Ministère de l'enseignement
supérieur et de la recherche scientifique
Université de M'sila



Faculté des Mathématiques et de l'Informatique
Département d'informatique

Mémoire de fin d'études
Pour l'obtention du diplôme d'ingénieur d'état en Informatique
Spécialité : Systèmes d'informations avancés

Thème

Système Automatique d'Identification par Empreinte Digitale (AFIS)

Réalisé par :

- ARIHIR Mohamed
- AROUSSI Noureddine
- DJENIDI Zineddine

Dirigé par :

Mr : BELHADJ Foudil

Promotion Juin 2010 / 2011.

Remerciement

Tout d'abord, nous présentons nos remerciements les plus sincères à notre encadreur Mr ***Belhadj Foudil****, pour sa disponibilité et son encadrement, ainsi que son soutien et ses conseils tout au long de l'année. C'est grâce à-il que nous avons pu mener à bien ce travail.*

Nous tenons ensuite à remercier chacun des membres du jury pour nous avoir fait l'insigne honneur d'accepter d'examiner notre travail.

Nos remerciements et notre et notre gratitude vont aux enseignants du département informatique.

Nous remercions également toute personne ayant contribué à notre éducation et notre formation.

Dédicaces

Nous dédions ce modeste travail:
À nos chers parents, pour leurs soutiens, leurs patiences et leurs amours
À toutes nos familles
À tous nos amis
Et à tous nos enseignants

Mohamed, Noureddine, Zineddine

Résumé

L'identification des personnes est une tâche qui suscite un intérêt grandissant dans la société cybernétique d'aujourd'hui. Le but étant de restreindre les accès aussi bien physiques que logiques (web sécurisé, accès distant). Le but de ce travail est d'arriver à un système d'identification automatique d'empreinte digitale robuste en terme de taux d'erreur (FAR, FRR) et en terme de temps de matching en se basant sur un algorithme qui fonctionne en temps linéaire. Un prétraitement de l'empreinte est nécessaire pour l'extraction des caractéristiques objets de matching.

Mot clé : Identification, Empreinte digitale, Matching, Temps linéaire.

Abstract

The identification of a person is a task which arouses a growing interest in the today cybernetic company. The objective is to restrict the accesses as well physical as logical (safe Web, distant access). The objective of this work is to arrive at an automatic fingerprint identification system with acceptable error rate (FAR, FRR) and matching time based on a linear time algorithm. A preprocessing of the fingerprint is necessary for the extraction of the characteristics included in matching step.

Key word : Identification, Fingerprint, Matching, Linear time.

ملخص

تحديد هوية الأشخاص عملية تثير اهتماما متزايدا في الشركات المعلوماتية اليوم، والهدف هو الحد من الوصول الى الموارد، مادية كانت أو غير ذلك. الهدف من هذا العمل هو التوصل الى نظام اتوماتيكي للتعرف على البصمة مع تحديد نسبة خطأ مقبولة (معدل القبول الخاطئ (FAR) معدل الرفض الخاطئ (FRR)) والزمن المستغرق للمقارنة استنادا إلى معادلة زمنية خطية . المعالجة الأولية للبصمة ضرورية لاستخلاص خصائصها اللازمة في عملية المطابقة.

مفاتيح: تحديد هوية الأشخاص، البصمة، المطابقة، معادلة زمنية خطية.

Sommaire

Liste des tableaux..... IV
Liste des figures VI
Liste des abréviations..... VII
Introduction générale 1

Chapitre I • **La biométrie**

I.1. Introduction 5
I.2. Les domaines d'application de la biométrie 6
I.2.1. Contrôle d'accès 6
I.2.2. Authentification des transactions 6
I.2.3. Répression 6
I.2.4. Personnalisation 7
I.3. Les modes de système biométrique 7
I.4. Principe de fonctionnement 8
I.4.1. Enrôlement 9
I.4.2. Contrôle (vérification/identification) 9
I.5. Les différents techniques biométriques 10
I.5.1. Les techniques biologiques 10
a). L'analyse ADN 10
b). Les groupes sanguins 11
c). Les cheveux et les poils 11
I.5.2. Les techniques comportementale 11
a). La signature dynamique 11
b). La dynamique de la frappe (au clavier) 12
I.5.3. Les techniques morphologiques (physiologique) 12
a). L'empreintes digitales 12
b). La géométrie de la main 13
c). L'iris 13
d). La rétine 14
e). La visage 14
f). La voix 15
I.6. Comparaison et évaluation dans solution biométrique 15
I.6.1. Critères de comparaison 16
I.6.2. Les erreurs de système biométrique 17

a). Les erreurs de système de vérification 17
b). Les erreurs d'un système d'identification 20
I.6.3. Évaluation des performances 22
•Évaluation de technologie 22
•Évaluation de scénario 22
•Évaluation opérationnelle 23
I.7. Conclusion 23

Chapitre II • **Identification automatique d'empreinte digitale**

II.1. Historique 25
II.2. Caractéristiques des empreintes Digitales 25
II.3. Architecture d'un système d'empreinte digitales 27
a). L'enrôlement 27
b). Vérification 28
c). Identification 28
II.4. Matching d'empreinte digitale 29
II.4.1. Image 29
II.4.2. Minuties 30
II.4.2.1. Représentation des minuties 30
II.4.2.2. extraction des modèle biométrique 31
a). La méthode classique 32
I. Amélioration 32
I.1. Normalisation 33
I.2. Orientation de crête 33
I.3. Fréquence de crête 34
I.4. Le masque de région 36
I.5. Filtre de Gabor 37
II. Binarisation 38
III. Squelettisation 40
IV. Extraction de minuties 42
V. Extraction de point core 43
VI. Post-traitement et filtrage de minuties 45
b). La méthode directe 47
II.4.3. Description de texture 48
II.5. Conclusion 49

Chapitre III • **Algorithmes de matching**

III.1. Introduction 51
III.2. Extraction de point core 51
III.3. Matching 52
III.3.1. La représentation de vecteur des voisins les plus proches (NNV) .. 52
III.3.2. Matching de NNV 53
III.3.3. Incorporation de matching NNV en arbre binaire 54
III.3.4. Analyse de complexité de temps 56
III.4. Critiques et solution 56

III.4.1. Critiques 56
III.4.2. Solution 57
III.5. Conclusion 60

Chapitre IV • **Conception, mise en œuvre & évaluation**

IV.1. Introduction 62
IV.2. conception 63
IV.2.1. Analyse des besoins 63
IV.2.2. Spécification fonctionnelle 63
IV.2.2.1. Les acteurs de système 63
IV.2.2.2. Détermination de cas d'utilisation 64
IV.2.2.3 Diagramme de séquence de cas d'utilisation « Matching » 65
IV.2.2.4. Diagramme de séquence de cas d'utilisation « Statistique » 67
IV.3. Mise en œuvre 68
IV.3.1. Outils de développement 68
● Microsoft Visual Studio 2010 68
IV.3.2. Description des différentes classes 69
V.3.3. Présentation de l'application 73
IV.4. Evaluation 76
●Commentaires sur les résultats 79
IV.5. Conclusion 79
Conclusion générale 80
Bibliographie 81

Liste des Tableaux

Tableau	Titre	Page
Tableau IV.1	Les acteurs du système……..	64
Tableau IV.2	Les principales classes de notre système……..	69
Tableau IV.3	Les différents paramètres de notre système……..	72
Tableau IV.4	Résultat du test sur la base FVC2002 (DB1, DB2)…...	76
Tableau IV.5	Résultat du test sur la base FVC2004(DB1, DB2)……	76

Liste des figures

FIGURE	TITRE	PAGE
Figure I.1	Le marché mondial de la biométrie (2009-2014)……..	5
Figure I.2	Diagramme des tâches d'inscription, de vérification et d'identification……..	8
Figure I.3	Fonctionnement d'un système biométrique……..	9
Figure I.4	Part du marché des différentes méthodes biométriques……..	10
Figure I.5	Comparaison des principales techniques biométriques……	17
Figure I.6	Lien entre FMR et FNMR……..	19
Figure I.7	Un exemple des courbes de FMR (t) et de FNMR (t), où les points correspondant à EER, à ZeroFNMR, et à ZeroFMR sont accentués……	20
Figure I.8	Les points typiques de différentes applications montrées sur un ROC courbent……	21
Figure II.1	Caractéristiques Des Empreints Digitales……..	25
Figure II.2	Les différents types de minutie (crêtes en noir)……	26
Figure II.3	Les classes d'empreinte, boucle(a), spirale(b), arc(c), double boucle(d)……	26
Figure II.4	La phase d'enrôlement……..	27
Figure II.5	Le mode de vérification……..	28
Figure II.6	Le mode d'identification……	29

Figure II.7	Les caractéristiques principales des minuties……..	31
Figure II.8	représentation d'une empreinte par sa carte de minuties……	31
Figure II.9	Processus d'augmentation d'une image……..	32
Figure II.10	Fenêtre d'orientation et x-signature……..	35
Figure II.11	Région des crêtes et pas de crête.. ……..	36
Figure II.12	k-plus-proches voisins. ……	37
Figure II.13	L'image après l'étape de binarisation. ……..	38
Figure II.14	Squelettisation de l'image d'empreinte……	41
Figure II.15	Détection des minuties……..	42
Figure II.16	Les types de connexion d'une minutie. ……	43
Figure II.17	Les types des fausses minuties. ……	46
Figure II.18	Méthode d'extraction directe. ……	48
Figure III.1	Point core. ……..	52
Figure III.2	La représentation de vecteur des voisins les plus proches….	53
Figure III.3	Exemple d'un arbre binaire. ……..	54
Figure III.4	Arbre binaire d'une empreinte digitale. ……..	54
Figure III.5	Points d'échantillonnage organisés dans un modèle circulaire autour d'une minutie……	58
Figure IV.1	Diagramme de cas d'utilisation de système. ……..	64
Figure IV.2	Diagramme de séquence de cas d'utilisation « Matching »…..	66
Figure IV.3	Diagramme de séquence de cas d'utilisation « statistiques ».	67
Figure IV.4	La forme principale.……..	73
Figure IV.5	Volet de détails.……..	74
Figure IV.6	Teste et expérimentations.……..	75
Figure IV.7	ROC de FVC2002(DB1, DB2) de l'algorithme BT.……..	77
Figure IV.8	ROC de FVC2002(DB1, DB2) de l'algorithme MBT……	77
Figure IV.9	ROC de FVC2004(DB1, DB2) de l'algorithme BT……	78
Figure IV.10	ROC de FVC2004(DB1, DB2) de l'algorithme MBT……..	78

Liste des Abréviation

IBG	International Biométrique Group.
CNIL	Commission Nationale de l'Informatique et des Libertés.
PIN	Personal Identfication Number.
FAR	False Acceptaion Rate.
FRR	False Reject Rate.
EER	Equal Error Rate.
FTC	Failure To Capture.
FTE	Failure To Enroll.
FTM	Failure To Match.
NNV	Nearest Neighbor Vector.
BT	Binary Tree.
MBT	Modified Binary Tree.
AFIS	Autoamted Fingerprint Identification System.
UML	Unified Modeling Language.
ROC	Receiver Operating Characteristic.
BDD	Base De Données.
FVC	Fingerprint Verification Competition.

Introduction générale

Actuellement, l'accès sécurisé est un domaine très important grâce aux réseaux informatiques présents dans tous les domaines. En général, on distingue deux catégories d'identification : la reconnaissance ou l'identification de l'utilisateur qui consiste à associer une identité à une personne. En d'autres termes, elle répond à des questions du type : *«Qui suis-je?»,* et l'authentification ou la vérification de l'utilisateur qui permet de confirmer une identité proclamée ou de l'infirmer. Le système doit alors répondre à la question suivante : *«Suis-je réellement la personne que je suis en train de proclamer?»*. L'application de reconnaissance est utilisée par la police et les services d'immigration dans les aéroports, ainsi que dans la recherche de bases de données criminelles. Elle figure aussi dans les applications civiles où l'authentification des cartes de crédit, de permis de conduire et des passeports est de plus en plus courante. En revanche, l'application d'authentification est souvent utilisée dans tous les domaines nécessitant un accès contrôlé tels que celui des applications bancaires, les endroits hautement sécurisés comme les sièges de gouvernement, et de site Internet.

La méthode d'identification personnelle traditionnelle, tel que le mot de passe, ou carte d'identité, est échouée à des attaques et est incapable de satisfaire le besoin de sécurité du système information interconnecté. Le mot de passe risque d'être oublié et la carte d'identité peut être perdu, mal placé. Différentes techniques d'authentification basées sur les caractéristiques physiques de l'utilisateur ont été développées par la recherche scientifique. Ces techniques généralement appelées méthodes biométriques. L'avantage principal de ce qu'on appelle mot de passe biométrique est lié au fait qu'il ne pourrait pas être volé, oublié ou transmis à une autre personne. En effet, chaque membre de la population possède sa propre caractéristique biométrique, et elle est relativement stable. La biométrie s'adresse l'identification automatique à individuel

basée sur ses traits physiologique, par exemple, l'odeur, le sang, l'urine, l'ADN, l'empreinte, la forme de la main, le traits du visage ou comportementaux, par exemple, la dynamique de la signature, la façon d'utiliser un clavier d'ordinateur.

Dans notre travail, nous sommes intéressés par la modalité d'empreintes digitales pour réaliser un système Automatique d'identification d'empreinte digitale.
L'empreinte digitale est unique et immuable, elle est formée de crêtes (ridges) et de vallées (valleys). Les minuties, qui représentent des discontinuités locales dans le flot des crêtes, fournissent les caractéristiques qui sont utilisées pour l'identification. Des détails, comme la location ou l'orientation de la minutie, sont fournis pendant la phase d'extraction des minuties.

Cependant, l'image d'empreintes est rarement en parfait état, elle est souvent entachée de bruit empêtrant l'extraction d'une carte de minuties correcte. Cette dégradation peut être le cumul de plusieurs facteurs : dispositif de prise de vue, état de la peau (sècheresse, humidité, plis, blessures, etc.), pression du doigt, etc. Une étape critique dans ce travail, consistera d'abord à une extraction fiable des minuties à partir de l'image d'empreintes. Il sera question d'implémenter une série de techniques, alliant plusieurs domaines du traitement de l'information pour aboutir à une carte de minuties aussi fiable que possible.

Un autre point, un protocole de vérification devrait être défini dans le but de comparer deux empreintes et d'attester de leur similarité ou de leur différence. Ici, les protocoles sont indéterministes, ils sont fondés sur des comparaisons probabilistes impliquant une fiabilité variable se traduisant dans la pratique par de possibles rejets et acceptations à tort. En utilisant les minuties comme attribut de comparaison, cette dernière sera traitée comme un problème d'appariement de points "Point Pattern Matching ". Actuellement ce problème et considéré comme un problème d'optimisation. Dans notre cas nous avons étudié un algorithme fonctionnant en temps linéaire (alignement des minuties de l'empreinte dans une arbre binaire) et nous avons essayé de l'optimiser en terme de résultats de taux de fausse acceptation et de taux de faux rejets.

Des expériences sur des images d'empreintes réelles doivent être menées afin d'évaluer les méthodes étudiées. Tout cela est pour but d'utiliser des métriques de performance (globalement : le taux de faux rejets et le taux de fausses acceptations)

pour constater le comportement du système. Il sera question de comparer la performance en utilisant un noyau de minutie simple et un noyau issu. Cette étape est déterminante pour l'élaboration d'une politique de vérification globale.

Tous ces points seront traités en substance dans les chapitres suivants :

Le premier chapitre est dédié à la présentation de la biométrie. Nous y expliquerons ainsi l'architecture d'un système de vérification et d'identification biométrique, et les différentes techniques biométriques.

Le deuxième chapitre est dédié pour la modalité d'empreinte digitale, où nous expliquerons l'architecture d'un système de vérification et d'identification par empreinte digitale, et les caractéristiques immuables de l'empreinte, ainsi les différentes étapes pour les extraire.

Le troisième chapitre présente un algorithme de matching en un temps lineaire. Nous présentons ses points faibles puis nous proposons des solutions d'amélioration.

Le quatrième chapitre sert à exprimer le modèle utilisé pour la conception et la mise en œuvre de notre système et les différentes étapes de conception.

Nous terminons le présent de la mémoire par une conclusion générale.

Chapitre

I

La biométrie

I.1. Introduction

Dans un monde de plus en plus numérique, l'authentification personnelle fiable est devenue une importance d'interface de l'activité homme-ordinateur, sécurité nationale, commerce électronique, et accès aux réseaux informatiques où l'établissement de l'identité d'une personne est essentiel. Les mesures de sécurité existantes se fondent sur des approches basées sur la connaissance (mots de passe, Code PIN…) ou des approches basées sur la possession (carte d'identité, passeports, Carte à Puce…) pour contrôler l'accès aux espaces physiques et virtuels. Jusqu' à présent, de telles méthodes ne sont pas très fiables. Des techniques telles que des badges et des cartes d'accès peuvent être partagées ou volées. Des mots de passe et des Codes PIN peuvent être volés électroniquement. En outre, ils ne peuvent pas différencier entre l'utilisateur autorisé et une personne ayant accès aux techniques ou à la connaissance. La biométrie telle que l'empreinte digitale, le visage et la copie de voix offre des moyens fiable de l'authentification personnelle qui peuvent aborder ces problèmes et gagnent l'acceptation de citoyen et de gouvernement.

Selon l'international biométrique group(IBG), la biométrie a généré un marché de plus de 3422 millions $US en 2009 est devrait dépasser les 9368 millions $US en 2014 (voir figure I.1).

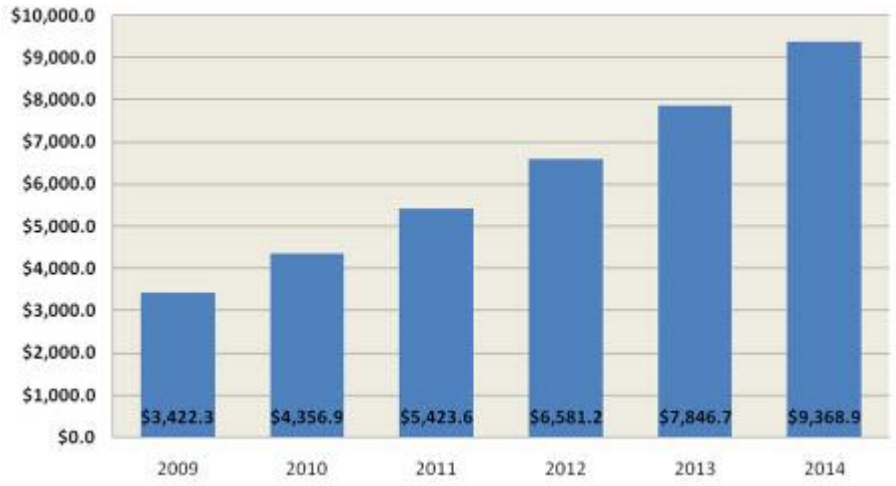


Figure I.1. Le marché mondial de la biométrie (2009-2014). [1].

La biométrie est un ensemble des méthodes automatisées d'identifier une personne basée sur une caractéristique physiologique ou comportementale. Les technologies biométriques deviennent la base d'un choix étendu fortement fiable de l'identification et des solutions personnelles de vérification. Les exemples des caractéristiques physiologiques incluent la main ou images de doigt, caractéristiques faciales, et l'identification d'iris. Les caractéristiques comportementales sont des traits qui sont appris ou acquis. La vérification de signature dynamique, la vérification de haut-parleur, et la dynamique de frappe sont des exemples des caractéristiques comportementales. [Joh-06]

I.2. Les domaines d'application de la biométrie

On peut distinguer quatre grands types d'applications de la biométrie : le *contrôle d'accès* (Access control), l'*authentification des transactions* (transaction authentication), la *répression* (lawenforcement) et la *personnalisation* (personalisation).

I.2.1. Contrôle d'accès

Le contrôle d'accès peut être lui-même subdivisé en deux sous catégories : le *contrôle d'accès physique* et le *contrôle d'accès virtuel*.

