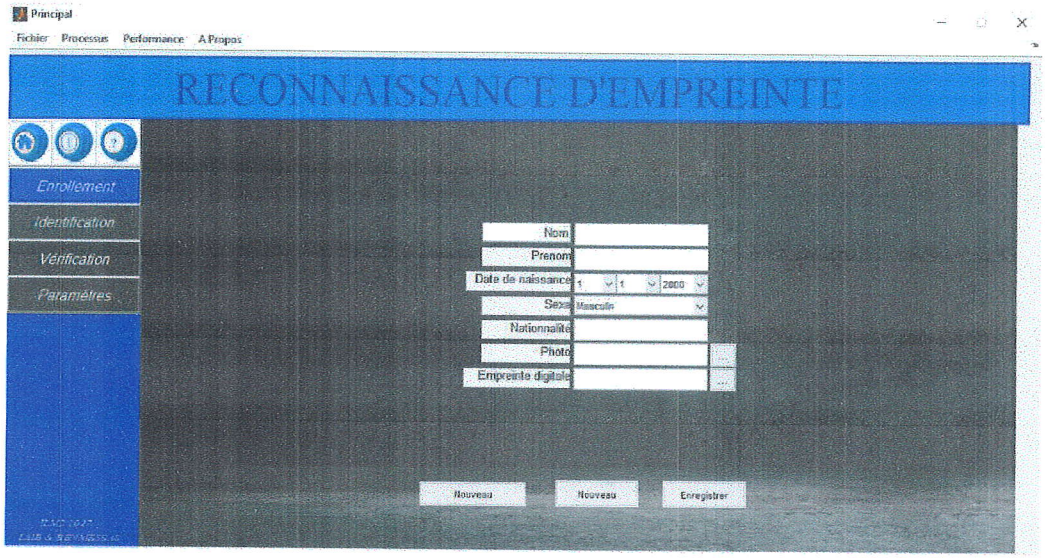


FIGURE 4.7 – Onglet Onrollemnt.

2.2.4 Onglet Identification

Pour charger une image, il suffit de cliquer sur le bouton'Charger l'empreinte à identifier', une fenetre de dialogue s'affiche permettant de choisir l'image. une fois l'image à identifier est sélectionner on clique sur le bouton 'Lancer l'identification' pour déclencher la phase d'appariement.

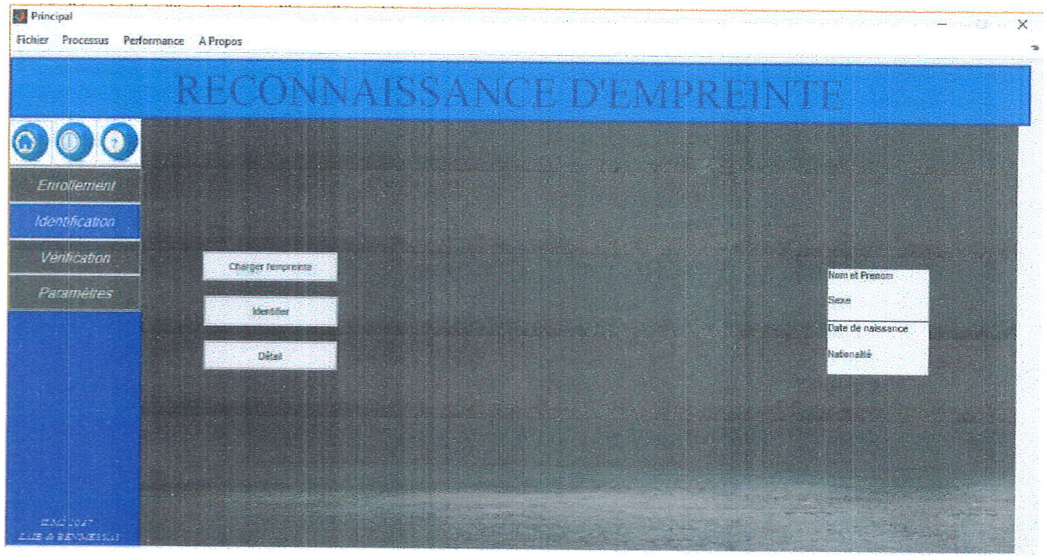


FIGURE 4.8 – Onglet Identification.