On parle de contrôle d'accès physique lorsqu'un utilisateur cherche à accéder à un lieu sécurisé.

On parle de contrôle d'accès virtuel dans le cas où un utilisateur cherche à accéder à une ressource ou un service.

I.2.2. Authentification des transactions

L'authentification des transactions représente un marché gigantesque puisqu'il englobe aussi bien le retrait d'argent au guichet des banques, et les paiements par cartes bancaires, les transferts de fond, les paiements effectués à distance par téléphone ou sur internet, *etc*….

I.2.3. Répression

Une des applications les plus immédiates de la biométrie à la répression est la criminologie. La reconnaissance d'empreintes digitales est l'exemple le plus connu. Elle fut acceptée dès le début du XXe siècle comme moyen d'identifier formellement un

individu et son utilisation s'est rapidement répandue. Il existe aussi des applications dans le domaine judiciaire.

T-Netix propose ainsi des solutions pour le suivi des individus en liberté surveillée en combinant technologies de l'internet et de reconnaissance du locuteur.

I.2.4. Personnalisation

Les technologies biométriques peuvent être aussi utilisées afin de personnaliser les appareils que nous utilisons tous les jours. Cette application de la biométrie apporte un plus grand confort d'utilisation. Afin de personnaliser les réglages de sa voiture, Siemens propose par exemple d'utiliser la reconnaissance des empreintes digitales. Une fois l'utilisateur identifié, donc la voiture ajuste automatiquement les sièges, les rétroviseurs, la climatisation, *etc*….

I.3. Les modes de système biométrique

Une question importante dans la conception d'un système pratique biométrique est de déterminer comment un individu est reconnu. Selon le contexte de l'application, un système biométrique peut être appelé soit un système de vérification ou d'un système d'identification :

- **Un système de vérification :** authentifie l'identité d'une personne en comparant la caractéristique biométrique capturée à ses propres calibres biométriques préenregistrés dans le système. Il conduit la comparaison un-a-un pour déterminer si l'identité réclamée par l'individu est vraie. Un système de vérification rejette ou accepte la réclamation soumise de l'identité « Suis-je bien la personne que je prétends être ? ».
- **Un système d'identification :** identifie un individu en recherchant dans la base de données entière de calibre pour le matching. Il conduit un-à-beaucoup de comparaisons pour établir l'identité de l'individu. Dans un système d'identification, le système établit l'identité d'un sujet (ou les échouer si le sujet n'est pas inscrit dans la base de données de système) sans soumis devant réclamer une identité (qui suis je ?).

Le terme d'authentification est également fréquemment utilisée dans le domaine biométrique, parfois comme synonyme pour la vérification ; réellement, dans la langue de

la technologie de l'information, l'authentification d'un utilisateur signifie pour faire le système savoir l'identité d'utilisateur indépendamment du mode (vérification ou identification). [dav-03]

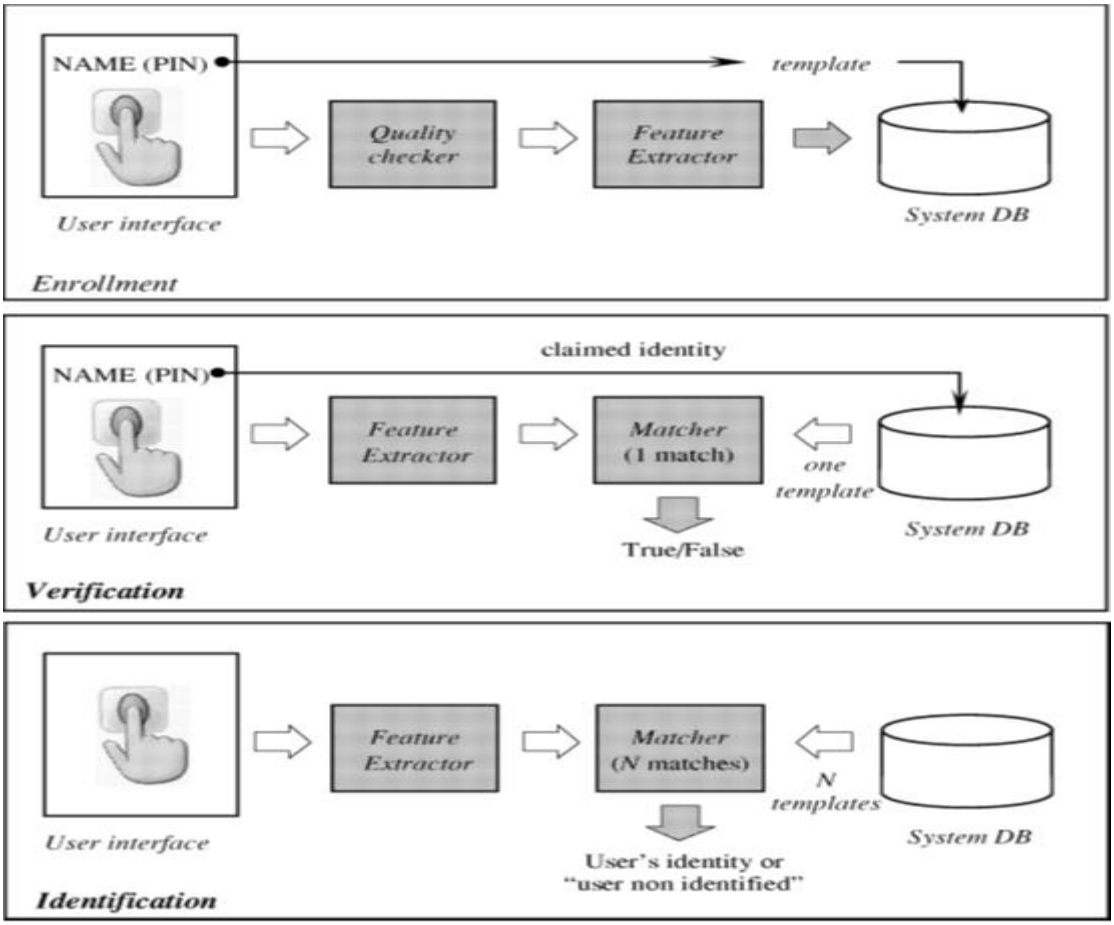


Figure I.2. Diagramme des tâches d'inscription, de vérification et d'identification.

I.4. Principe de fonctionnement

Le cycle macroscopie d'un processus d'identification biométrique se décompose en deux grandes catégories l'enrôlement et le contrôle.



Figure I.3. Fonctionnement d'un système biométrique. [1]

I.4.1. Enrôlement

L'enrôlement des personnes est la phase initiale de création de gabarit biométrique et de son stockage en liaison avec une identité déclarée. Les caractéristiques physiques sont transformées en un gabarit représentatif de la personne et propre au système de reconnaissance.

Durant cette phase, des données additionnelles propres à la personne qui s'enrôle sont enregistrées comme par exemple ses nom et prénom et un identifiant personnel(PIN). Cette étape n'est pas effectuée qu'une seule fois.

I.4.2. Contrôle (vérification/identification)

C'est l'action de contrôle des données d'une personne afin de procéder à la vérification de son identité déclarée ou dans une investigation à la recherche de l'identité de cette personne. Cette étape se déroule à chaque fois qu'une personne se présente devant le système.

A ces deux processus s'ajoutent souvent les deux processus suivants :

- **Le rafraîchissement ou l'actualisation :** le système biométrique peut périodiquement corriger le gabarit de référence lors d'un contrôle de façon à prendre en compte des évolutions des données personnelles de la personne, en particulier pour les systèmes de contrôle.
- **La fin de vie :** le gabarit et autres données de références propres à la personne sont détruites pour prendre en compte la suppression de la personne du système de contrôle.

I.5. Les différents techniques biométriques

Les empreintes digitales continuent à être la principale technologie biométrique en terme de part de marché, près de 50% du chiffre d'affaires total (hors applications judiciaires). La reconnaissance du visage avec 12% 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. Toutes ces méthodes sont classées on trois catégories: les techniques biologique, les techniques comportementale et les techniques morphologique. [1]

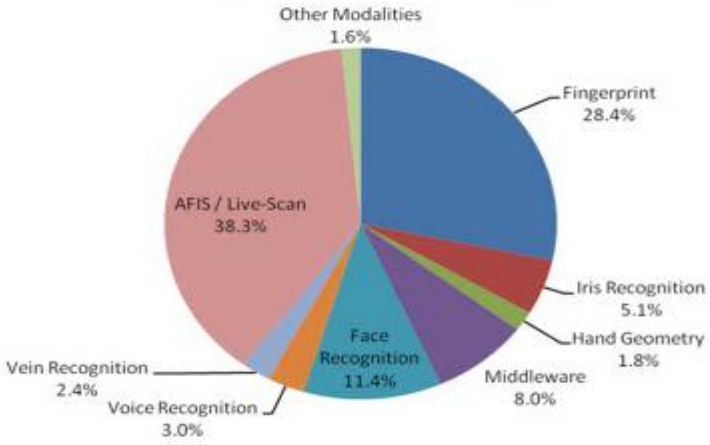


Figure I.4. Part du marché des différentes méthodes biométriques.

I.5.1. Les techniques biologiques

Ces systèmes portant sur l'analyse des empreintes génétiques (ADN, groupes sanguins, cheveux et les poils,...). Leur fonctionnement est expliqué dans ce qui suit :

a). L'analyse ADN : l'analyse des empreintes génétiques est devenue en quelques années l'un des outils majeurs de la criminalistique, la science de l'identification des indices matériels. L'analyse de l'ADN est couramment utilisée en criminologie pour identifier une personne à partir d'un morceau de peau, d'un cheveu ou d'une goutte de sang. Souvent les échantillons d'ADN trouvés sur le lieu du crime sont trop infimes pour être analysés. Mais il existe des appareils d'amplification en chaîne par

polymérase (PCR) qui utilisent le même procédé naturel que l'ADN pour recopier et amplifier. Ils procurent ainsi aux criminalistes des brins d'ADN répliqués et exploitables. [1]

b). Les groupes sanguins : Karl Landsteiner, pathologiste américain d'origine autrichienne, découvrit pour la première fois l'existence de groupes sanguins chez l'être humain en 1900. Il mit en évidence le système ABO, dont le premier test de reconnaissance fut développé en 1915 puis amélioré jusqu'en 1931. Landsteiner découvrit également les groupes P et MN ainsi que le facteur rhésus et fut lauréat du prix Nobel pour ses travaux en 1930. Ses travaux furent par la suite utilisés à des fins judiciaires, premier exemple d'application directe de travaux scientifiques à la criminalistique. Cependant, il est à noter que l'utilisation de ces marqueurs biologiques a plus d'intérêt dans l'exclusion d'une personne que dans l'identification de celle-ci. [1]

c). Les cheveux et les poils : Rudolf Virchow, médecin et anthropologue allemand, étudia pour la première fois en 1869 des poils humains. En 1910, V. Balthazard, professeur de médecine à la Sorbonne, et M. Lambert publièrent la première étude approfondie. En effet, l'examen de cheveux ou de poils permet d'identifier certaines caractéristiques propres à chaque individu, et d'apporter des éléments d'identification intéressants**.** [1]

I.5.2. Les Techniques Comportementale

Cette technique consiste à observer et analyser le comportement des individus dans des actions répétitives et usuelles tel que :

a). La signature dynamique : chaque personne a un style d'écriture unique. On peut donc définir, à partir de la signature d'une personne, un modèle qui pourra être employé pour effectuer une identification. De plus, la signature est utilisée dans beaucoup de pays comme élément juridique ou administratif. Elle permet de justifier de la bonne foi d'une personne ou de la confondre devant des documents signés. Le grand avantage des systèmes biométriques à base de signature réside dans la reconnaissance de cette méthode comme une forme acceptable juridiquement pour

l'identification des personnes. Cependant, en raison des grandes variations de signature pour une même personne, pour des systèmes tant à base d'analyse statique que dynamique, il est difficile d'atteindre une très haute exactitude d'identification. [1]

b). La dynamique de la frappe (au clavier) : la dynamique de la frappe est propre à chaque individu. Il s'agit en quelque sorte de la graphologie des temps modernes car nous écrivons plus souvent avec un clavier qu'avec un stylo. Les éléments analysés sont : vitesse de frappe, suite de lettre, temps de frappe, pauses... [1]

I.5.3. Les Techniques Morphologiques (Physiologique)

Cette technique s'attache à quantifier les caractéristiques corporelles et aussi, cette approche est la plus utilisé dans la pratique vu sa précision et la facilite de son utilisation. Nous allons présenter dans cette section plusieurs exemples, leurs avantages et inconvénients.

a). l'empreinte digitale : 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 préfèrera 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. [1]

b). La géométrie de la main: ce type de mesure biométrique est l'un des plus répandus, notamment aux Etats Unis. Cela consiste à mesurer plusieurs caractéristiques de la main (jusqu'à 90) tel que la forme de la main, longueur et largeur des doigts, formes des articulations, longueurs inter-articulations, etc... La technologie associée à cela est principalement de l'imagerie infrarouge; d'une façon générale, le système présente des FAR (False Acceptation Rate) assez élevés, surtout entre personnes de la même famille ou bien encore des jumeaux. [1]

c). L'Iris : pour les deux techniques suivantes, il faut tout d'abord faire la distinction entre l'iris et la rétine. Autrement dit, l'étude de l'iris va se porter sur la partie de l'œil. En ce qui concerne l'iris, l'individu se place en face du capteur (caméra CCD/CMOS) qui scanne son iris. Celui-ci représente quelque chose de très intéressant pour la biométrie car il est à la fois toujours différent (même entre jumeaux, entre l'œil gauche et le droit, etc...), indépendant du code génétique de l'individu, et très difficilement falsifiable. En effet, l'iris présente une quasi-infinité de points caractéristiques (que certains comparent en nombre à ceux de l'ADN), qui ne varient pratiquement pas pendant la vie d'une personne contrairement *à la couleur de l'iris* qui elle peut changer. Mais cela n'a aucune influence car les images d'iris obtenues par les capteurs sont en noir et blanc. Le seul problème de cette technique est liée à la mesure en elle-même qui peut être source d'erreurs ou de problèmes. Ainsi, on peut quasiment dire que le nombre de problèmes rencontrés lors de cette mesure augmente proportionnellement avec la distance entre l'œil et la caméra.

D'autres problèmes se posent à cause des reflets (nécessité d'avoir un éclairage restreint et maitrisé), et lors de la détection de faux yeux (photos) et autres fraudes. Pour ces dernières, on peut faire appel à certaines caractéristiques dynamiques de l'œil qui prouveront son authenticité : réactivité de la pupille (dilatation/rétraction) par rapport à la quantité de lumière, étude de l'iris dans l'infrarouge et l'ultraviolet, etc... [1]

d). La rétine : cette mesure biométrique est plus ancienne que celle utilisant l'iris, mais, elle a été moins bien acceptée par le public et les utilisateurs, sans doute à cause de son caractère trop contraignant : la mesure doit s'effectuer à très faible distance du capteur (quelques centimètres) qui effectue ensuite un balayage de la rétine. Il est physiquement impossible d'effectuer une mesure rétinienne à une distance de 30cm ou plus sur un sujet mobile comme on peut le voir dans certains films. Cette méthode requiert des sujets coopératifs et entrainés.

Pourtant cette technique semble être tout aussi fiable que celle de l'iris; elle se base sur le fait que le schéma et le dessin formé par les vaisseaux sanguins de la rétine (la paroi interne et opposée de l'œil) est unique pour chaque individu, différent entre jumeaux et assez stable durant la vie de la personne. La mesure peut ainsi fournir jusqu'à 400 points caractéristique du sujet, que l'on peut comparer aux 30 à 40 points fournis par une empreinte digitale!

En conclusion, la mesure rétinienne est la plus difficile à utiliser mais également la plus dure à contrefaire. [1]

e). Le Visage : il s'agit ici de faire une photographie plus ou moins évoluée pour en extraire un ensemble de facteurs qui se veulent propres à chaque individu. Ces facteurs sont choisis pour leur forte invariabilité et concernent des zones du visage tel que le haut des joues, les coins de la bouche, etc... On évitera d'autre part les types de coiffures, les zones occupées par des cheveux en général ou toute zone sujette à modification durant la vie de la personne.

Il existe plusieurs variantes de la technologie de reconnaissance du visage. La première est développée et supportée par le MIT et se nomme "Eigenface". Elle consiste à décomposer le visage en plusieurs images faites de nuances de gris, chacune mettant en évidence une caractéristique particulière :

Une autre technique appelée "feature analysis" se base sur la précédente en y rajoutant des informations sur les distances inter-éléments, leurs positions, etc... Elle se dit plus souple quant aux éventuelles modifications pouvant survenir : angle de prise de vue, inclinaison de la tête, etc..., ils viennent ensuite des techniques moins

utilisées à l'heure actuelle, basée sur des réseaux neuronaux, sur des méthodes plus techniques et moins souples. [1]

f). La Voix : la voix d'une personne se caractérise par beaucoup de paramètres. Chaque personne possède une voix propre que l'on peut analyser par enregistrement avec un micro. Les sons se caractérisent par une fréquence, par une intensité et par une tonalité. Le traitement informatique tient compte des distorsions liées au matériel utilisé, et sait analyser un son de mauvaise qualité tel qu'une transmission téléphonique ou radiophonique. La fatigue, le stress ou un rhume peuvent provoquer des variations de cette voix. La fraude est possible en enregistrent à son insu, la voix d'une personne autorisée, à moins d'obliger la personne contrôlée à lire un texte aléatoire.

La reconnaissance d'un locuteur offre l'avantage d'être bien acceptée par l'utilisateur, quelle que soit sa culture. De plus, s'il s'agit de sécuriser une transaction téléphonique, la voix est la seule information disponible. On distingue les systèmes à texte prédéterminé (texte dépendent) où l'utilisateur doit répéter un texte qu'il ne choisit pas, et les systèmes où la personne peut parler librement (text independent). La phase d'apprentissage utilise généralement plusieurs modèles du locuteur pour tenir compte de la variabilité de son discours. La phase de reconnaissance consiste à segmenter le signal de parole en unités qui sont ensuite classées. Ces unités peuvent être des mots ou des phonèmes. La performance est sujette à la qualité du signal, qui dépend de la variabilité de la voix du locuteur dans le temps comme dans le cas de maladie (e.g. rhume), des états émotionnels (e.g. angoisse, joie) et de l'âge, des conditions d'acquisition de la voix telles que le bruit et la réverbération de la qualité des équipements tels que le microphone, sans oublier le fait que différentes personnes peuvent avoir des voix similaires.[1]

I.6. Comparaison et évaluation dans une solution biométrique

Un certain nombre d'éléments d'identification biométriques sont utilisés dans diverses applications.

Chaque biométrique a ses forces et ses faiblesses et le choix dépend généralement de l'application. Aucun biométriques unique ne devrait répondre efficacement aux exigences

de toutes les applications. Le matching entre les données biométriques et une demande est déterminée en fonction de les caractéristiques de la demande et les propriétés des données biométriques. [Dav-03]

I.6.1. Critères de comparaison

Toute caractéristique physiologique et / ou de comportement peut être utilisé comme un identificateur biométrique de reconnaître une personne tant qu'il satisfait à ces exigences:

- ***Universalité* :** ce qui signifie que chaque personne devrait avoir les données biométriques.
- ***Unicité :*** ce qui indique que les deux personnes doivent être suffisamment différentes en terme de leurs éléments d'identification biométriques.
- ***Mesurable :*** ce qui indique que les données biométriques peuvent être mesurées quantitativement.

Toutefois, dans un système pratique biométriques. Il y a un certain nombre d'autres questions qui devraient être envisagées, y compris :

- ***Performance :*** qui se réfère à la précision obtenue de la reconnaissance, la vitesse, la robustesse, les ressources nécessaires pour atteindre la précision de la reconnaissance souhaitée et la vitesse, ainsi que des facteurs opérationnels et environnementaux qui influent sur la précision de la reconnaissance et vitesse.
- ***Acceptabilité :*** qui indique la mesure dans laquelle les gens sont prêts à accepter une par- notamment d'identification biométrique dans leur vie quotidienne.
- ***Mise en échec :*** qui reflète la façon dont il est facile de tromper le système par des méthodes frauduleuses. [Dav-03]

La figure I.5 représente une comparaison entre des principales technologies biométriques :

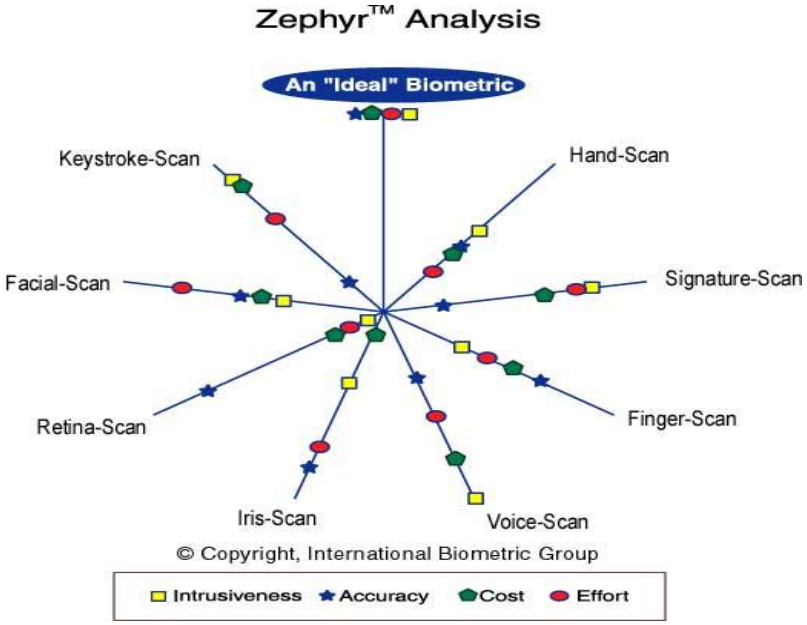


Figure I.5. Comparaison des principales techniques biométriques. [1]

Légende :

- ***Effort :*** Effort requis pour l'utilisateur lors de la mesure.
- ***Intrusiveness:*** Décrit dans quelle mesure l'utilisateur perçoit le test comme intrusif.
- ***Cost :*** Coût de la technologie (lecteurs, capteurs, etc...)
- ***Accuracy :*** Efficacité de la méthode (capacité à identifier quelqu'un).

I.6.2. Les erreurs de système biométrique

Un système biométrique typique de vérification engage deux types d'erreurs : confondant des mesures biométriques de deux doigts différents pour être du même doigt (false match) et confondre de deux mesures biométriques du même doigt d'être de deux doigts différents (false non-match). Noter que ces deux types d'erreurs sont également souvent dénotés en tant que l'acceptation fausse (false acceptation rate (FAR)) et rejet faux (false rejection rate (FRR))

a). Les erreurs de système de vérification

De la perspective de conception, le problème de vérification biométrique peut être formulé comme suit. Laisser le calibre biométrique stocké d'une personne être représenté comme T et l'entrée acquise pour l'identification soit représentée par *I*. Puis les hypothèses nulles et alternatives sont :

- ***H*0** : *1* ≠ T, entrée ne vient pas de la même personne que le calibre ;
- ***H*1** : I = T, entrée vient de la même personne que le calibre.

Les décisions associées sont comme suit.

- ***D*0**: la personne n'est pas qui elle prétend être ;
- ***D*1** : la personne de est qui elle prétend être.

La vérification implique d'assortir T et *I* en utilisent une mesure de similitude s (T, *I*). Si les points assortis sont moins que le seuil (t) de système, alors décider *D*0, sinon décider *D*1 La terminologie ci-dessus est empruntée à la théorie de communication, où le but est de détecter un message dans la présence de bruit. Ho est l'hypothèse que le signal reçu est seul bruit, et *H*1 est l'hypothèse que le signal reçu est message plus le bruit. Une telle formulation d'essai d'hypothèse contient en soi deux types d'erreurs :

- **Type *I*** : False accepte (*D*1 est décidé quand Ho est vrai) ;
- **Type *II*** : false rejet (*D*0 est décidé quand le *H*1 est vrai).

Le taux de fausse acceptation (false match rate ou FMR) est la probabilité de type (*I)* d'erreur et le taux de faux rejet (false non match rate ou FNMR) et la probabilité de type(*II*) d'erreur

- FMR= P(*D*1\ Ho =true);
- FNMR = P(*D*0\ *H*1 = true).

Note que (1-FNRM) est la puissance de l'essai de l'hypothèse.

Pour évaluer l'exactitude d'un système biométrique doit rassembler des points produits d'un certain nombre de paires d'empreinte digitale du même doigt (la distribution p (s \ *H*1 = vrai) de tels points s'appelle traditionnellement la distribution de légitimes), et des points produits d'un certain nombre de paires d'empreinte digitale de différents doigts (la distribution p (s \ H0 = vrai) de tels points s'appelle traditionnellement la distribution d'imposteur). [Dav-03]

$$FNMR = \int_{0}^{t} p(s \setminus H1 = vrai)ds$$

$$FMR = \int_{t}^{1} p(s \setminus H0 = vrai)ds$$

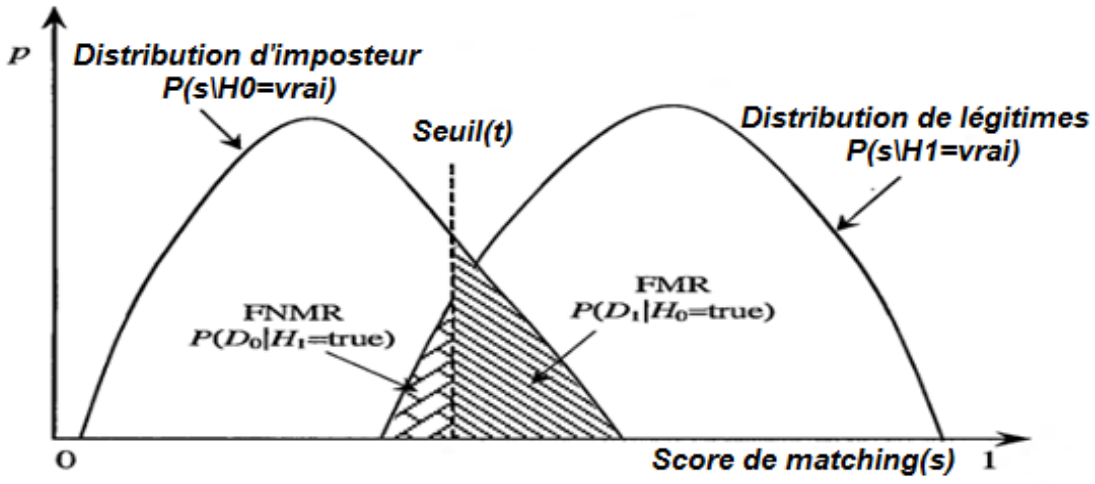


Figure I.6. Lien entre FMR et FNMR. [Dav-03]

D'autre index « compacts » sont également utilisés pour additionner récapitulent l'exactitude d'un système de vérification.

- **Le taux d'Égal-Erreur (Equal Error Rate EER) :** Dénote le taux d'erreur au seuil (t) pour lequel le taux fausse d'acceptation(false match rate) et le taux faux rejet(false non match rate) sont identiques : FMR (t) = FNMR (t). Dans la pratique, parce que les distributions assorties de points ne sont pas continués (en raison du nombre fini de paires assorties et de la quantification des points de rendement), un point exact d'EER ne pourrait pas exister.
- **ZeroFNMR :** est le plus bas FMR auquel FNRM=0
- **Manque de capturer (Failure To Capture FTC) :** le taux est associé à la fonction automatique de capture d'un dispositif biométrique et dénote le pourcentage des périodes des échouer de dispositif de capturer automatiquement le biométrique quand il est présenté à une sonde. Un manque élevé de capturer le taux rend le dispositif biométrique difficile à employer.
- **Le manque de s'inscrire (Failure To Enroll FTE) :** le taux dénote le pourcentage des périodes que les utilisateurs ne peuvent pas s'inscrire dans le système d'identification.

- **Le manque d'assortir (Failure To Match FTM) :** le taux est le pourcentage des périodes que l'entrée ne peut pas être traitée ou assortie contre un calibre valide en raison de la qualité insuffisante.

La courbe dite **ROC (Receiver Operating Characteristic)**, permet de représenter graphiquement la performance d'un système de vérification pour des différentes valeurs de *t***.**

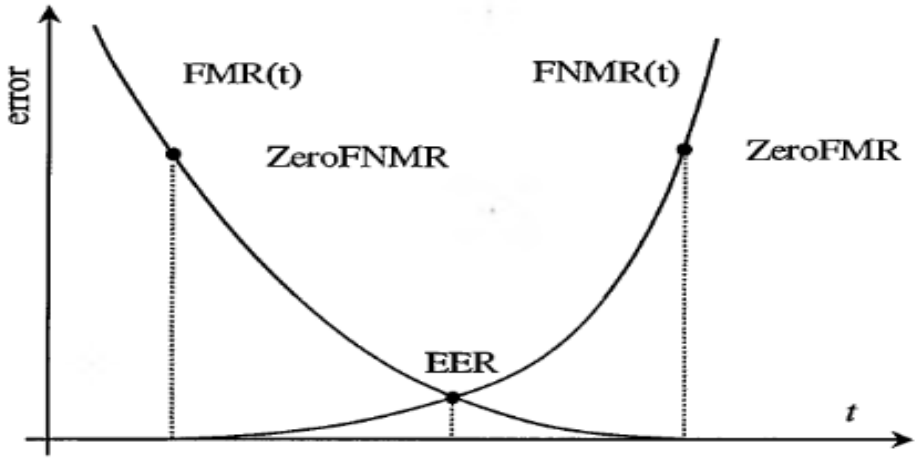


Figure I.7. Un exemple des courbes de FMR (t) et de FNMR (t), où les points correspondant à EER, à ZeroFNMR, et à ZeroFMR sont accentués.

b). Les erreurs d'un système d'identification

Comment les définitions des erreurs présentées ci-dessus pour la vérification d'empreinte digitale se prolongent à l'identification d'empreinte digitale de doigt ?

Dans quelques prétentions de simplification, une évaluation de l'exécution en mode d'identification peut être impliquée par les évaluations d'erreur en mode de vérification.

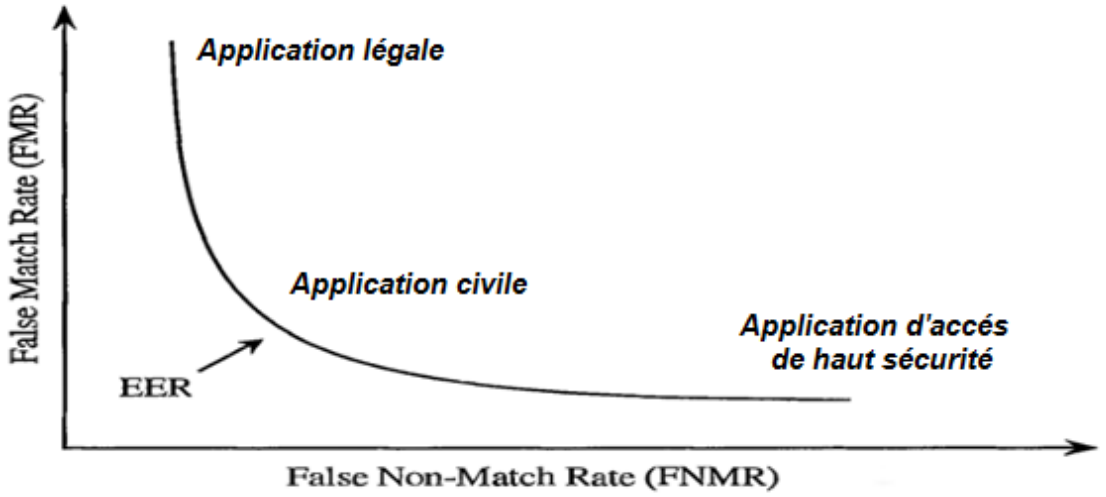


Figure I.8. Les points typiques de différentes applications montrées sur un ROC courbent.

Supposons qu'aucun mécanisme d'indexation/récupération n'est disponible (c.-à-d., la base de données entière contenir N calibres doit être recherché), et qu'un calibre simple pour chaque utilisateur est présent dans la base de données. Laisser $FNMR_N$ et FMR_N dénoter le taux de fausse acceptation d'identification et taux de faux rejet, respectivement, alors :

- $FNMR_N$ = FNMR; la probabilité de faux rejet de l'entrée contre le calibre d'utilisateur est le même qu'en mode de vérification.
- $FMR_N = 1 – (1 - FMR)^N$; une fausse d'acceptation se produit quand l'entrée s'assortit faussement un ou plusieurs calibres dans la base de données. FMR_N est alors calculé en tant qu'un moins la probabilité qu'aucune fausse d'acceptation n'est faite avec les calibres de base de données. Dans l'expression au-dessus (1 - FMR) est la probabilité que l'entrée n'assortit pas faussement à un seul calibre, et $(1- FMR)^N$ est la probabilité qu'il n'en assortit pas faussement des calibres de base de données. Si FMR est très petit, alors l'expression ci-dessus peut être rapproché par (FMR_N = N* FMR), et donc nous pouvons déclarer que la probabilité de fausse d'acceptation augmente linéairement avec la taille de la base de données. [Dav-03]

Si les calibres dans la base de données ont été classifiés/indexés alors seulement une partie de la base de données est recherchés pendant l'identification et ceci a comme conséquence une formulation différente de $FNMR_N$ et de FMR_N :

- $FNMR_N$ = RER + (1-RER)* FNMR, où RER (Retrieval Error Rate) le taux d'erreur de récupération est la probabilité qu' un calibre de base de données correspondant au doigt recherché est incorrectement lancé par le mécanisme de récupération.
- $FMR_N = 1 – (1 - FMR)^{N-P}$, où P (également appelé le taux de pénétration) est le pourcentage moyen de la base de données recherchée pendant l'identification d'une empreinte digitale d'entrée. [Dav-03]

I.6.3. Évaluation des performances

Phillips et autres spécialistes (2000) définissent trois types d'évaluation des systèmes biométriques : évaluation de technologie, évaluation de scénario, et évaluation opérationnelle.

- **Évaluation de technologie :** le but d'une évaluation de technologie est de comparer des algorithmes de concurrence d'une technologie simple. Seulement des algorithmes conformes avec un protocole donné d'entrée-sortie sont examinés (des dispositifs de détection et les aspects d'application ne sont pas tenus compte). L'essai de tous les algorithmes est effectué sur une ou plusieurs bases de données. Bien que des données d'échantillon puissent être distribuées pour le développement ou la tenir du but priori à l'essai, l'essai réel doit être fait sur les données qui n'ont pas été précédemment vues par les développeurs d'algorithme, puisque la base de données est fixe, les résultats des essais de technologie sont qu'on peut répéter. FVC2000 (Maio et autres, 2002a) et FVC2002 (Maio et autres, 2002b) sont des exemples des évaluations de technologie des algorithmes de vérification d'empreinte digitale discute en plus détail la comparaison des algorithmes de matching d'empreinte digitale. [Dav-03]
- **Évaluation de scénario :** le but de l'évaluation de scénario est de déterminer la performance de système global dans un prototype ou une application simulée.

L'essai est réalisé sur un système complet dans un environnement qui modèle une application réelle de cible. Chaque système examiné a son propre dispositif d'acquisition. La collecte de données à travers tous les systèmes examinés doit être effectuée dans le même environnement avec la même population. Les résultats d'essai sont qu'on peut répéter seulement dans la mesure où le scénario modelé peut être soigneusement contrôlé (UKBWG, 2002). [Dav-03]

- **Évaluation opérationnelle :** Le but de l'essai opérationnel est de déterminer l'exécution d'un système biométrique complet dans un environnement d'application spécifique avec une population de cible spécifique. Généralement les résultats d'essai opérationnels ne sont pas qu'on peut répéter en raison des différences inconnues et non documentées entre les environnements opérationnels (UKBWG, 2002). [Dav-03]

I.7. Conclusion

Dans ce chapitre nous avons vu l'émergence des méthodes biométriques de reconnaissance et par la même nous avons aperçu les quelques problèmes qui en découlent par voie de comparaison nous concluons que les empreintes digitales est une modalité qui offre un compromis intéressant entre précision, cout et acceptation, chose qui explique sa pré pondération sur le marché.

Dans la suite de ce travail, nous allons axer sur le problème de vérifications des empreintes digitales.

Chapitre

II

Système automatique d'identification d'empreinte digitale

II.1. Historique

Le XXème rapport de la CNIL (2000) nous rappelons que les chinois ont été les premiers à utiliser les empreintes digitales pour authentifier les contrats, sous la dynastie Tang (618-906), la logique gestionnaire, ou marchande, est donc présente dès les prémisses de la biométrie. Ce procédé a été repris par Francis Gallon, statisticien, cousin de Darwin. Scotlland Yard adopte ce procédé en 1901 pour identifier les criminels. Le procédé s'est rapidement diffusé dans les laboratoires de police scientifique, en France en 1903 avec Alphonse Bertillon, qui avait déjà fait adopter l'anthropométrie par la police dès 1882. Rappelons d'ailleurs que Alphonse Bertillon est lui-même le fils de Louis Adolpe Bertillon, qui a créé la société d'anthropologie avec Broca. [Syl-04]

II.2. Caractéristiques des empreintes digitales

Une empreinte digitale est constituée d'un ensemble de lignes localement parallèles formant un motif unique pour chaque individu (figure II.1), on distingue les crêtes (ou stries, ce sont les lignes en contact avec une surface au touché) et les vallées (ce sont les creux entre deux stries) les crêtes contiennent en leur centre un ensemble de pores régulièrement espacés.

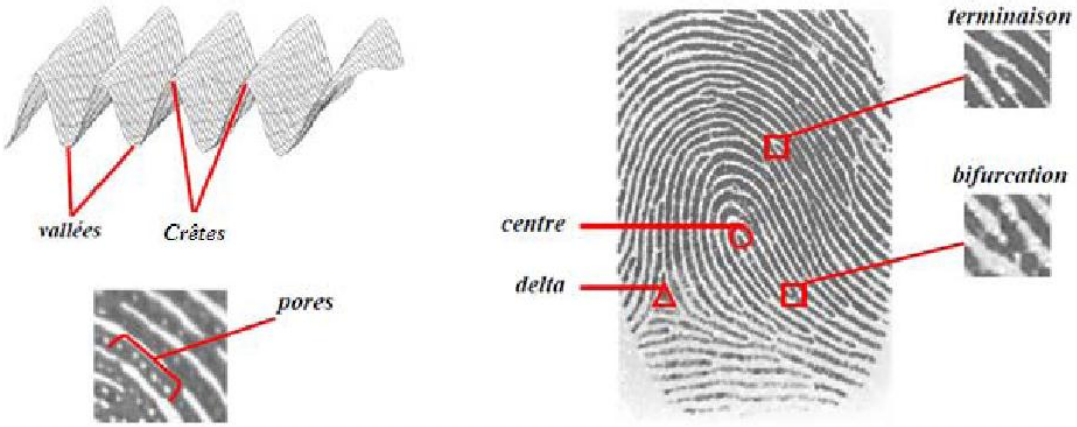
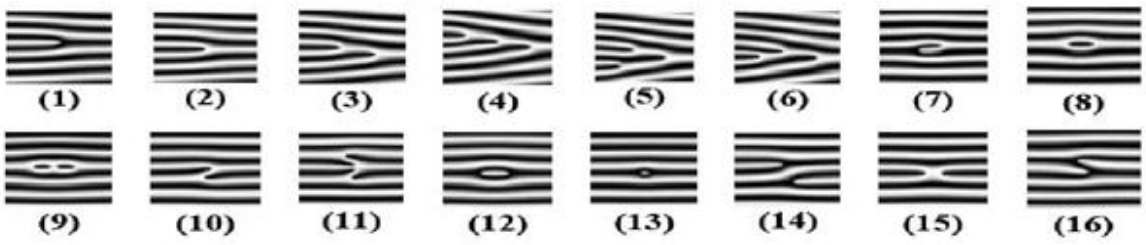


Figure II.1. Caractéristiques Des Empreints Digitales.