Quand le processus d'identification est terminé, les résultats de l'empreinte choisit (Informations de la personne, Sa photo et son emreinte) seront affichées.

2.2.5 Onglet Vérification

l'authentification des utilisateurs se faite en entrant son Nom et son Prénom avec l'image de son empreinte.

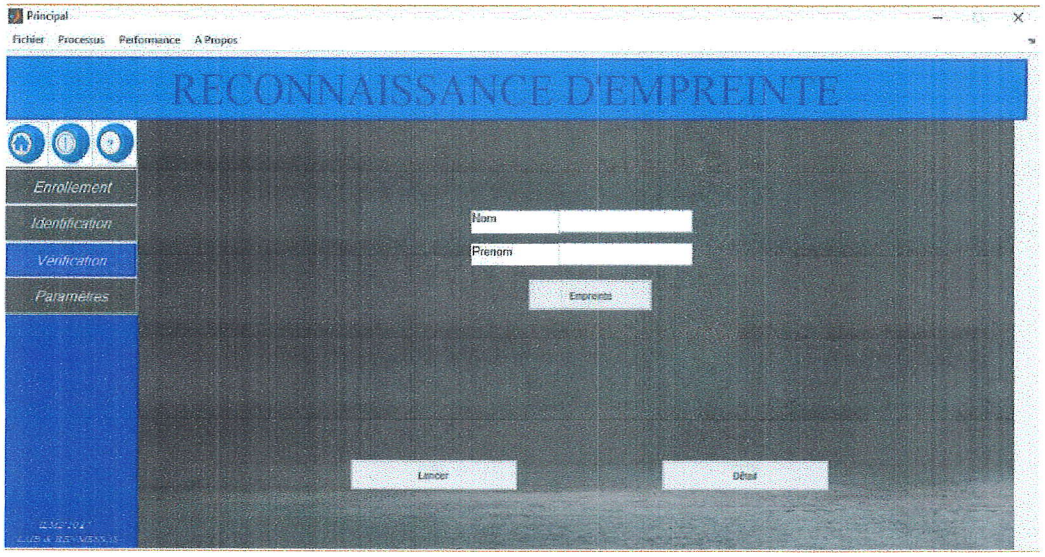


FIGURE 4.9 – Phase d'authentification.

Le système compare le hache de cette image avec ceux stockées dans la base, si les deux haches sont de similaires, un message de confirmation s'affiche, sino un message indique la dissimilarité des empreintes.

2.2.6 Onglet Paramètre

Cet onglet nous permet de modifier les seuil utilisés dans le système et quelques autres réglages concernant l'application.

2.2.7 Onglet A propos

Contient les informations du concepteur(s) et une petite déscription de l'application.

2.2.8 Onglet Fermeture

Il suffit de cliquer sur le bouton 'Quitter' pour quitter le système, et 'Oui pour Pour confirmer', sinon 'Annuler' comme dans la figure