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

Une étude a montré 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. (Figure II.2)



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 II.2. Les différents types de minutie (crêtes en noir).

La position et le nombre de cores et de deltas permettent de classifier les empreintes en catégorie selon leur motif général. On distingue principalement trois grandes familles (Figure II.3). Ces trois types d'empreintes regroupent 95% des doigts humains, il existe aussi des dessins beaucoup plus rares comme les doubles boucles. [Vér-06]

- Les boucles (loop) représentent 60% des empreintes rencontrées.
- Les spirales (whorl) représentent 30% des empreintes rencontrées.
- Les arcs (arch) représentent 5% des empreintes rencontrées.



Figure II.3. Les classes d'empreinte, boucle(a), spirale(b), arc(c), double boucle(d).

L'ensemble Formé par la disposition des points singuliers constitue un motif unique pour chaque individu. En effet il a été montré que l'empreinte digitale se forme au cours du troisième mois de la vie fœtale, le motif général est influencé par les gènes héréditaires mais l'apparition des détails (minuties) est créée de manières accidentelles par des pressions variables aléatoires sur les surfaces tactiles. Ainsi l'empreinte est unique pour tout individu y compris pour des vrais jumeaux et a été montré que les méthodes de reconnaissance actuelles permettent d'identifier efficacement les jumeaux. 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. [Vér-06]

II.3. Architecture d'un système d'empreinte digitale

Les systèmes d'empreinte digitales peuvent fournir trois modes de fonctionnement, à savoir, **l'enrôlement**, **l'authentification** (**la vérification** ou **l'identification**). Dans ce qui suit, les figures illustreront l'exemple d'un système biométrique utilisant l'empreinte digitale comme modalité.

a). L'enrôlement : est la première phase de tout système d'empreinte digitales. Il s'agit de l'étape pendant laquelle un utilisateur est enregistré dans le système pour la première fois et où une ou plusieurs modalités biométriques sont capturées et enregistrées dans une base de données. Cet enregistrement peut s'accompagner par l'ajout d'information biographique dans la base de données. (Figure.II.4)

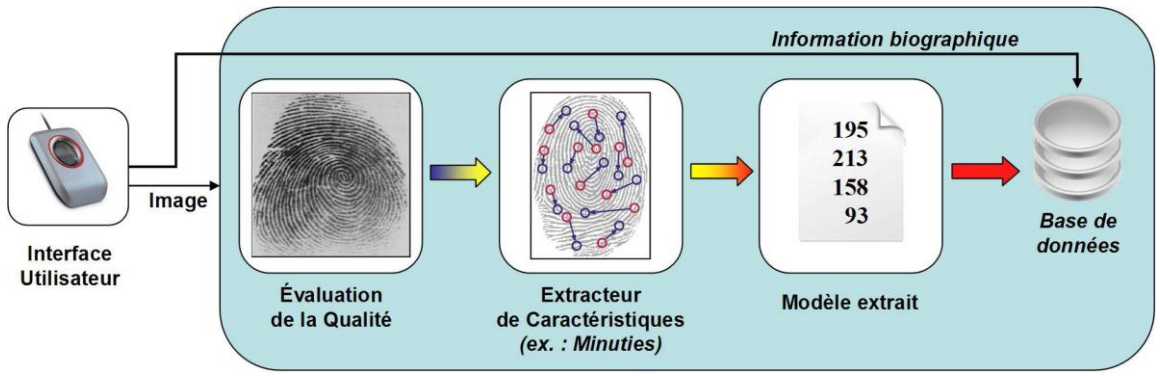


Figure II.4. La phase d'enrôlement.

b). La vérification: Lorsqu'un système d'empreinte digitale opère en **mode vérification** (Figure.II.5), l'utilisateur affirme son identité et le système vérifie si cette affirmation est valide ou non.

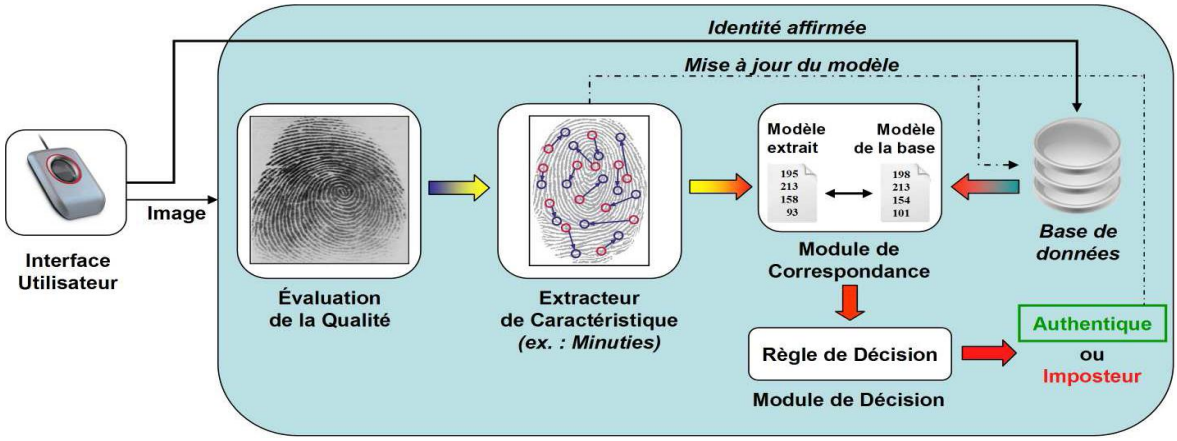


Figure II.5. Le mode de vérification.

Pour illustrer ce principe, prenons la situation où un utilisateur (*M. X*) souhaite retirer de l'argent à un distributeur de billets en entrant son code personnel d'identification (code PIN) et en présentant une modalité biométrique. Le système acquiert alors les données biométriques et va les comparer **uniquement** avec le modèle enregistré correspondant à *M. X*. On parle alors de **correspondance 1**:**1**. Ainsi, si l'entrée biométrique de l'utilisateur et le modèle enregistré dans la base de données correspondant à l'identité affirmée possèdent un degré de similitude élevé, l'affirmation est validée et l'utilisateur est considéré comme étant un **authentique**. Dans le cas contraire, l'affirmation est rejetée et l'utilisateur est considéré comme étant un **imposteur**. En résumé, un système biométrique opérant en mode vérification répond à la question "*Suis-je bien M. X*?".

c). L'identification : dans un système d'empreinte digitale opérant en **mode identification** (Figure. II.6), l'utilisateur ne dévoile pas explicitement son identité. Cependant, l'affirmation implicite faite par l'utilisateur est qu'elle est une des personnes déjà enrôlées par le système. Ainsi, l'échantillon biométrique de l'individu est comparé avec les modèles de **toutes les personnes** de la base de données. On parle alors de **correspondance 1**:**N**. La sortie du système biométrique est constituée par l'identité de la personne dont le modèle possède le degré de similitude le plus

élevé avec l'échantillon biométrique présenté en entrée. Typiquement, si la plus grande similarité entre l'échantillon et tous les modèles est inférieure à un seuil de sécurité minimum fixé, la personne est **rejetée**, ce qui implique que l'utilisateur n'était pas une des personnes enrôlées par le système. Dans le cas contraire, la personne est **acceptée**.

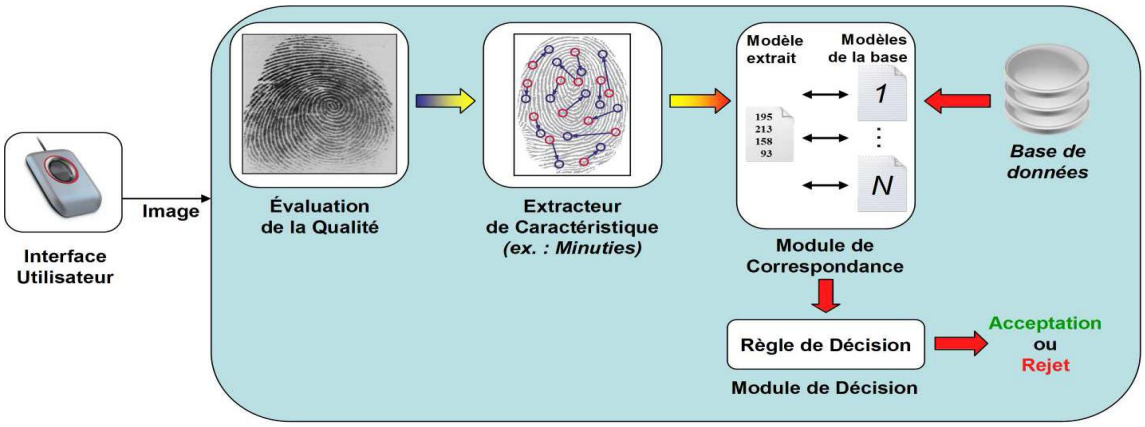


Figure II.6. Le mode d'identification.

Un exemple de système opérant en mode identification serait l'accès à un bâtiment sécurisé: tous les utilisateurs qui sont autorisés à entrer dans le bâtiment sont enrôlés par le système ; lorsqu'un individu essaye de pénétrer dans le bâtiment, il doit d'abord présenter ses données biométriques au système et, selon la détermination de l'identité de l'utilisateur, le système lui accorde le droit d'entrée ou non. En résumé, un système biométrique opérant en mode identification répond à la question "*Suis-je bien connu du système ?*". [Nic-09]

II.4. Matching d'empreinte digitale

II.4.1. Image

Dans cette représentation, c'est toute l'image qui est considérée comme une représentation possible. Le matching est réalisé par corrélation. La corrélation entre les deux images $I_1(x, y)$, $I_2(x, y)$ est donnée dans le domaine spatial par :

$$I_c(k,l) = \sum_x \sum_y I_1(x+k, y+1) I_1(x,y)$$

Le corrélateur établit la correspondance par la recherche de la magnitude du pic dans l'image de corrélation I_c. La position du pic indique la translation entre les images et la valeur du pic informe sur le degré de similarité. L'exactitude de cette corrélation se dégrade avec les transformations de l'image comme les phénomènes de translation et de rotation. Le problème de distorsion peut être surmonte par les méthodes de corrélation locales proposées dans [Nan-04] et dans [Var-00]. Un autre inconvénient est en relation avec la taille conséquente de l'image à sauvegarder durant l'inscription.

II.4.2. Minuties

Le but d'un algorithme de matching est de comparer deux images ou deux gabarits et de retourner le score de similarité qui correspond à la probabilité que deux empreintes se correspondent. A l'exception des algorithmes bases sur la corrélation, la plupart des algorithmes extraient des caractéristiques dans le but de faire le matching. Les détails de minutie constituent la représentation la plus populaire de toutes les représentations existantes. Elles répondent efficacement au problème de taille pose précédemment.

II.4.2.1. Représentation des minuties

Les minuties représentent des discontinuités locales et marquent les positions ou la crête se termine ou bifurque. Cela constitue les types de minutie les plus fréquentes [Bel-06]

Chaque minutie peut être décrit par un nombre d'attributs tels que la position (x, y), l'orientation q et d'autres informations susceptibles d'aider à le matching. Cependant, la plupart des algorithmes considèrent seulement sa position et son orientation :(Figure.II.7)

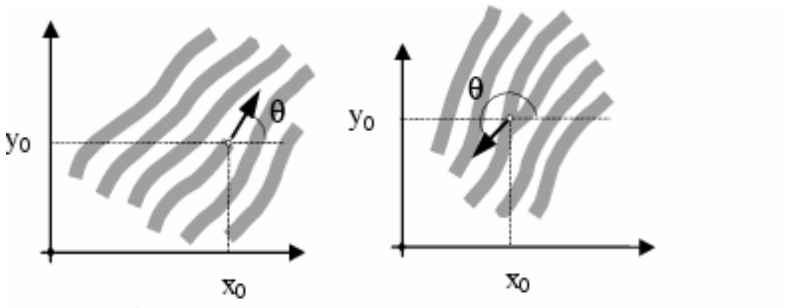


Figure II.7. Les caractéristiques principales des minuties.

Les minuties peuvent être appariées en considérant le problème comme un problème de matching de primitives point (point pattern matching). La (Figure.II.8) est une représentation en minuties d'une empreinte :

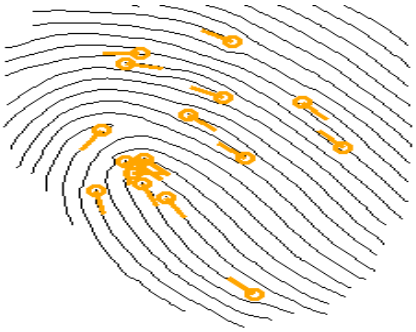


Figure II.8. Représentation d'une empreinte par sa carte de minuties.

II.4.2.2. Extraction de modèle biométrique

Il y a plusieurs techniques dans la littérature pour l'automatisation de l'extraction des caractéristiques des empreintes digitales. Nous distinguons deux catégories :

- La méthode classique qui consiste à extraire l'information à partir d'un squelette binaire (noir et blanc) de l'image filtrée.
- La méthode directe qui consiste à extraire les minuties directement à l'image en niveau de gris.

a). La méthode classique

Cette méthode est la plus répandue, elle consiste à extraire les minuties à partir d'un squelette de l'image filtrée (adoucissement de l'image). Le processus d'extraction suit les phases suivantes : amélioration, binarisation, squelettisation et extraction de minuties.

I. **Amélioration :** une fois l'image chargé un prétraitement est effectué au niveau de l'image afin d'améliorer la qualité de crêtes, cette qualité peut être endommagé par l'état de l'épiderme (sécheresse, humidité, plis, etc.) Ou les bruit du capteur d'acquisition. L'algorithme qui nous avons utilisé est proposé par le groupe de *Lin Hong, Yifei Wan*, et *Anil Jain* à l'Université de l'Etat du Michigan [Lin-98] en appliquant le filtre de Gabor.

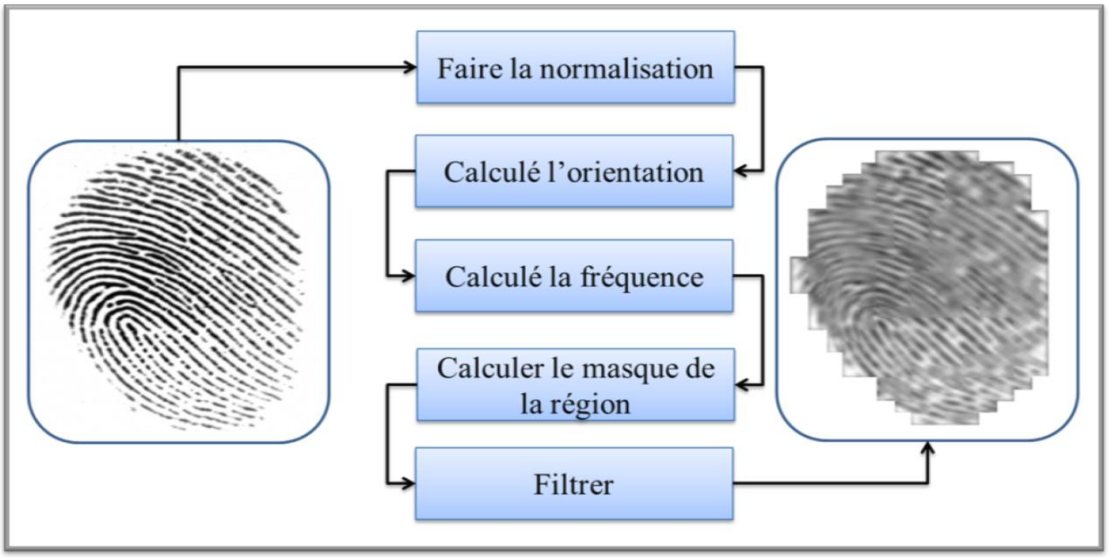


Figure II.9. Processus d'augmentation d'une image.

Soit l'image entrée *I*, est une image en niveau gris *N x M* pixels où *I(i ,j), i=1…N, j=1…M* représente la couleur en niveau de gris à ligne *i-ème* et colonne *j-ème*.

Le moyen de l'empreinte digitale est définit par :

$$M(I) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} I(i, j)$$

Et la variance de l'empreinte digitale est définit par :

$$VAR(I) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} (I(i, j) - M(I))^2$$

I.1. Normalisation

L'image originale est normalisée par son moyen *M(I)* et sa variance *VAR(I)*. La matrice *G(I)* indique l'image normalisée en niveau de gris et *G(i, j)* représente sa couleur à *(i , j)*.

$$G(i, j) = \begin{cases} M_0 + \sqrt{\dfrac{VAR_0 (I(i, j) - M)^2}{VAR}} & ,if \quad I(i, j) > M \\ M_0 - \sqrt{\dfrac{VAR_0 (I(i, j) - M)^2}{VAR}} & , \quad autrement \end{cases}$$

Où M_0 et VAR_0 sont la moyenne et la variance désirées. Cette étape ne change pas la clarté de la structure d'empreinte digitale. Le but de normalisation est la réduction de la variation de niveau gris entre les crêtes et les vallées. Le résultat fait faciliter les étapes suivantes.

I.2. Orientation de crête

L'image d'orientation de l'image *I* est définit comme une image *O* de *NxN*, où *O(i, j)* représente l'orientation de crête locale de pixel *(i, j)*. L'orientation de crête est couramment spécifiée par un bloc. Donc, l'empreinte digitale est divisée par un ensemble de bloc de *w x w* non-chevauchés. Une seule valeur est définit pour un bloc.

En regardant l'empreinte digitale et ses textures, il y a quelques approches pour calculer l'orientation. Dans le cadre de notre travail, nous utilisons la

méthode d'estimation de l'orientation de moindre carré.

Algorithme

- Étape 1 : Diviser *G* par bloc de *w x w* (16 x 16).
- Étape 2 : Calculer les gradient $\partial_x(i, j)$ et $\partial_y(i, j)$ pour chaque pixel(i, j).On utilise également la matrice de Roberts, Sobel ou bien Marr-Hildreth.
- Étape 3 : Estimer l'orientation locale de chaque bloc. Pour le point central (i, j) utilisant l'équation suivante :

$$V_x(i, j) = \sum_{u=i-\frac{w}{2}}^{i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{j+\frac{w}{2}} 2\partial_x(u, v)\partial_y(u, v)$$

$$V_y(i, j) = \sum_{u=i-\frac{w}{2}}^{i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{j+\frac{w}{2}} (\partial_x^2(u, v) - \partial_y^2(u, v))$$

$$O(i, j) = \frac{V_y(i, j)}{V_x(i, j)}$$

$$\theta(i, j) = tg^{-1}O(i, j)$$

I.3. Fréquence de crête

L'image de fréquence de l'image I est définie comme une image F de N x N, où F(i ,j) représente la fréquence de la crête locale de pixel (i ,j). Dans la partie où la densité de minutie est basse ou pas de minutie, l'empreinte digitale se compose des crêtes et des vallées comme sinusoïdal forme. C'est une des caractéristiques importantes d'empreinte digitale que la fréquence locale représente. Ci-dessous est les étapes de calcul de la fréquence d'une image :

Algorithme

- Étape 1 : Diviser *G* par bloc de *w x w* (16 x 16).
- Étape 2 : Pour chaque bloc centré à pixel *(i, j)*, calculer la fenêtre orientée de *l x w* qui est définit le système de coordonnées crête

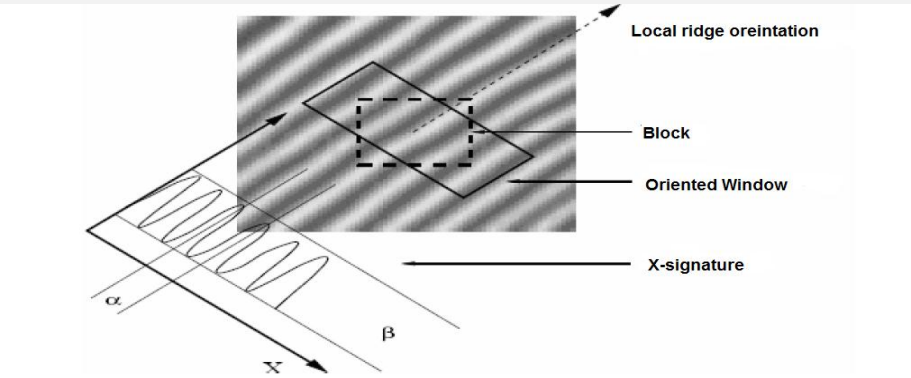


Figure II.10. Fenêtre d'orientation et x-signature.

- Étape 3 : Pour chaque bloc centré à pixel *(i, j)*, calculer le x- signature, *X[0], X[1], ...X[l-1]* de crêtes et de vallées basées sur la fenêtre orientée :

$$X[k] = \frac{1}{w}\sum_{d=0}^{w-1} G(u,v) \qquad ,k = 0....l-1$$

$$u = i + (d - \frac{w}{2})\cos(\theta(i,j)) + (k - \frac{1}{2})\sin(\theta(i,j))$$

$$v = j + (d - \frac{w}{2})\sin(\theta(i,j)) - (k - \frac{1}{2})\cos(\theta(i,j))$$

S'il n'y a pas de minutie dans la fenêtre, la x-signature est la même. On peut estimer la fréquence F(i, j) basée sur x-signature.

$$F(i,j) = \frac{1}{T(i,j)}$$

Où T(i, j) est le nombre moyen de pixel entre deux sommets consécutifs de

x-signature.

I.4. Le masque de région

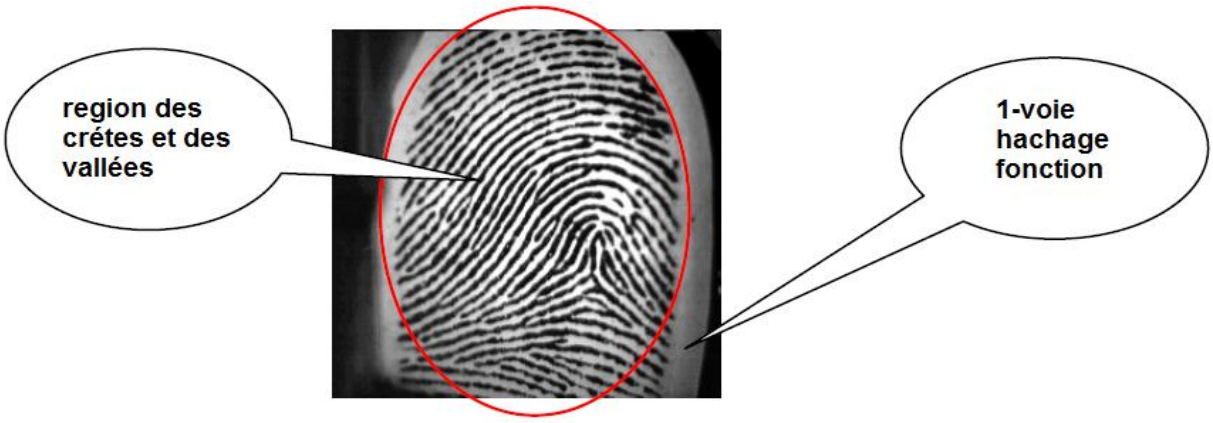


Figure II.11. Région des crêtes et pas de crête.

Le masque de région est définit comme une image $R_{N \times M}$ où R(i, j) indique la catégorie du pixel (i, j). Un pixel doit être ou bien le pixel pas-de-crête-et-vallée (irrécouvrable) ou bien le pixel crête-et-vallée (recouvrable). Le masque de région est pour but de détection de région soit recouvrable, soit irrécouvrable.

Le masque de région est basé sur l'amplitude (α), fréquence (β), et variance (γ). Soit *X[0], X[1], ...X[l-1]* la x-signature de bloc centré à *(i, j)*. Trois Caractéristiques sont calculées par :

α =moyenne de hauteur de la crête – moyenne de profondeur de vallée.

$$\beta = \frac{1}{T(i,\ j)}$$

$$\gamma = \frac{1}{l}\sum_{i=1}^{l}(X[i]-(\frac{1}{l}\sum_{i=1}^{l}X[i]))^2$$

Pour chaque fenêtre, on obtient un vecteur $(\alpha\ ,\ \beta\ ,\ \gamma)$. Les régions d'image sont classifiées par deux groupes : recouvrable et

irrécouvrable en utilisant l'algorithme k-plus-proche voisin. Deux classes initiales sont choisies :

- Classe irrécouvrable : 4 bloques/ coin x 4 coins = 16 bloques initiales.
- Classe recouvrable : 16 bloques au centre
- La distance euclidienne est utilisée pour estimer la distance.

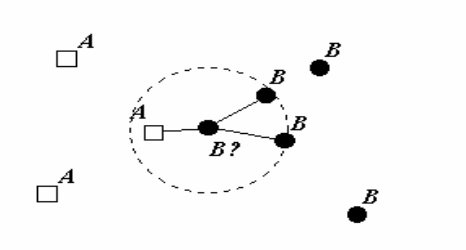


Figure II.12. k-plus-proches voisins.

I.5. Filtre de Gabor

Le filtre de Gabor est directement lié aux ondelettes de Gabor, puisqu'ils peuvent être conçus pour le nombre de *érosions et de rotations*. Cependant, généralement l'expansion n'est pas appliquée pour des ondelettes de Gabor, puisque ceci exige le calcul des ondelettes de biorthogonal, qui peuvent prendre très du temps. Par conséquent, habituellement, une batterie de filtres se composant de Gabor filtre avec de diverses balances et des rotations est créée. Les filtres sont enlacés avec le signal, ayant pour résultat un prétendu espace de Gabor. Ce processus est étroitement lié aux processus dans le cortex visuel primaire. Les relations entre les activations pour un endroit spatial spécifiques sont très distinctives entre les objets dans une image. En outre, des activations importantes peuvent être extraites à partir de l'espace de Gabor afin de créer une représentation clairsemée d'objet.

$$h(x, y : \phi, f) = \exp\left\{-\frac{1}{2}\left[\frac{(x\cos\phi)^2}{\delta_x^2} + \frac{(y\sin\phi)^2}{\delta_y^2}\right]\right\}\cos(2\pi fx\cos\phi),$$

Ou ϕ est l'orientation et f est la frequence. σ_x et σ_y sont les constance d'enveloppe Gaussian pour l'axis x et y respectivement. Dans cette application. Appliquant ce filtre pour une image, trois paramètres doivent être spécifiés : la fréquence de crêtes et de vallées, l'orientation de crêtes, la normalisation d'image :

$$\varepsilon(i,j) = \begin{cases} 255 \quad , \quad si\ R(i,j) = 0 \\ \sum_{u=-\frac{w}{2}}^{\frac{w}{2}} \sum_{v=-\frac{w}{2}}^{\frac{w}{2}} h(u,v,O(i,j),F(i,j))G(i-u,j-v) \quad , \quad auterment \end{cases}$$

Où w=11 est la taille de filtre Gabor.

Le résultat de cette étape est une image en niveau de gris. Elle est convertie en image binaire par un seuil moyen.

II. Binarisation *:* Les approches de binarisation classiques reposent sur le choix d'un seuil global T.

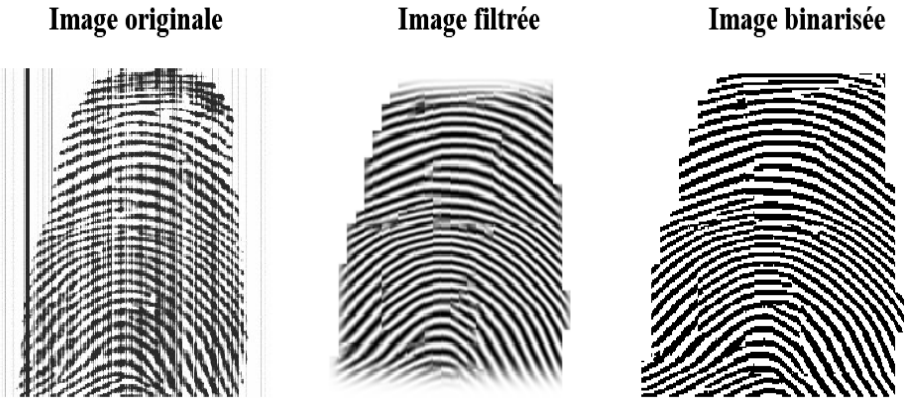


Figure II.13. L'image après l'étape de binarisation.

Le seuillage est une opération de base dans le traitement d'images. Et, un bon algorithme commence toujours par une bonne base ! Dans notre travail nous avons utilisé Le seuillage d'Otsu.

Dans le traitement d'images, la méthode du seuillage d'Otsu (1979) est utilisée pour la décision automatique de niveau de binarisation, basée sur la forme de l'histogramme. L'algorithme suppose que l'image se compose de deux classes de base : Premier plan et fond. Il calcule alors un seuil optimal qui minimise la pondération de variance inter-class de ces deux classes. On le montre mathématiquement que la minimisation inter-class correspond la maximisation inter-class. [Ots-79]

Formulation mathématique :

Let q1 et q2 représentent l'évaluation des probabilités de classe définies comme :

$$q_1(t) = \sum_{i=1}^{t} P(i) \qquad et \qquad q_2(t) = \sum_{i=t+1}^{I} P(i)$$

Et les sigmas sont les différents inter-class définis comme :

$$\sigma_1^2(t) = \sum_{i=1}^{t} [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \;\; et \;\; \sigma_2^2(t) = \sum_{i=t+1}^{t} [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)}$$

Et les moyens de classe :

$$\mu_1(t) = \sum_{i=1}^{t} \frac{iP(i)}{q_1(t)} \qquad et \qquad \mu_2(t) = \sum_{i=t+1}^{I} \frac{iP(i)}{q_2(t)}$$

Ici, P représente l'histogramme d'image. Le problème de la minimisation inter- class peut être exprimé comme problème de maximisation inter-class. Il peut écrire comme la différence de la variance totale et la variance inter-class :

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = q_1(t)[1 - q_1(t)][\mu_1(t) - \mu_2(t)]$$

Finalement, cette expression peut sans risque être maximisée et la solution est *t* qui est :

$$\sigma_b^2(t)$$

Pseudo-code pour la Binarisation

Pour chaque seuil potentiel T:

- Séparer les Pixel dans deux groupes selon le seuil.
- Trouver la moyenne de chaque groupe.
- Ajuster la différence entre les moyennes.
- Multiplier le nombre de Pixel de des deux groupe.

Si nous incorporons un peu de maths à cet algorithme, une telle explication évolue :

- Calculer l'histogramme et les probabilités de chaque intensité.
- Initialiser $q_i(0)$ initial et $\mu_i(0)$.
- Étape par toute l'intensité possible de maximum de seuils.
- Mise à jour q_i et μ_i.
- Calcul $\sigma_b^2(t)$

  Le seuil désiré correspond au maximum.

III. Squelettisation : la **squelettisation** est une classe d'algorithmes utilisée en analyse de formes. Elle consiste à réduire une forme en un ensemble de courbes, appelées **squelettes**. La squelettisation est un outil d'analyse de forme non-scalaire, qui conserve les propriétés topologiques de la forme d'origine ainsi que les propriétés géométriques, selon la méthode employée.

Un grand nombre d'approches de squelettisation existent dans la littérature, du certainement au rôle central que joue cette étape en traitement d'image dans diverses applications : reconnaissance de caractères, analyse de document et autres. Hung [Hun-93] utilise un algorithme proposée par Arcelli et Baja dans [Arc-84], Ratha et autres [Rat-95] adoptent une

technique de squelettisation incluse dans la librairie HIPS. Finalement Coetze et Botha [Coe-93] utilisent la méthode de Baruch [Bar-88].

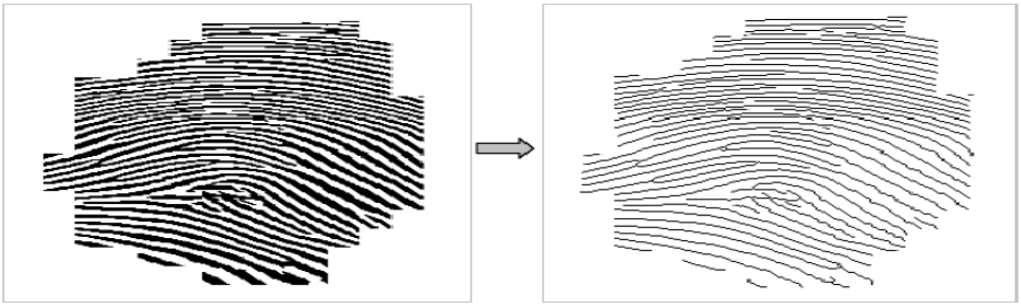


Figure II.14. Squelettisation de l'image d'empreinte.

Dans notre travail nous avons utilisé l'algorithme de **Hilditch** [2]**,** sont pseudo-code est donne comme suite :

Pseudo-algorithme pour la squelettisation

DEBUT
I : image binaire en entrée ;
Change : booléen ;
Change = vrai ;
Tant que (change = vrai) Faire
DTQ
Change = faux ;
Pour i allant de 1 a longueur de l'image I
Pour j allant de 1 a largeur de I
DPR
SI (Image [i ,j] = 1) Alors
DSI
Parcourir le voisinage 3x3 du pixel (i ,j)
SI une des configurations C se présente Alors
DSI

Image [i ,j]= 0;

Change = vrai;

FSI

/* Les configurations structurelles C sont : *

0	0	0	0	0			1			1		1	1	1	1		0	0		1		0	0
	1		0	1	1	0	1	1	1	1	0		1		1	1	0	0	1	1	1	1	0
1	1	1		1	1	0	0			0	0	0	0	0	0	1		0	0		1	1	

FSI

FPR

FTQ

FIN.

IV. Extraction de minuties : Les deux étapes de préparation à l'extraction (binarisation et squelettisation) ont grandement facilite cette phase. En effet nous disposons un pixel blanc et un pixel noir pour chaque point du squelette, on obtient les nombre *CN*(CrossingNumber) de crêtes portant de ce point et nous pouvons donc déterminer si un pixel est une minutie ou non.

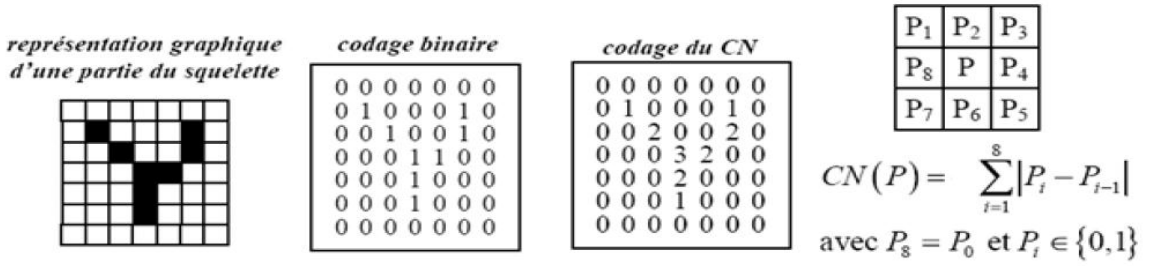


Figure II.15. Détection des minuties.

Ainsi pour chaque pixel *P* appartenant à une strie (c'est-à-dire pour chaque pixel ayant une valeur de 1) le calcul de *CN* peut prendre cinq valeurs :

CN(P)=0 : dans ce cas il s'agit d'un pixel isolé et nous n'en tenons pas compte car même si ce type de minutie il est très rare et à ce stade du traitement de l'image il est probablement dû à un résidu de bruit.

CN(P)=1 dans ce cas nous avons à faire à une minutie de type *terminaison*

CN(P)=2 c'est le cas le plus courant, le pixel se situe sur une série, il n'y a pas de minutie.

CN(P)=3 nous sommes en présence d'une *bifurcation* triple.

CN(P)=4 nous sommes en présence d'une *bifurcation* quadruple. Ce type de minutie étant assez rare il est probablement due à du bruit et nous l'ignorons.

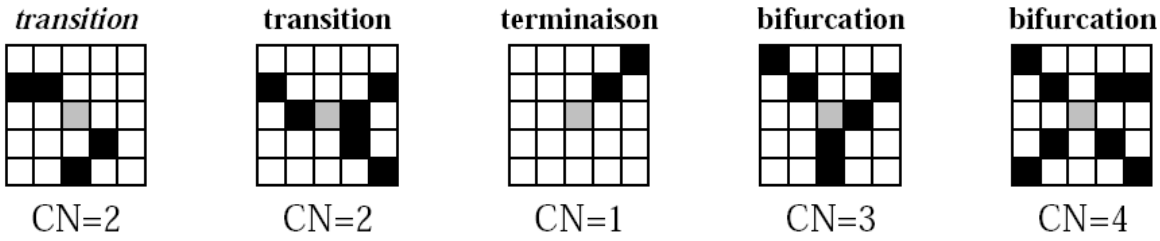


Figure II.16. Les types de connexion d'une minutie.
(Dans chaque cas on considère le pixel gris au centre de carré).

V. Extraction de point core

Le point core est le centre de la région d'intérêt. Localiser le point core est une étape essentielle qui peut influencer la précision de matching. Nous définissons le point core d'une empreinte digitale comme le point de courbure maximale des crêtes dans les images d'empreintes digitales et les axes de références sont définis comme des axes de symétrie au point core. Le point core et les axes de référence établissent un invariant pour une empreinte digitale. La détection de cette référence repose sur la méthode dite : Indice du Poincaré.

La localisation du point core passe par étapes suivantes :

Pseudo-algorithme pour la détection du point core

1. Estimation du champ d'orientation O.
2. Lissage du champ d'orientation dans le voisinage local. Le champ d'orientation lissé est représenté par O, tel que :

$$\phi_x = \cos(2O(i, j))$$

Et $$\phi_y = \sin(2O(i, j))$$

Avec ϕ_x et ϕ_y sont les composantes x et y du champ vectoriel, respectivement.

$$\phi_x'(i, j) = \sum_{u=-\frac{W_\phi}{2}}^{\frac{W_\phi}{2}} \sum_{v=-\frac{W_\phi}{2}}^{\frac{W_\phi}{2}} W(u,v)\phi_x(i-uw, j-vw)$$

Et $$\phi_y'(i, j) = \sum_{u=-\frac{W_\phi}{2}}^{\frac{W_\phi}{2}} \sum_{v=-\frac{W_\phi}{2}}^{\frac{W_\phi}{2}} W(u,v)\phi_x(i-uw, j-vw)$$

ou W est un filtre moyenneur à deux dimensions dont la taille est …..(la taille du filtre par défaut est 5x5). Le champ vectoriel lissé O à (i, j) est calculé comme suit :

$$O'(i, j) = \frac{1}{2}\tan^{-1}\left(\frac{\phi_y'(i, j)}{\phi_x'(i, j)}\right)$$

3. Calcul de ξ, la composante sinus O de l'image d'orientation.

$$\xi(i, j) = \sin(O'(i, j))$$

4. Soit A l'image étiquette initialisée à O, //utilisée pour indiquer le point core.
5. Parcours de A. Pour chaque pixels (i, j) dans … intégrer les intensités des pixels dans les régions Ri et Rii et assigner aux pixels correspondants dans A la valeur de leur différence :

$$A(i, j) = \sum_{R_I} \xi(i, j) - \sum_{R_{II}} \xi(i, j)$$

6. Détection de la valeur maximum dans A et assignation de ses coordonnées comme étant celles du point de core.

Si la région d'intérêt set dans les frontières de l'empreinte et tous les secteurs représentent une alternance et de vallées alors l'image est acceptée sinon elle est rejetée

VI. Post-traitement et filtrage de minuties : dans ce paragraphe, on considère les erreurs introduites durant les phases d'extraction précédentes. L'extraction peut engendrer ces deux types d'erreur :

- **Oubli de minuties** : l'algorithme d'extraction échoue quant à la détection de vraies minuties existantes sur l'empreinte. En général parce que cette minutie est dans une région bruitée ou parce que la structure des crêtes sont assez endommagées.
- **Ajout de minuties** : l'algorithme d'extraction identifie faussement une structure de crêtes altérée comme étant une minutie.