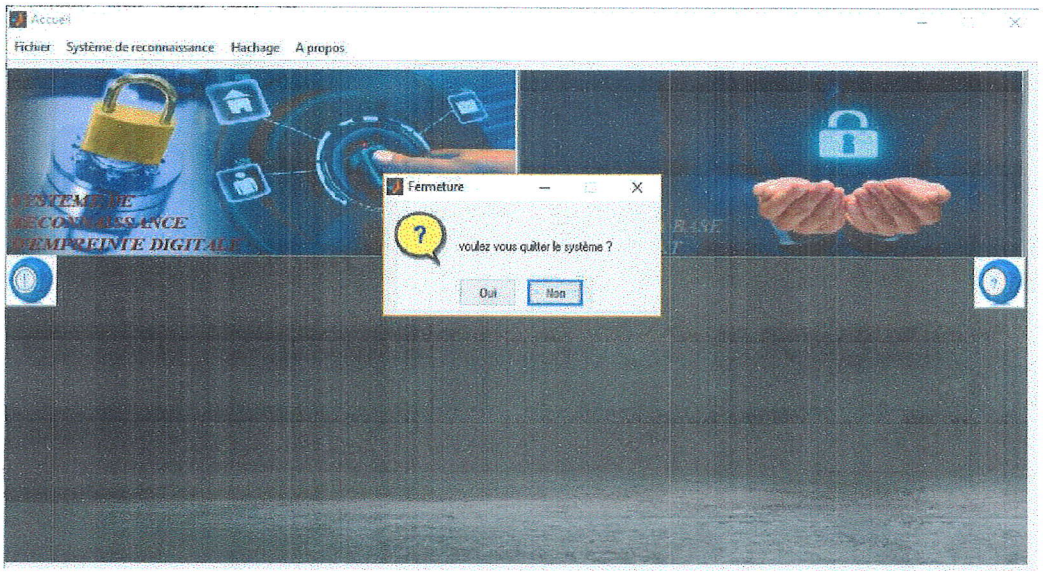


FIGURE 4.10 – Fermeture de système.

3 Tests et Evaluation

Cette section présente une évaluation générale de notre système, les expériences et les tests portant sur la performance et le temps d'exécution, sont réalisés sur un ordinteur ayant les caractéristiques suivantes :

- Mémoire vive(RAM) : 4 Go
- Processeur : Intel i5 2.0 GHz

Pour évaluer les systèmes de reconnaissances des empreintes digitales, plusiers bases de données standards sont utilisées. Parmi celles les plus courantes la base FVC2002 sur laquelle nous allons effectuer nos tests.

3.1 Présentation de la base de données

Nous effectuons les tests sur la base de données **FVC2002 - DB2** fournie publiquement dans le cadre de la compétition internationale pour la vérification d'empreintes digitales. Les images sont acquises avec un capteur optique d'une résolution de 569 dpi, générant des images de 560×296 pixels. Elle est composée de 80 images pour 10 individus et 8 échantillons par individu.

FIGURE 4.11 – Excemple d'images de la base.

3.2 Mesures de performance

3.2.1 Protocole utilisé pour le test d'identification

afin de calculer la bonne classification des personnes, on a pris la totalité de la base, 80 images dans la phase d'apprentissage (10 personne avec 8 acquisitions par personne). et on a pris 80 images de test à identifier (chaque image est comparée avec 80)

pour chaque image de test le système va calculer le score de similarité entre cette image et l'ensemble des images d'apprentissage.

Résultas expérimentales :

En utilisant ces scores, nous traçons l'authentique Et la distribution de l'imposteur pour des images sans attaques, avec rotation de 10°et avec bruit de 0,05 :

Les résultats obtenus des images sans attaques :

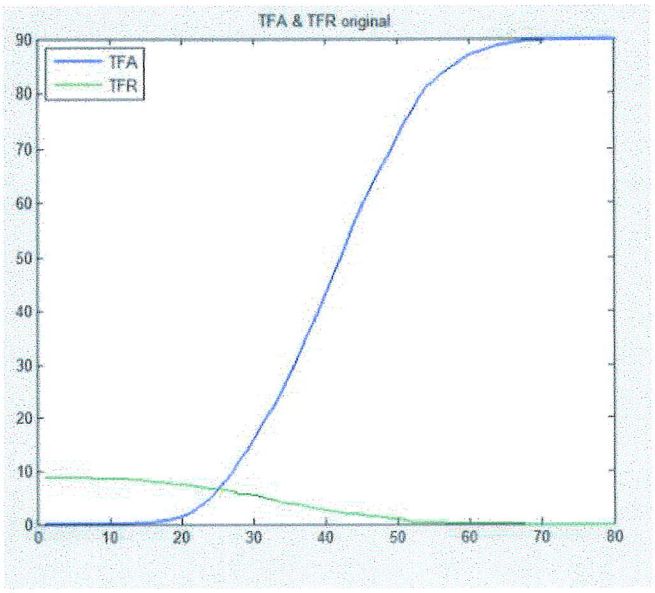


FIGURE 4.12 – TFA & TFR d'image sans attaque.

comme le montre la Fig. 4.12 un EER (taux d'erreur égal) inférieur à 1 % a été atteint sur les images sans attaques. qui est un bon taux indiquant l'éfficacité du système qui a aussi marqué les taux suivants :

- Taux d'identification : 85,2 %
- TFA : 6,4 %
- TFR : 8,4 %

Les résultats obtenus des images avec rotation de 10° :

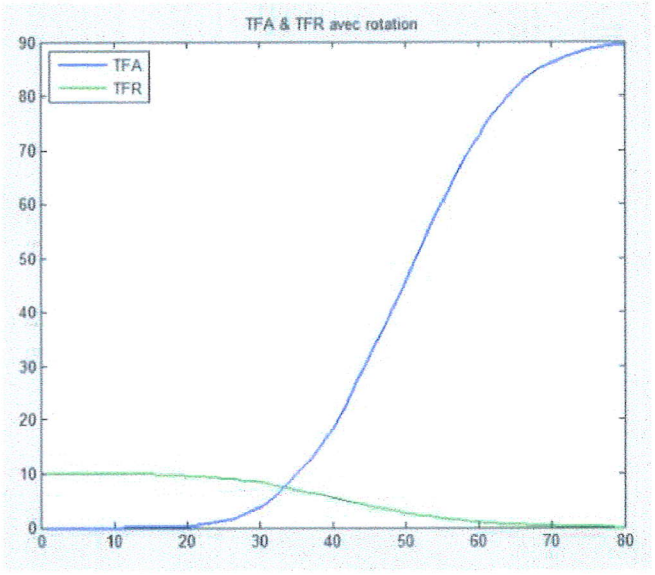


FIGURE 4.13 – TFA & TFR d'image avec rotation de 10°.

En appliquant une rotation de 10°les résultats sont aussi bien avec un EER (taux d'erreur égal) inférieur à 1 % a été aussi atteint avec l'applacation de cette attaque accéptable comme le montre la Fig. 4.13 avec les taux suivants :

- Taux d'identification : 82,5 %
- TFA : 8,4 %
- TFR : 9,1 %

Les résultats obtenus des images bruitée :

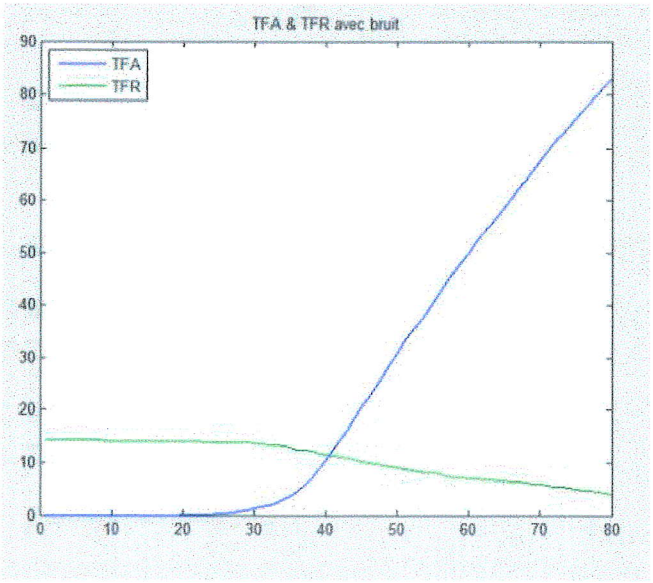


FIGURE 4.14 – TFA & TFR d'mage bruitée .