Par conséquent, l'étape d'extraction de minuties est souvent suivie par un post traitement qui essaye d'éliminer les fausses minuties. Ce raffinement peut considérablement améliorer l'exactitude des résultats. Les méthodes existantes qui filtrent les fausses positives peuvent être classifiées en deux catégories :

- **Post traitement structurel :** ce genre de post traitement identifie les fausses minuties en se basant sur des règles de type heuristique qui utilisent la location relative des minuties détectées, la longueur des crêtes et d'autres informations. Jain donne dans [Jan-97] une liste de ces quelques heuristiques. Cependant, ces règles sont très spécifiques à l'algorithme d'extraction. Xiao et Raafat [Xia-91] donnent une taxonomie des structures résultantes de l'étape de squelettisation et qui peuvent induisent les fausses minuties et proposent des heuristiques pour les éliminer. Hung [Hun-93] propose un algorithme base graphe qui exploite la dualité entre crêtes et vallées.
- **Filtrage de l'image en niveau de gris** : ces techniques utilisent les valeurs des niveaux de gris du voisinage de la minutie considérée. Dans [Par-00], une région 64 x 64 du voisinage entourant la minutie est normalisée. Des filtres de Gabor horizontaux sont utilisés afin d'améliorer la région et les pixels du

centre de chaque région 32 x 32 sont utilisés comme des attributs pour distinguer les voisinages d'une minutie et les régions qui n'avoisinent aucune minutie. La classification est faite en utilisant la méthode LVQ (Learning Vector Quantizer) qui utilise les 1024 pixels du centre comme entrée, et classifie les régions comme étant des crêtes, des bifurcations ou sans minutie. Maio et Maltons [Mai-98] proposent un réseau de neurones pour le filtrage des minuties. La dimension de l'ensemble des caractéristiques est réduite en utilisant la transformée de Karhunen Leove. La classification est opérée par un réseau de neurones de poids partages qui utilisent les images positive et négative. Malgré le fait que cette méthode aboutisse à une réduction significative dans le nombre des fausses détections, elle augmente le nombre des minuties oubliées. Récemment, Wu et all [Wu-03] développèrent un filtre optimal pour la détection des plis dans l'image d'empreinte. Dans cette approche, le voisinage contenant le pli est représenté en utilisant un rectangle paramètre. Cependant, cette méthode ne peut pas détecter les fausses minuties qui occurrent dans les régions de l'empreinte qui sont très sèches ou assez poussiéreuses.

Dans notre cas, nous avons utilisé la taxonomie suivante [Tic-00] :

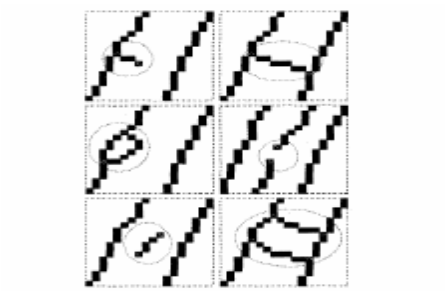


Figure II.17. Les types des fausses minuties.

Son pseudo-algorithme est donne comme suite :

Pseudo-algorithme pour l'élimination des fausses minuties

DEBUT
Pour chaque minutie candidate M
DPR
Déterminer un bloc B de taille W x W autour de la minutie ;
SI (M est une fin) Alors
DSI
Etiqueter par 1 tous les pixels qui sont connectes a M se trouvant dans B;
Comptabiliser le nombre de transition T01de 0 a 1 sur les bords du bloc B en faisant un parcours
suivant les aiguilles d'une montre;
SI (T01 = 0) Alors valider la minutie M;
FSI
SI (M est une bifurcation) Alors
DSI
Etiqueter par 1, 2, 3 les trois pixels voisins de M;
Etiqueter tous les pixels sur le bloc B connectes a ces trois voisins ;
Comptabiliser le nombre de transition T01 de 0 a 1, T02 de 0 a 2, T03 de 0 a 3;
SI (T01 = 1 et T02 = 1 T03 = 1) Alors valider la minutie M.
FSI
FPR ; FIN

La taille du bloc dépend de la distance inter-crête (W = 23, par exemple).

b). La méthode directe

Elle consiste à extraire la carte des minuties directement sur l'image brute en niveaux de gris ce qui permet d'éviter les étapes de binarisation et squelettisation. [Mai-97]

Dans cette méthode l'extraction des minuties se scinde en deux étapes :

a. Estimation de l'image d'orientation en examinant le gradient de l'intensité des pixels selon x et y.
b. Suivie des crêtes selon l'image d'orientation et détection, en parallèle, des minuties en appliquant un ensemble de règles sur les crêtes de l'empreinte digitale.

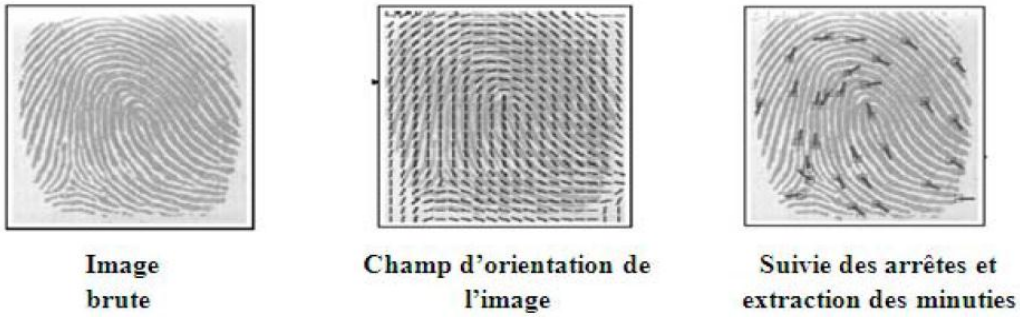


Figure II.18. Méthode d'extraction directe.

II.4.3. Description de texture

Les algorithmes de matching bases sur les minuties ne s'exécutent pas bien quand la surface de l'empreinte est petite étant donné que le nombre de minuties n'est pas suffisant pour le matching. De plus, le processus d'extraction des minuties est plus enclin à l'erreur dans ce genre de cas. Prabhakar et al [Pra-99] proposent une méthode basée sur un banc de filtres pour effectuer le matching de cette sorte d'images. Ils décrivent un nouveau descripteur de texture appelé "code d'empreinte".

Les informations de texture (moyennes et variances) sont extraites en effectuant une sectorisation [Ros-02] de l'image autour d'un point de référence (le point core); et le matching est base sur le calcul de la distance euclidienne entre deux codes. L'inconvénient de cette approche est que le point core doit être localise de façon exacte. Cela n'est pas possible pour des images de mauvaise qualité. De plus, la performance est inférieure au matching basé sur les minuties. Cependant, une décision qui combine les paramètres de texture et les minuties aurait un meilleur rendement.

II.5. Conclusion

Nous avons présenté dans ce chapitre l'empreinte digitale comme une modalité de système biométrique. Nous avons expliqué les différentes caractéristiques des empreintes digitales, ainsi nous avons abordé le fonctionnement algorithmique de chaque module du systeme.

Chapitre

III

Algorithmes de matching

III.1. Introduction

Un grand nombre d'algorithmes d'identification d'empreinte digitale est basé sur la corrélation d'image, la texture et les bancs de filtres (filterbanks), les algorithmes basés sur les minuties sont les plus utilisés couramment. C'est parce que les minuties dans les empreintes digitales sont la plupart des dispositifs distinctifs et fiables. La plupart des algorithmes basés sur les minuties ne se fondent pas sur le pré-alignement des empreintes digitales et suivent des processus de matching qui fonctionnent avec la complexité de temps de $O(n^2)$ (souvent $O(n^3)$), où n est le nombre de minuties. Cette complexité peut être réduite si nous devions pré-aligner les deux empreintes digitales d'une façon quelconque de sorte que le problème réduise à O(n). Les caractéristiques globaux des empreintes digitales telles que des points cores et de delta sont des points de borne limite dont l'endroit est conformé à travers différentes impressions du même utilisateur.

Dans ce chapitre nous avons étudié l'algorithme de matching proposé par **Mayur D Jain('Microsoft (R&D) India Private Ltd.)** et autres. Il consiste à pré-aligner l'empreinte dans une structure arborescente (arbre binaire) à partir de son point core tel que chaque minutie représente un nœud et contient deux fils(gauche et droit) qui sont les minuties les proches à la minutie père, le but de cet algorithme et de réduire le temps d'appariement de l'ordre $O(n^2)$ à l'ordre O(n).

III.2. Extraction de point core

Il y a plusieurs approches proposées pour la détection des points singuliers (core et delta). De loin, la méthode la plus populaire est celle proposée par Kawagoe et Tojo qui est basé sur le calcul de l'index de Poincaré. En outre, récemment une improvisation intéressante sur l'index de Poincaré proposé par Bazen et Gerez. Cependant, ceux-ci les méthodes basées sur l'index de Poincaré sont très sensibles pour ébruiter et mener à la détection des fausses singularités dans des images de qualité faible. [May-06]

La figure.III.1 montre le point de noyau extrait à partir d'une image d'empreinte digitale.

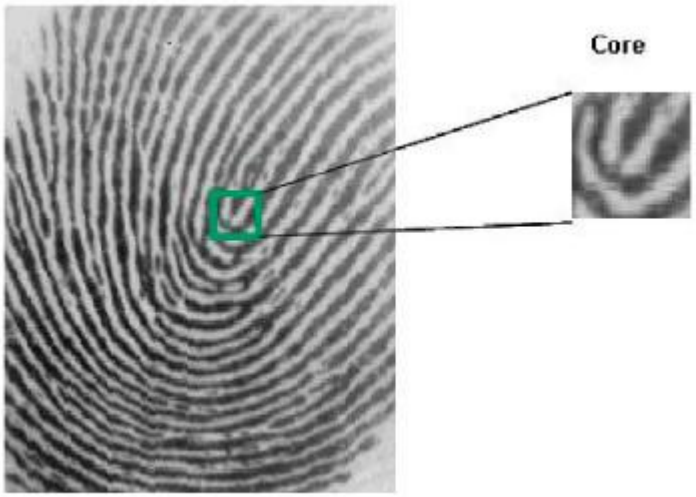


Figure III.1. Point core.

III.3. Matching

L'algorithme proposé par **Mayur D Jain('Microsoft (R&D) India Private Ltd.)** est un algorithme de matching qui fonctionne en un temps linéaire basé sur la construction d'arbre binaire. Ce dernier est construit à partir du point core qui est extrait. Les étapes de l'algorithme peuvent être mis en tant que (1) représentation des NNV (2) matching NNV, (3) incorporation NNV dans un arbre binaire(4) analyse de complexité de temps de matching. [May-06]

III.3.1. La représentation de vecteur des voisins les plus proches (NNV)

Chaque minutie dans l'empreinte digitale est représentée par les attributs (x, y, Ɵ). Pour construire une structure de NNV on utilise les « k » voisins les plus proches de la minutie. Comme le schéma 2 démontre un exemple de NNV utilisant quatre voisins les plus proches de la minutie. Si « a » est la minutie centrale, b, c, d et e sont ses quatre minuties voisines. Nous appelons la ligne qui relie la minutie centrale à la minutie voisine par le bord et chaque minutie comme sommet. On construit NNV en extrayant les quatre caractéristiques suivantes : [May-06]

1. La distance euclidienne entre la minutie centrale et leurs voisins, D_i, $i=1..4$.
2. La différence d'angles entre deux bords, α_i, $i=1..4$.

3. La différence d'angles entre l'orientation de minuties et son bord respectif, $\beta_{i,}$ $i=1..4$.
4. La différence d'angles entre l'orientation de la minutie centrale et l'orientation de son voisine, γ_i , $i_{=1..4.}$

Le vecteur NNV peut être représenté par ($D_{i,}$ α_i , β_i , γ_i) , $i=1..4$.

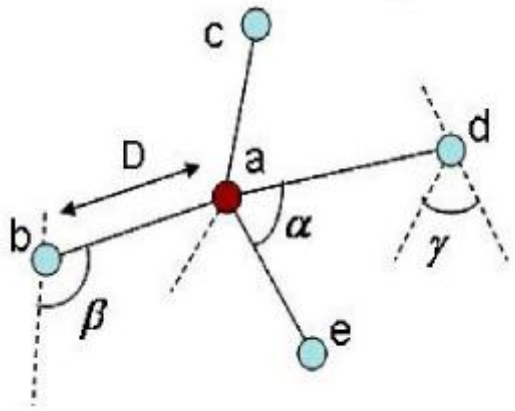


Figure III.2. La représentation de vecteur des voisins les plus proches.

III.3.2. Matching de NNV

L'algorithme de matching est basé sur l'assortiment de NNV formé de voisinage local et en propageant le matching en incorporant les points de minuties en structure arborescente binaire. On supposé que $(D_{1i},\alpha_{1i},\beta_{1i},\gamma_{1i})$ et $(D_{2i},\alpha_{2i},\beta_{2i},\gamma_{2i})$, $i=1$ *à* 4 deux NNVs de deux empreints digitales . Pour comparer les deux NNVs ; nous calculons NNV Score a partir la relative distance et les différences d'angle du α,β,γ :

1. Calculer la distance relative :

$$Rd_i = \left|D_{1i} - D_{2i}\right| / \min(D_{1i}, D_{2i}) \quad (1)$$

2. Calculer la différence d'angles :

$$\Delta\alpha_i = \left|\alpha_{1i} - \alpha_{2i}\right| \quad (2)$$

$$\Delta\beta_i = \left|\beta_{1i} - \beta_{2i}\right| \quad (3)$$

$$\Delta\gamma_i = \left|\gamma_{1i} - \gamma_{2i}\right| \quad (4)$$

$$NNVScore = f_d \sum_{i=0}^{4} Rd_i + f_\alpha \sum_{i=0}^{4} \Delta\alpha_i + f_\beta \sum_{i=0}^{4} \Delta\beta_i + f_\gamma \sum_{i=0}^{4} \Delta\gamma_i \quad (5)$$

Où f_d est le facteur de pondération pour la distance relative et f_α, f_β, f_γ sont les facteurs de pondération pour les orientations respectives, Ces facteurs de pondération sont une combinaison linéaire de la somme pour calculer les valeurs de NNV. [May-06]

III.3.3. Incorporation de matching NNV en arbre binaire

La construction de l'arbre binaire se produit simultanément dans l'empreinte de référence et l'empreinte digitale d'essai. Dans l'arbre binaire, des nœuds sont ajoutés de gauche à droite dans chaque niveau après seulement l'accomplissement du niveau précédent suivant les indications de figure III.3. Nous maintenons l'arbre binaire comme liste tels que chaque nœud pourrait être accédé dans le temps *O(l)*. [May-06]

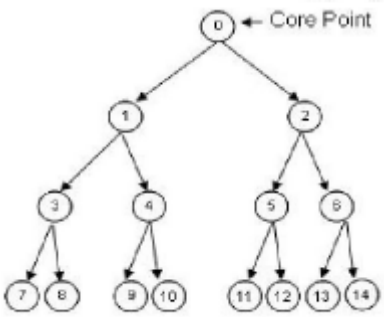


Figure III.3. Exemple d'un arbre binaire.

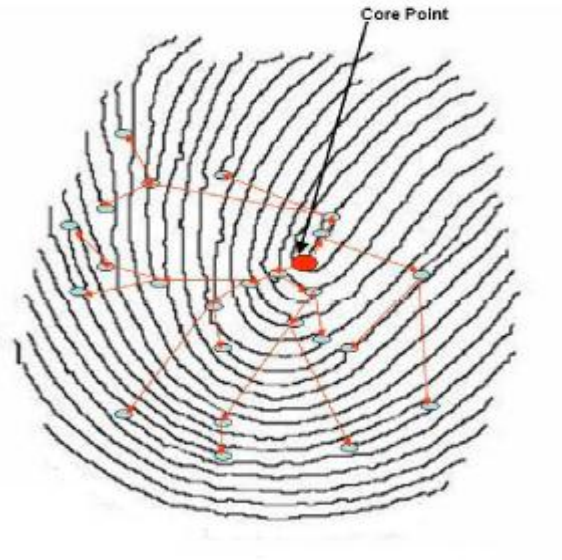


Figure III.4. Arbre binaire d'une empreinte digitale.

Supposer que R et S sont deux listes de minuties pour l'empreinte digitale de référence et l'empreinte digitale de test alors chaque minute de NNV est comparée comme expliqué ci-après et ajoutée dans leurs arbres binaires respectifs, un nul est ajouté à l'arbre binaire de l'empreinte digitale de test quand le matching échoue. Ceci aide principalement en propageant l'arbre binaire même lorsque le matching échoue. [May-06]

Algorithme de matching binary tree(BT)

1. Let R and S be set of minutiae pairs corresponding to reference and test images
2. Let CR & CS be core points from two fingerprints.
3. Let M represent set of matched minutiae pairs <R,S>.
4. Let RQ & SQ represent a FIFO queue.
5. Let Bt1 & Bt2 represent binary tree list of two fingerprints with roots (CR Cs).
6. Initialize a.ENQUEUE (RQ,CR).

   b.ENQUEUE (SQ,CS)
7. While (RQ is not empty or SQ is not empty)
   a. rq=DEQUEUE(RQ)
   b. sq=DEQUEUE(SQ)
   c. (<rql,rq2>,<sql,sq2>)=Find two nearest minutiae of (rq,sq)
   d. ENQUEUE (RQ,(rql,rq2))
   e. ENQUEUE(SQ,(sqj,sq2))
   f. If (NNV Match(rqj,sqj) = true)

      Add (rql,sql) as leftnode to Btl,Bt2list

      M= M + <rql,sql>

      else

      Add (rqj,dummy) as leftnode to Btl,Bt2list
   g. If (NNV Match(rq2 sq2) = true)

      Add (rq2 ,sq2) as rightnode to Btl,Bt2list

      M= M + <rq2,sq2>

      else

      Add (rq2,dummy) as rightnode to Btl,Bt2list
8. RETURN M (Match count obtained from M)

Finalement le nombre maximum de comparaison retenue pour calculer le score d'appariement. Le score est calculé comme suite :

$$Score = \frac{M^2}{Mr * Ms}$$

Où (*M)* représente le nombre de minuties comparé et (Mr) et (Ms) représente le nombre de minutie de l'empreinte digitale de référence et d'essai.

III.3.4. Analyse de complexité de temps

Afin d'examiner la complexité de matching des minuties, nous devons examiner la construction de l'arbre binaire. En se basant sur l'algorithme de la **Mayur D Jain**. Nous pouvons voir que les minuties sont manipulées dans une file d'attente pour la formation de l'arbre binaire, chaque opération de file d'attente comme ENQUEUE en file d'attente ou DEQUEUE prendre *O(l)*. Et à chaque nœud avant de l'ajouter à l'arbre binaire, nous effectuons la correspondance de NNV en considérant les K voisins dont la complexité est O= (k^2)=O(l) nous exécutons ces opérations sur chaque minutie. Ainsi la complexité moyenne de temps de NNV dans un arbre binaire prend Le TM = 0(l) 0(k2)0(l) 0(n) = 0 (n). [May-06]

III.4. Critiques et solutions

III.4.1. Critiques

- **la construction de l'arbre binaire**

L'arbre binaire est construit selon l'algorithme suivant :

- Soit E1 et E2 deux impressions de la même empreinte.
- Soit (A) l'arbre binaire de l'empreinte E1, et (B) l'arbre binaire de E2
- Soit (NA) un nœud de l'arbre (A) et (NB) un nœud de l'arbre (B)
- Trouver les deux minutes (LA, RA) les plus proches de la minutie (NA) dans l'empreinte E1.
- Trouver les deux minutes (LB, RB) les plus proches de la minutie (NB) dans l'empreinte E2.
- Faire le matching de RA avec RB et LA avec LB.