En ajoutant un bruit de 0,O5 qui est considéré comme une attques accéptable, on a marqué un EER (taux d'erreur égal) prèsuqe 1 % en appliquanr ce niveau de bruit comme le montre la Fig. 4.14 et les taux suivant :

- Taux d'identification : 81,9 %
- TFA : 8,7 %
- TFR : 9,4 %

Les résultats des différents tests sont représentés dans le tableau suivant :

	TAUX D'IDENTIFICATION	TAUX FAUSSE ACCEPTATION	TAUX FAUX DE REJET
Image originale	85,2 %	6,4 %	8,4 %
Rotation 10°	82,5 %	8,4 %	9,1 %
Bruit 0,05	81,9 %	8,7 %	9,4 %

4 Conclusion

Après une série de tests effectués sur notre système pour évaluer ses performances en utilisant des attaques accéptables, on a enregisté un taux de reconnaissance proche de 85 % et un taux d'erreur égale proche de 1 %. Nous pouvons dire

que les résultats obtenus sont satisfaisants et le système a prouvé son éfficacité et sa robustesse.

Conclusion et perspectives

Conclusion

Les fonctions de hachage perceptuel sont fortement inspirées des fonctions de hachage cryptographique. Comme ces dernières sont très sensibles au contenu binaire des données à hacher, les fonctions de hachage perceptuel sont proposées comme une solution alternative pour une principale application aux données multimédias et spécialement aux images. Une image numérique tel que l'image d'empreinte digitale peut subir différentes formes de transformations ou de manipulations qui peuvent affecter son contenu binaire et/ou visuel. Certaines applications peuvent avoir besoin d'appliquer certaines manipulations acceptables afin d'améliorer la qualité de l'image originale telles que la compression, le filtrage, ou même d'effectuer d'autres opérations permettant l'amélioration de l'image en question. Certaines applications peuvent également nécessiter une compression avec pertes pour satisfaire les contraintes de ressources sur la bande passante ou d'espace de stockage. Ces manipulations acceptables modifient uniquement les valeurs de pixels, qui se traduit par différents niveaux de distorsion visuelle de l'image, mais le contenu de l'image, qui porte la même information visuelle vers le récepteur, est encore conservé. L'authentification des images devrait se baser sur leurs contenus visuels et non pas sur leurs contenus binaires. Par conséquent, pour authentifier une image d'empreinte digitale, il faut tolérer des manipulations acceptables que pourrait subir une image. Les fonctions de hachage perceptuel sont des solutions potentielles dans ces cas-là permettant d'établir une "correspondance perceptuelle" entre l'image originale et l'image à authentifier.

L'objectif principale de ce mémoire a été de proposer une méthode de hachage perceptuel basée sur des caractéristiques extraites en basant sur le détecteur de SIFT et Harris afin d'extraire des caractéristiques stables qui seront ensuite utilisés pour générer la signature numérique (hach) en utilisant le Shape Context.

Afin de mener cette étude, nous étions convaincu dés le début que les fonctions de hachage cryptographique et les fonctions de hachage perceptuel se différent les unes des autres seulement dans la nature des données à hacher. Pour ces raisons, nous avons fixé les exigences que les signatures perceptuelles doivent vérifier. Les fonctions de hachage perceptuel basées sur le Shape Context doivent générer une signature :

- Courte : la signature doit être courte de l'ordre de quelques centaines de bits.
- avoir la même signature pour des données multimédia de même contenus visuels.
- Sécurisée et déscriminante : impossible de générer les données originales à partir de leurs signatures et en même temps avoir des signatures totalement

différentes pour des données multimédia n'ayant pas le même contenu visuel.

Afin de respecter notre modèle de système de hachage perceptuel, idéalement robuste et sécurisé, nous avons fixé l'objectif de cet analyse est de cerner les raisons menant à l'instabilité des signatures perceptuelles et non pas de chercher des caractéristiques stables pour générer des signatures stables.

Perspectives

Nous pensons que cette étude a juste permis de poser les bases d'une méthode d'authentification innovante, mais de nouvelles perspectives et de futurs challenges, en termes d'amélioration de la robustesse du système de hachage perceptuel, de sont envisageables sur différents plan. Plus précisément nous pensons aux perspectives suivants :

- Appliquer l'analyse théorique sur d'autres systèmes de hachage perceptuel existants.
- Prendre en compte d'autres types d'attaques pour analyser les signatures perceptuelles.
- Étudier la robustesse à d'autres types de caractéristiques extraites.
- Améliorer la robustesse de la méthode de hachage perceptuel proposée dans le chapitre 3 pour d'autres types de caractéristiques extraites et tester sa robustesse à d'autres types d'attaques.

Bibliographie

[1] S. Akrouf, "Une Approche Multimodale pour l'Identification du Locuteur", Thèse de doctorat, Université Ferhat Abbas Sétif, 2011.

[2] Alfred A. Manuel,'Element de MATLAB', Université de genève,15 Octobre 2004.

[3] K. Aloui, "Caractérisation du Cerveau Humain : Étude de la Faisabilité en Biométrie", Thèse de doctorat, École Nationale d'Ingénieurs de Tunis (ENIT), 2012.

[4] W. Zhao, R. Chellappa, P.J. Phillips and A.Rosenfeld, Face recognition : A literature survey , ACM Computing Surveys (CSUR), Volume 35, Issue 4, December 2003.

[5] W.A. Barrett, A survey of face recognition algorithms and testing results , Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers, pp. 301-305, 1997.

[6] A.K. Jain, A. Ross, and S. Pankanti, A prototype hand geometry-based verification system , in Proc. of 2nd Int l Conf. on Audio- and Video-based Biometric Person Authentication, pp. 166 171, March 1999.

[7] G.O. Williams, Iris Recognition Technology , IEEE Aerospace and Electronics Systems Magazine, Volume 12, Issue 4, pp. 23 -29, April 1997.

[8] R.P. Wildes, Iris Recognition : An Emerging Biometric Technology , Proceedings of the IEEE, Volume 85, Issue 9, pp. 1348 -1363, Sept. 1997.

[9] S.J. Vaughan-Nichols ; Voice authentication speaks to the marketplace , Computer, Volume : 37 , Issue 3, pp. 13-15, March 2004.

[10] B.H. Juang and T. Chen, The past, present, and future of speech processing , IEEE Signal Processing Magazine, Volume : 15 Issue : 3, pp. 24-48, May 1998.

[11]] L.L. Lee, T. Berger, and E. Aviczer, Reliable On-Line Human Signature Verification Systems , IEEE Trans. on PAMI, Vol. 18, No. 6, pp.643-647, June 1996.

[12] W.J. Babler, Embryologic Development of Epidermal Ridges and Their Configurations , Dermatoglyphics :Science in transition. Birth defects, New York, Wiley-Liss, pp. 95-112, 1991.

[13] N. Yager and A. Amin, "Fingerprint verification based on minutiae features : a review", Pattern Analysis and Applications, Vol. 7, No. 1, pp. 94-113, April 2004.

[14] H. Ailisto and M. Linholm, "A reviex of fingerprint image enhancement methods", International Journal of Image and Graphics, Vol. 3, No. 3, pp. 401-424, 2003.

[15] D. Maio and D. Maltoni, "Direct Gray-Scale Minutiae Detection In Fingerprints", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, pp. 27-40, 1997.

[16] Z. Bian, D. Zhang and W. Shu, "Knowledge-based fingerprint post-processing", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 16, No. 1, pp. 53-67, 2002.