Le problème est que les minuties RB et LB ne sont pas toujours les mêmes minuties que RA et LA car les positions de RB et LB peuvent changer car l'image suite à un bruit additif ou que la structure des crêtes et assez endommagée. Aussi les minuties peuvent subir un changement suite à une rotation ou une translation de l'une des deux images.

Les problèmes précités induisent à une fausse construction de l'arbre binaire de l'une des empreintes.

- **Le descripteur de la minutie**

Le descripteur proposé par **Mayur D Jain** pose le même problème que précédemment car il consiste à trouver les quatre voisins les plus proches de la minutie puis calculer les facteurs D_i, α_i , β_i , γ_i Un changement de la position ou une disparition de l'une des quatre voisins conduit à un changement directe de la valeur du descripteur et ça refléter sur le score de matching du deux minuties semblable du deux copie de la même empreinte.

III.4.2. Solution

- ✓ pour le premier problème on propose la solution suivante :
  - Trouver les deux minutes (LA, RA) les plus proches de la minutie (NA) dans l'empreinte E1.
  - Trouver les N minuties (M_1, M_2 …. M_n) les plus proches de minutie B.
  - Faire le matching de RA avec {M_1,M_2,…,M_n).
  - Choisir M_i qui a le maximum score de matching avec RA.
  - Faire le matching LA avec {M_1,M_2,…,M_n}-{M_i}.
  - Choisir M_k qui à le maximum score de matching avec LA.
- ✓ Pour le deuxième problème nous avons utilisé comme descripteur de minutie le descripteur proposé par TICO puisque il est Indépendant de la rotation et de la translation de l'image de l'empreinte mais dépend de l'orientation de la minutie.

TICO propose de caractériser l'endroit de chaque minutie en ce qui concerne le modèle d'empreinte digitale d'entrée basé sur un descripteur qui comporte des informations sur le champ d'orientation dans une large région autour du point de minutie. Puisque l'orientation d'arrête montre typiquement de petites variations spatiales entre les pixels de

voisinage, des vastes zones du champ d'orientation peuvent être reconstruites des angles d'orientation prévus dans un nombre relativement petit de points de prélèvement. [Tic-01]

Les points de prélèvement assignés à chaque minutie peuvent être organisés dans un modèle circulaire autour de la position de minutie $m=[x, y]^T$, comme illustré dans la (figure. III 1).

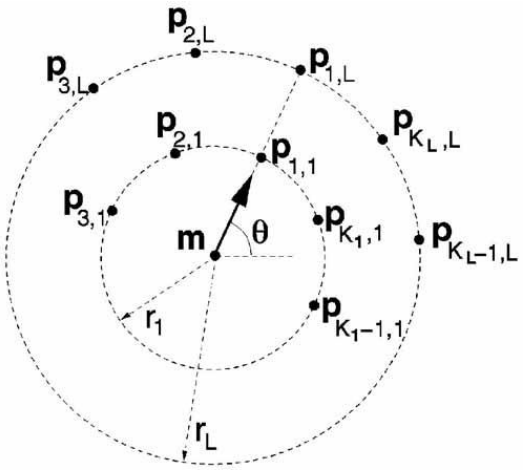


Figure III.5. Points d'échantillonnage organisés dans un modèle circulaire autour d'une minutie.

Le modèle circulaire se compose de L cercles concentriques des rayons r_ℓ , $(1 \le \ell \le L$), chacun de eux comportant des points de K_ℓ , échantillonnage $P_{k,\ell}$, $(1 \le k \le K_\ell)$, également distribués le long de sa circonférence. L'utilisation de la direction(θ) de minutie comme référence, les points sur chaque cercle peuvent être ordonnés d'une façon dans le sens contraire des aiguilles d'une montre commençant par le point situé avec le respect à la position de minutie. On dénote par $\theta_{k,\ell}$ l'orientation locale d'arête prévue en $P_{k,\ell}$, nous définissons le descripteur de minutie comme suit :

$$\mathrm{f} = \left\{ \left\{ \lambda\left(\theta_{k,l}, \theta\right) \right\}_{k=1}^{K_\ell} \right\}_{\ell=1}^{L} \qquad (2)$$

Le descripteur de minutie (2) est invariable à la rotation et la translation et, par conséquent, lui peuvent caractériser l'endroit de minutie en ce qui concerne le modèle d'empreinte digitale indépendamment de la position et de l'orientation du doigt sur le

capteur d'acquisition. En outre, une propriété utile du descripteur proposé se compose de son indépendance en ce qui concerne n'importe quelle autre minutie détectée dans l'image d'entrée.

Algorithme de matching modifie(MBT)

6. Let R and S be set of minutiae pairs corresponding to reference and test images
7. Let CR & CS be core points from two fingerprints.
8. Let M represent set of matched minutiae pairs <R, S>.
9. Let RQ & SQ represent a FIFO queue.
10. Let Bt1 & Bt2 represent binary tree list of two fingerprints with roots (CR Cs).

6. Initialize a.ENQUEUE (RQ,CR).
   b.ENQUEUE (SQ,CS)
9. While (RQ is not empty or SQ is not empty)
   h. rq=DEQUEUE(RQ)
   i. sq=DEQUEUE(SQ)
   j. - (<rq1,rq2>)= Find two nearest minutiae of (rq)
      - (<sql,sq2,…,sqn>)=Find n nearest minutiae of (sq)
   k. - Get <rq1, sqj> where score <rq1, sqj> is max.
      Remove sqj from (<sql,sq2,…,sqn>).
      - Get <rq2, sqk> where score <rq2, sqk> is max.
   l. ENQUEUE (RQ,(rql,rq2))
   m. ENQUEUE(SQ,(sqj,sqk))
   n. If (Match(rq1,sqj) = true)
      Add (rql,sqj) as leftnode to Btl,Bt2list
      M= M + <rql,sqj>
      else
      Add (rq1,dummy) as leftnode to Btl,Bt2list
   o. If (Match(rq2 sqk) = true)
      Add (rq2 ,sqk) as rightnode to Btl,Bt2list

```
            M= M + <rq2,sqk>
          else
            Add (rq2,dummy) as rightnode to Btl,Bt2list
10. RETURN M (Match count obtained from M)
```

III.5. Conclusion

Ce chapitre a été dédié à l'étape de matching qui est originalement un problème d'optimisation. Nous avons étudié l'algorithme de **Mayur D Jain** dont l'idée primordiale est d'organiser les minuties sous forme d'un arbre binaire. Ayant une complexité linéaire, cet algorithme présente l'inconvénient de taux élevés en terme de FAR et FRR. Nous avons proposé quelques modifications pour améliorer ses résultats.

Chapitre

IV

Conception, mise en œuvre & évaluation

IV.1. Introduction

Afin de mettre en œuvre un logiciel, il faut d'abord étudier son cycle de vie (en anglais software lifecycle), qui désigne toutes les étapes de développement d'un logiciel, de sa conception à sa disparition. L'objectif d'un tel découpage et de permettre de définir des jalons intermédiaires permettant la validation du développement logiciel, c'est-à-dire la conformité du logiciel avec les besoins exprimés, et la vérification du processus de développement, c'est -à-dire l'adéquation des méthodes mises en œuvre.

Notre conception sera confirmé aux activités suivante faisant référence à un model connu dit modèle en cascade.

- Analyse des besoins : c'est-à-dire l'expression, le recueil et la formalisation des besoins et des objectifs attendus.
- Spécification fonctionnelle : il s'agit de définir l'architecture logicielle de système en présentant les principales fonctionnalités à implémenter.
- Découpage modulaire (conception globale) : il s'agit de l'élaboration de tous les modules nécessaires.
- Conception détaillée : consistant à définir précisément le fonctionnement algorithmique de chaque module.
- Codage (implémentation ou programmation) : soit la traduction dans un langage de programmation des fonctionnalités définies lors de la phase de conception.

Tests : permettant de vérifier individuellement que chaque sous-ensemble du logiciel est implémenté conformément aux spécifications.

IV.2. conception

IV.2.1. Analyse des besoins

Nous avons comme objectif principal la réalisation d'un système automatique d'identification d'empreinte digitale (AFIS : Autoamted Fingerprint Identification System) qui réunit quelques exigences de fiabilité.

Nous avons spécifié ses exigences par les points suivants :

- Avoir un système d'identification d'empreinte digitale fiable en termes de taux d'erreur (FAR et FRR).
- Avoir un système rapide en terme de temps de match.
- Valider notre système par des tests en discutant ses performances (le taux d'erreur et le temps de matching)**.**

IV.2.2. Spécification fonctionnelle

La conception sera modélisée à l'aide de langage UML(Unified Modeling Language) qui est un langage de modélisation graphique à base de pictogrammes. Il est apparu dans le monde du génie logiciel, dans le cadre de la « conception orientée objet ». Couramment utilisé dans les projets logiciels, il peut être appliqué à toutes sortes de systèmes ne se limitant pas au domaine informatique. [3]

IV.2.2.1. Les acteurs du système

Un acteur représente un rôle joué par une entité externe (utilisateur humain, dispositif matériel ou autre système) qui interagit directement avec le système étudié. [Pas07]

Le tableau suivant présente les différents acteurs de notre système.

Acteur	Description
 Administrateur	L'administrateur a pour mission : - Lancer le matching. - Etablir des statistiques
BDD	Contient un ensemble des empreintes pour faire le test de système.
Empreinte	Objet de test de matching

Tableau IV.1. Les acteurs du système.

IV.2.2.2. Détermination des cas d'utilisation

Un *cas d'utilisation* (*use case*) représente un ensemble de séquences d'actions réalisées par le système et produisant un résultat observable intéressant pour un acteur particulier. [Pas-07]

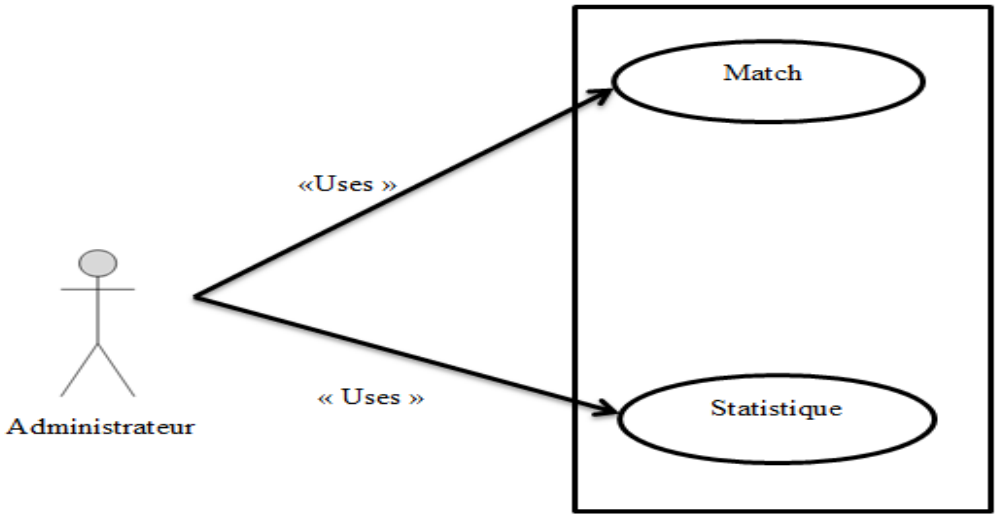


Figure IV.1. Diagramme de cas d'utilisation de système.

Description des cas d'utilisation

Description de cas d'utilisation de « Matching »
Titre : comparaison entre deux empreintes. **Description** : ce cas d'utilisation permet à l'administrateur de faire un matching entre deux empreintes. Lors de cette phase le système extrait les caractéristiques de chaque empreinte, et rend le résultat de matching entre les deux empreintes. **Acteurs :** administrateur, empreinte.

Description de cas d'utilisation « Statistiques »
Titre : Statistique **Description :** ce cas d'utilisation permet à l'administrateur d'établir des statistiques sur le système, en effet cette phase nous permet de faire des tests sur une base d'empreinte donnée. Ceci permet d'évaluer et de mesurer les performances de notre système. **Acteur :** l'administrateur, base d'empreinte.

IV.2.2.3. Diagramme de séquence de cas d'utilisation « Matching »

Dans ce cas, on présente la séquence des scénarios du cas d'utilisation Matching ».

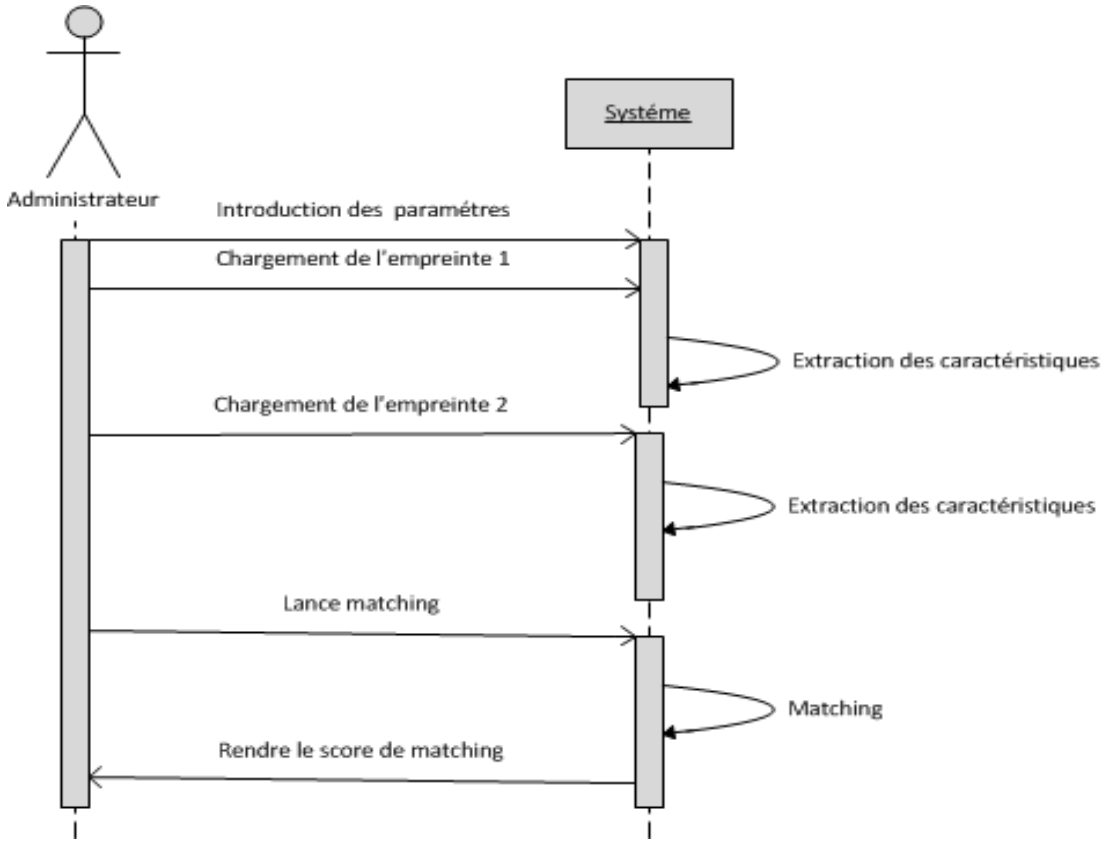


Figure IV.2. Diagramme de séquence de cas d'utilisation « Matching ».

Description des scénarios

Scénarios de cas d'utilisation « Matching »

- L'administrateur initialise les paramètres de matching.
- L'administrateur charge l'empreinte 1.
- Le système extrait les caractéristiques de l'empreinte 1.
- L'administrateur charge l'empreinte 2.
- Le système extrait les caractéristiques de l'empreinte 2.
- L'administrateur lance le matching.
- Le système fait le match et rend le score de matching.

IV.2.2.4. Diagramme de séquence de cas d'utilisation « Statistique »

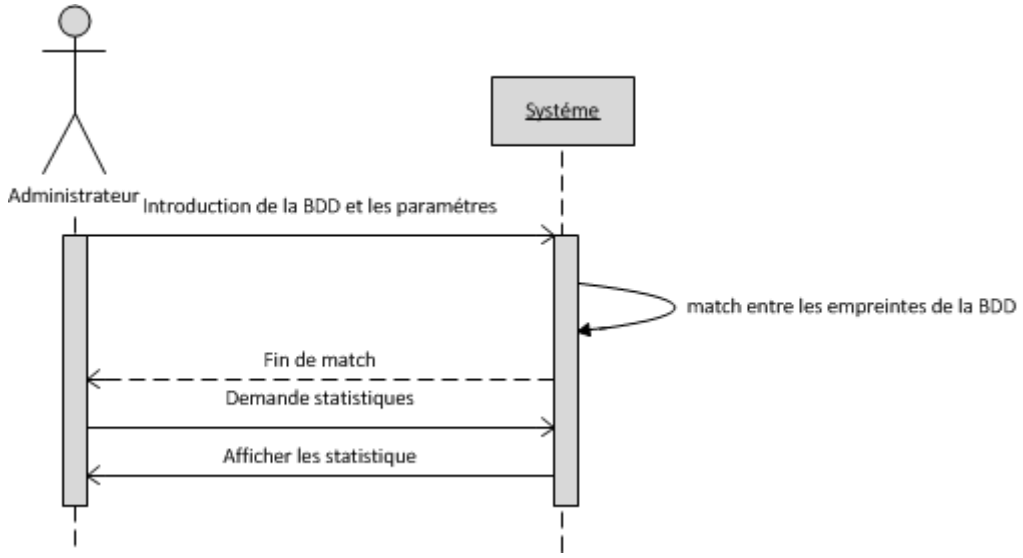


Figure IV.3. Diagramme de séquence de cas d'utilisation « statistiques ».

Description des scénarios

Scénarios de cas d'utilisation « statistiques »
- L'administrateur charge la BDD et initialise les paramètres. - Le système fait le matching entre les empreintes de la BDD et calcule les statistiques et informe l'administrateur la fin de l'exécution. - L'administrateur demande les statistiques - Le système affiche les statistiques.

IV.3. Mise en oeuvre

Après avoir conçu une solution qui, nous espérons, répondrait mieux aux objectifs fixés préalablement, nous entamons l'étape de réalisation de la plateforme logicielle de notre système.

Dans cette partie, nous mettrons en évidence les raisons de nos choix techniques (langage de programmation et les outils utilisés). Ensuite, nous allons citer les principales classes développées et décrire les différentes fenêtres de l'application.

IV.3.1. Outils de développement

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

Afin de réaliser notre système et son interface, nous avons eu recours aux outils : Microsoft Visual Studio 2010, .Net Framework.

Microsoft Visual Studio 2010 :

Microsoft Visual C# 2010 est un langage de programmation permettant la création d'une large gamme d'application s'exécutant sur le .Net Framework. C# est simple, puissant et orienté objet. Avec ses nombreuses innovations, C# permet le développement rapide d'applications tout en conservant la simplicité et l'élégance de style C.

La plateforme .Net contient une collection riche de classes et de bibliothèques prédéfinies qui facilitent l'implémentation des algorithmes.

Un Développement robuste qui permet la modification et le débogage efficace du code source.

Possibilité d'exploiter les différentes DLL (Dynamic Link Library) crées par d'autres langages que le Visual C# sous la même plateforme.