[17] J.D. Stosz and L.A. Alyea, "Automated system for fingerprint authentication using pores and ridge structure", Proceedings of SPIE in Automatic Systems for the Identification and Inspection of Humans, Vol. 2277,pp. 210-223, October 1994.

[18] V. Soifer, V. Kotlyar, S. Khonina and R. Skidanov, "Fingerprint identification using the directions field", ICPR proceedings, Vol. 3, pp. 586-590, 1996.

[19] L.C. Ern and G. Sulong, "Fingerprint Classification Approaches : An Overview", International Symposium on Signal Processing and its Applications, Kuala Lumpur, Malaysia, 13-16 August, 2001.

[20] S. Pankanti, S. Prabhakar and, A.K. Jain, On the Individuality of Fingerprints , IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 8, pp. 1010-1025, August 2002.

[21] D.P. Mital and E.K. Teoh, "An automated Matching Technique for Fingerprint Identification", Proceedings of 22nd International Conference on Industrial Electronics, Control, and Instrumentation, Vol.2, pp. 806-911, 1996.

[22] Jain, L. Hong and R. Bolle, On-Line Fingerprint Verification , IEEE Transactions on PAMI, Vol. 19, No. 4, pp. 302-314, 1997.

[23] A.K. Jain, S. Prabhakar and S. Pankanti, "Twin Test : On Discriminability of Fingerprints", Proc. 3rd International Conference on Audio- and Video-Based Person Authentication,, pp. 211-216, Sweden, June 6-8, 2001.

[24] R. Bolle, Guide to Biometrics. 2004.

[25] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal, 40(03) :614–634, 2001.

[26] Cheng et X. Li : Partial encryption of compressed images and videos. IEEE Transactions on Signal Processing, 48(8) :2439–2451, 2000.

[27] Van Droogenbroeck : Partial encryption of images for real-time applications. In Fourth IEEE Signal Processing Symposium, pages 11–15, Hilvarenbeek, The Netherlands, April 2004.

[28] P. Tuyls, A.H.M. Akkermans, T.A.M. Kevenaar, G.J. Schrijen, A.M. Bazen, and R.N.J. Veldhuis. Practical biometric authentication with template protection. In International Conference on Audio- and Video-Based Biometric Person Authentication, volume 3546, pages 436–446, 2005.

[29] A. Arakala, J. Jeffers, and K. J. Horadam. Fuzzy extractors for minutiae-based fingerprint authentication. In 2nd International Conference on Biometrics, 2007.

[30] A. Juels and M. Sudan. A fuzzy vault scheme. In Proceedings of IEEE Internation Symposium on Information Theory, Lausanne, Switzerland. IEEE Press, 2002

[31] Cox, M. L. Miller et J. A. Bloom : Digital watermarking. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002. ISBN 1-55860-714-5.

[32] Monga : Perceptually Based Methods for Robust Image Hashing. Phd dissertation, University of Texas at Austin, 2005.

[33] G. Lowe : Object recognition from local scale-invariant features. In Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2, ICCV '99, pages 1150– 1157, Washington, DC, USA, 1999. IEEE Computer Society. ISBN 0-7695-0164-8.

[34] G. Lowe : Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2) :91–110, November 2004.

[35] C.Harris et M.Stephens. A combined corner and edge detector. In *In Proc. of fourth elvey vision conference. pages 147-151, 1988.*

[36] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 4, pp. 509–522, Apr. 2002.

[37] Jakubowski, M.H., Venkatesan, R. : Randomized Radon Transforms for Biometric Authentication via Fingerprint Hashing. In : Proceedings of the 2007 ACM workshop on Digital Rights Management (2007)

Table des figures

1.1 Exemples de différentes caractéristiques biométriques : empreinte digitale(a), visage(b), main(c), iris(d), empreinte vocale(e), signature(f). 5
1.2 Architecture d'un système biométrique 7
1.3 (a) Densités de scores (de similarité) pour un système biométrique idéal. (b) Densités de scores pour un système biométrique réel. 8
1.4 (a) Courbe ROC. (b) Courbes DET. 10
1.5 Illustration des points de fonctionnement sur une courbe des taux d'erreurs en fonction du seuil de décision. 11
1.6 Illustration du FRR et du FAR. [1] 12
1.7 Evolution du marché international de la biométrie. 13
1.8 Part de marché des differentes méthodes biométriques. 14

2.1 Caractéristique d'une empreinte digitale 18
2.2 Les différents types de munitie 19
2.3 Les trois principales classes d'empreinte : boucle(a), spire(b), arche(c). 19
2.4 Architecture d'un système biométrique basé sur l'empreinte 21
2.5 synopsis du pré-traitement des images d'empreintes digitales en vue de l'extraction des minuties 23
2.6 Les points de vulnérabilités d'un système d'empreintes digitales . . . 25
2.7 Principe du fuzzy vault 29
2.8 Problèmatique du tatouage des images. 31
2.9 Fonction de hachage cryptographique 32

3.1 Exemple qui illustre les exigences d'un hachage perceptuel dans le scénario d'authentification de contenu. Les signatures perceptuelles des images (b) et (a) doivent être égales et différentes de celle de l'image (c). [32] 36
3.2 Présentation des quatre étapes d'un système de hachage perceptuel. . 38
3.3 Sélection des caractéristiques les plus pertinentes 38
3.4 Méthode de hachage perceptuel proposée. 41
3.5 Construction de la matrice de pixels utilisée pour construire le vecteur de description SIFT. Le voisinage (encadré en rouge) est tourné relativement à la direction dominante du voisinage (segment rouge). La matrice utilisée pour construire le vecteur de description est de taille 16*16. 43
3.6 Construction du vecteur de description SIFT à partir de la matrice d'orientations de taille 16*16 tournée. 43

3.7 Calcul des images de différences de gaussiennes et détection de l'extremum et du minimum local. 44
3.8 Construction d'un descripteur SIFT. 45
3.9 Localisation des points d'intérêt SIFT sur l'image. 45
3.10 Les 3 cas de changements d'intensité considérés. 46
3.11 Détecteur de Harris - Afin de retrouver les points d'intérêt, le détecteur de Harris calcule pour chaque pixel, la matrice d'auto-corrélation à partir des deux composantes des vecteurs gradients de l'image. Ensuite, la matrice de réponse du détecteur est obtenue à partir de ces matrices. Enfin, les points d'intérêt, ici marqués d'une croix verte, sont localisés à partir de cette réponse. 47
3.12 Localisation des points d'intérêt sur l'image résultante de SIFT. . . . 48
3.13 Localisation des points d'intérêt SIFT-Harris sur l'image. 48
3.14 Diagramme de shape context original et du hachage de shape context proposé : RSCH. (a) Les contextes de forme originale. (b) le hachage de shape context radiale. 49
3.15 Schéma de la méthode proposée. 51
3.16 Conversion de la représentation des minuties en vecteur entier. 52
3.17 Réglement des détails d'orientation minutieux. 53
3.18 Comparaison entre la requête et le modèle inscrit. 54

4.1 Interface principale. 57
4.2 Hachage perceptuel à base shape context. 57
4.3 Présentation 58
4.4 Rotation 58
4.5 Bruit 59
4.6 performance 59
4.7 Onglet Onrollemnt. 60
4.8 Onglet Identification. 60
4.9 Phase d'authentification. 61
4.10 Fermeture de système. 62
4.11 Excemple d'images de la base. 63
4.12 TFA & TFR d'image sans attaque. 64
4.13 TFA & TFR d'image avec rotation de 10°. 64
4.14 TFA & TFR d'mage bruitée 65

Liste des tableaux

3.1 Manipulations acceptables vs Manipulations malveillantes. 37