IV.3.2. Description des différentes classes

L'ensemble des classes et les principales méthodes développées sont présentées dans le tableau suivant :

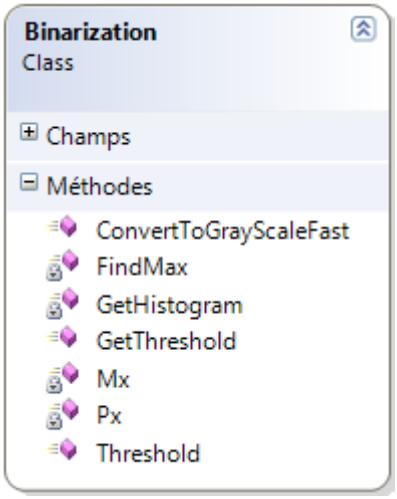
Classe	Signification
Enhancement	Améliore la qualité de l'image
Binarization	Transforme l'image d'empreinte digitale en image binaire.
Skeleton	Transforme l'image binaire en squelette.
FeatureExtractor	Extrait le point core et les minuties.
BinaryTreeMatching	Matching avec l'algorithme de [May-06]
MBinaryTreeMatching	Matching avec notre algorithme.
FVCDB	Calcul ldes statistiques.
Config	Contient les différents paramètres

Tableau IV.2. Les principales classes de notre système.

- **Classe Enhancement**

Les algorithmes de reconnaissance des empreintes digitales sont sensibles à la qualité des images d'empreintes digitales. Alors une étape d'amélioration devient nécessaire avant d'effectuer les étapes d'extraction des caractéristiques. Pour ce faire, nous avons utilisé une DLL **AFIS.** [4]

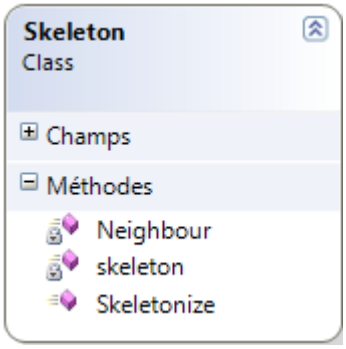
- **Classe Binarization**



Cette classe convertie une image de type BMP en une image Binaire.

- ***ConvertToGrayScalFast*** : Convertie l'image en image en niveau gris
- ***FindMax*** : retourne le maximum de niveaux de gris.
- ***GetHistogram*** : Etablit l'histogramme de l'image en niveau de gris.
- ***Threashold*** : donne le seuil de binarisation.

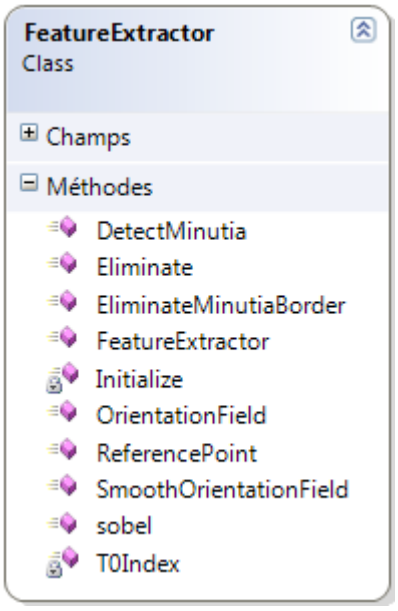
- **Classe Skeleton**



Cette classe convertit une image Binaire en une image Squelette.

- ***Neighbour* :** donne les 8 voisins d'une minutie.
- ***Skeletonize*** : faire la squelettisation.

- **Classe FeatureExtractor**

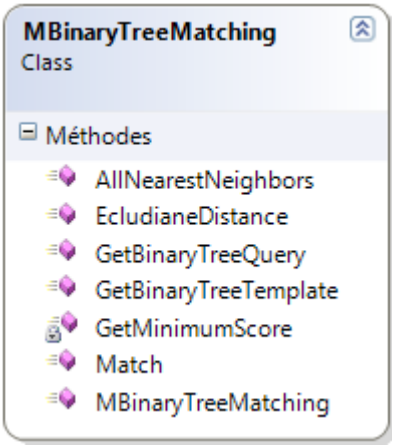


Cette classe utilisée pour l'extraction de point core et les minuties et pour l'élimination des fausses minuties.

- ***DetectMinutia*** : Extrait la liste des minuties.
- ***Eliminate*** : Elimination des fausses minuties.
- ***ElininateMinutiaBorder*** : Elimination des minuties qui se situent sur les bordures de l'empreinte digitale.
- ***OrientationField*** : Calcule le champ d'orientation.

- ***ReferencePoint*** : Calcule les coordonnées du point core.
- ***SmoothOrientationField*** : Calcule le champ d'orientation lissé.

- **Classe BinaryTreeMatching**

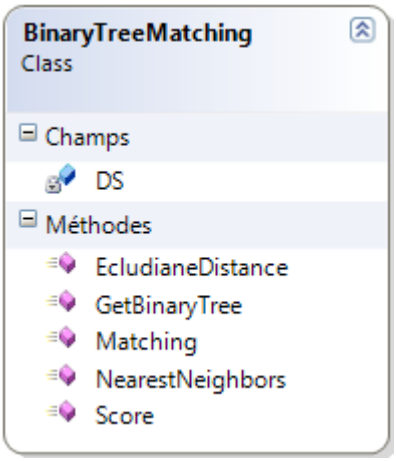


Cette classe fait le matching avec l'algorithme de Mayur D Jain [May-06] et rende le score.

- ***EcludianeDistance*** : Calcule la distance euclidienne entre deux minuties
- ***GetBinaryTree*** : Constuit l'arbre de l'empreinte (Query).
- ***Matching*** : Donne le score du matching
- ***NearestNeighbors*** : Donne les plus proches voisins d'une minutie.

Score : Matching entre deux minuties.

- **Classe MBinaryTreeMatching**

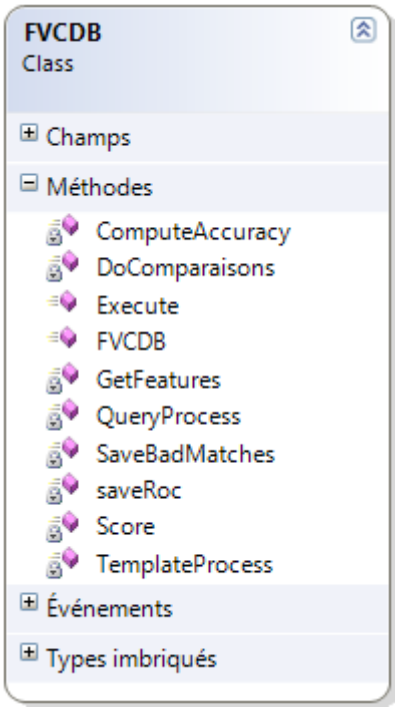


Cette classe fait le matching avec l'algorithme modifie par nous et donne le score de matching.

- ***AllNearestNeighbors*** : Donne les plus proches voisins d'une minutie.
- ***EcludianeDistance*** : Calcule la distance euclidienne entre deux minuties
- ***GetBinaryTreeQuery*** : Construit l'arbre de l'empreinte (Query).
- ***GetBinaryTreeTemplate*** : Construit l'arbre de l'empreinte (Template).
- ***GetMinimumScore*** : Donne le minimum score de matching entre les voisins d'une minutie.
- ***Match*** : Donne le score de matching.

- **Classe FVCDB**

Cette classe utilisée pour calculer les statistiques.



- ***ComputeAccuracy*** : Génération de la courbe ROC.
- ***DoComparaisons*** : Charge les listes des génuines, des imposteurs et listes des FNMRs, FMRs.
- ***GetFeatures*** : Charge la liste des empreintes.
- ***QueryProcess*** : Extraction des caractéristiques de l'empreinte(Query)
- ***SaveBadMatches*** : Sauvegarde les listes (fausses acceptation rates et fausses rejets rate) dans des fichiers « .CSV »
- ***SaveRoc*** : Sauvegarde le ROC dans un fichier « .CSV »
- ***Score*** : Donne le score de matching.
- ***TemplateProcess*** : Extraction des caractéristiques de l'empreinte(Template)

- **Classe Config**

Cette classe contient les différents paramètres du système.

Paramètre	Signification
MinThreashold	Le minimum seuil de matching
MaxThreashold	Le maximum seuil de matching
NumberMinutiae	Le nombre minimum de minuties pour décider qu'une personne est un génuine ou imposteur.
CountNeighbors	Le nombre de voisins d'une minutie.
NNVThreashold	Le seuil de match pour l'algorithme BinaryTree.

Tableau IV.3. Les différents paramètres de notre système.

IV.3.3. Présentation de l'application

On présente dans cette section les différentes interfaces de l'application.

a) L'interface principale

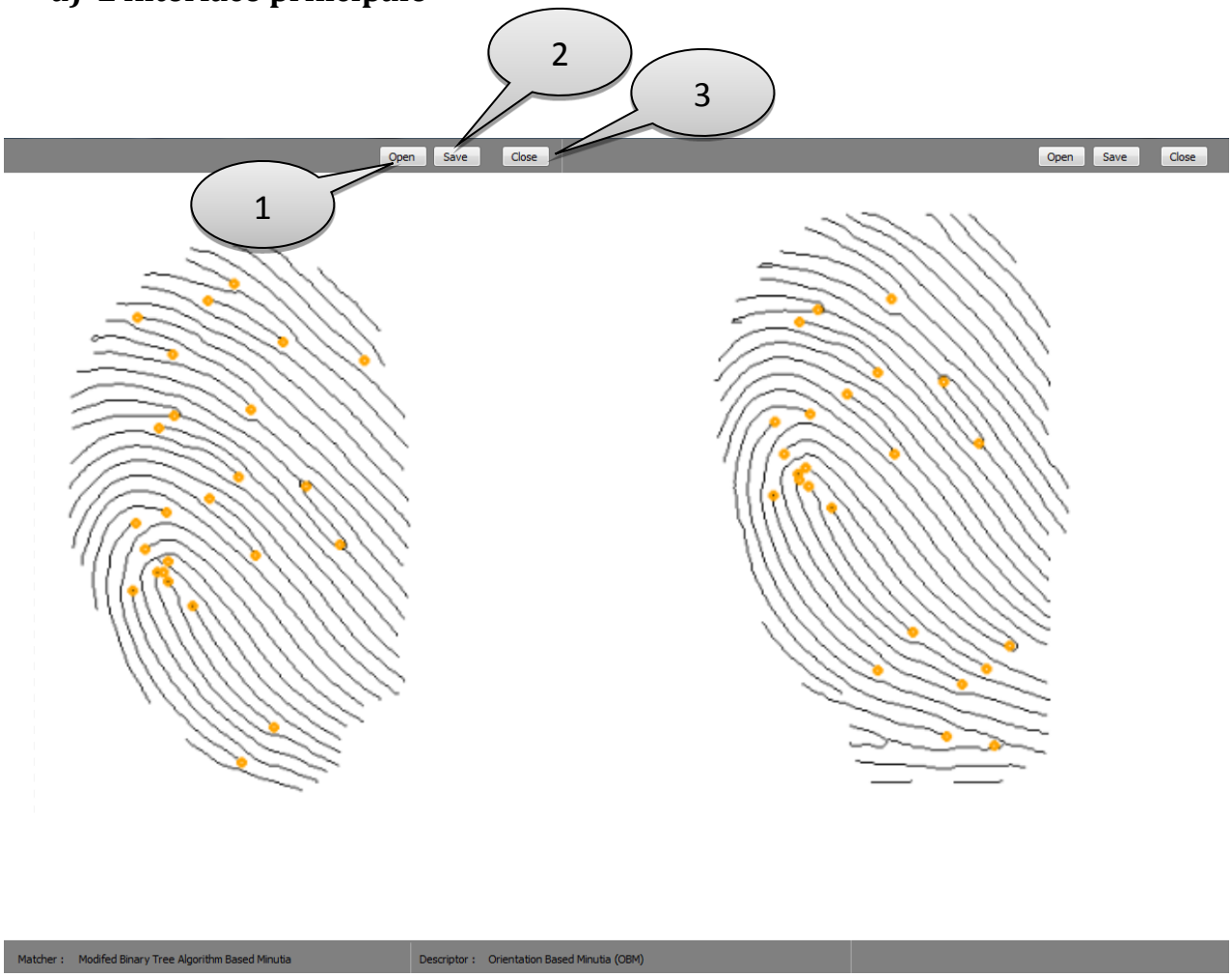


Figure IV.4. La forme principale.

1 Chargement de l'empreinte.

2 Sauvegarder l'empreinte.

3 Déchargement l'empreinte.

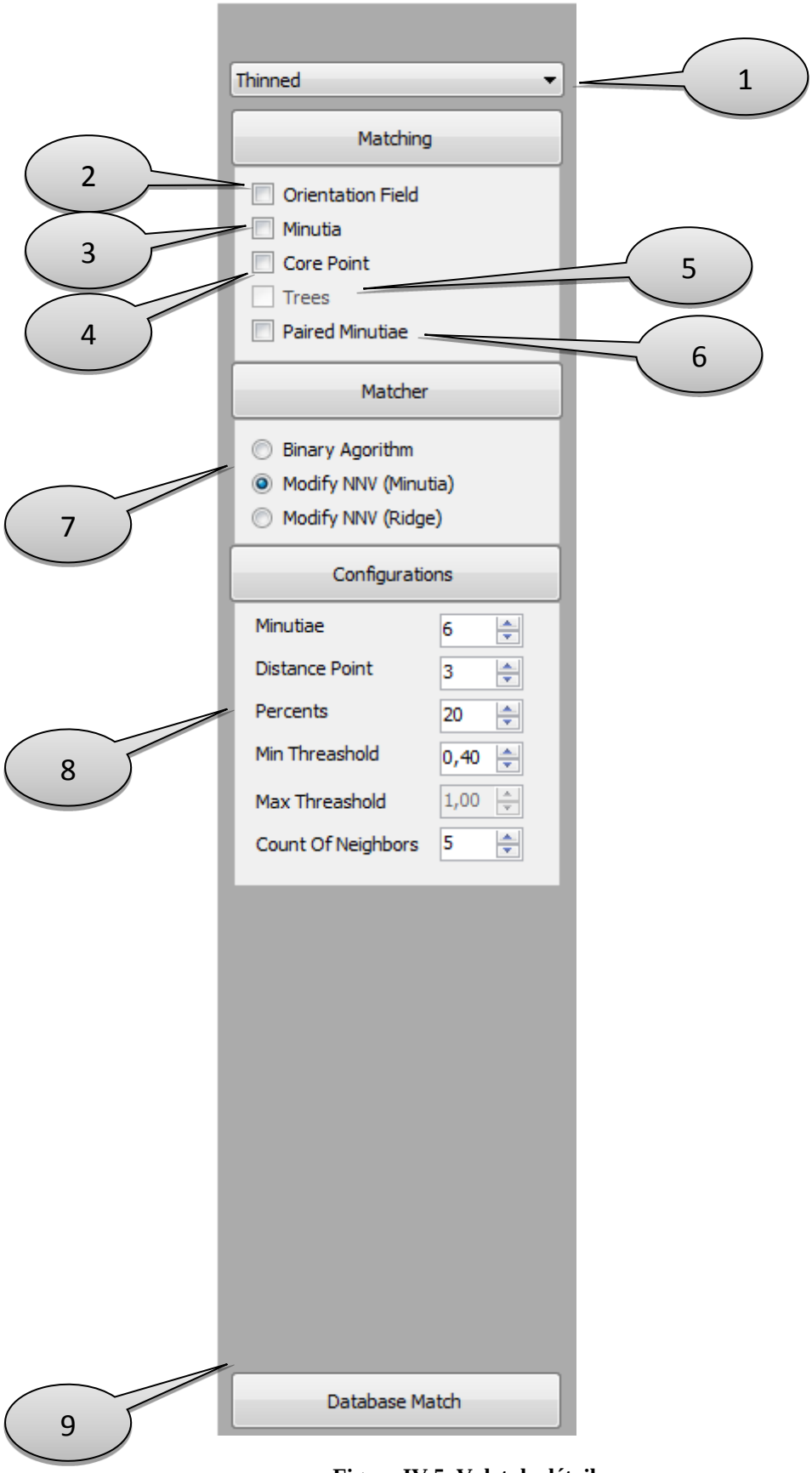


Figure IV.5. Volet de détails.

1. Affiche les différents filtres (Amélioration, Binarisation et squelettisation).
2. Affiche le champ d'orientation de l'empreinte.
3. Affiche la liste des minuties.
4. Affiche le point core.
5. Afficher les sous- arbres de l'empreinte.
6. Fait le match et affiche les paires de minuties et le score.
7. Choix de l'algorithme de match.
8. Choix des paramètres
9. Expérimentation et statistiques

b) L'interface d'expérimentation

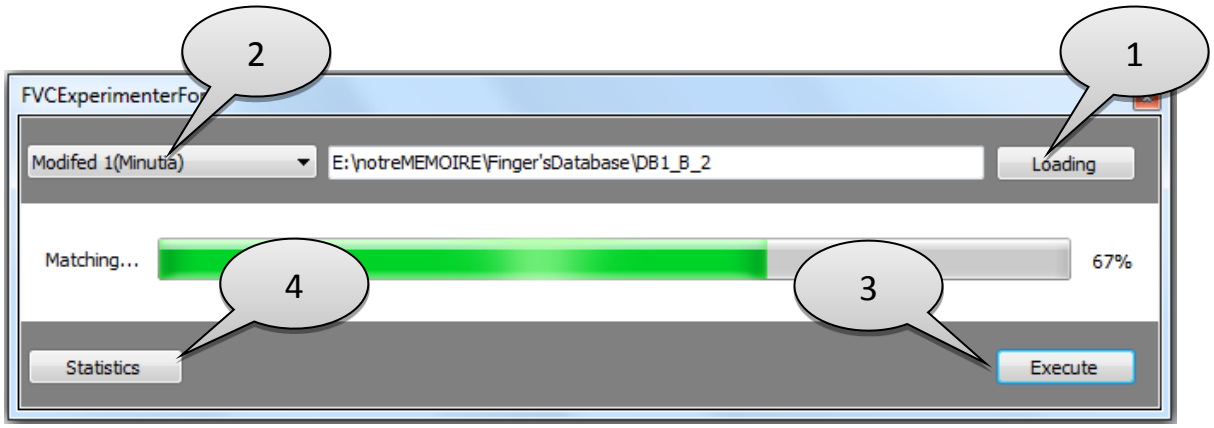


Figure IV.6. Teste et expérimentations.

1. Charge la BDD.
2. Choix de l'algorithme de match.
3. Exécute le test.
4. Affiche les statistiques.

IV.4. Evaluation

Afin d'évaluer l'algorithme de matching de **Mayur D Jain** et **notre algorithme**, nous avons utilisé les indices (FAR et FRR) et le temps de matching. Les bases FVC2002 (DB1, DB2) et FVC2004 (DB1, DB2) [5] ont été utilisées pour l'évaluation.

Le tableau suivant résume les résultats obtenus

		FAR%	FRR%	EER%	ZEROFMR%	TEMPS (ms)
Algorithme BT	DB1	69.166	28.928	43,73	99,643	1,917
	DB2	83.888	7.142	54,226	100	3,78906
Algorithme MBT	DB1	14.166	26.428	17,996	79,286	6.615
	DB2	17.222	14.285	13,750	54,643	8.235

Tableau IV.4. Résultat du test sur la base FVC2002 (DB1, DB2)

		FAR%	FRR%	EER%	ZEROFMR%	TEMPS (ms)
Algorithme BT	DB1	71.111	16.428	39,782	96,429	3,014
	DB2	72.222	18.928	42.52	95.00	2.15
Algorithme MBT	DB1	11.666	19.624	15,675	60,000	8,03906
	DB2	11.338	28.571	22,5	57.5	8.723

Tableau IV.5. Résultat du test sur la base FVC2004(DB1, DB2).

Les tests a été effectué sur un ordinateur (Pentium IV, 3.6 Ghz, 1G de RAM).

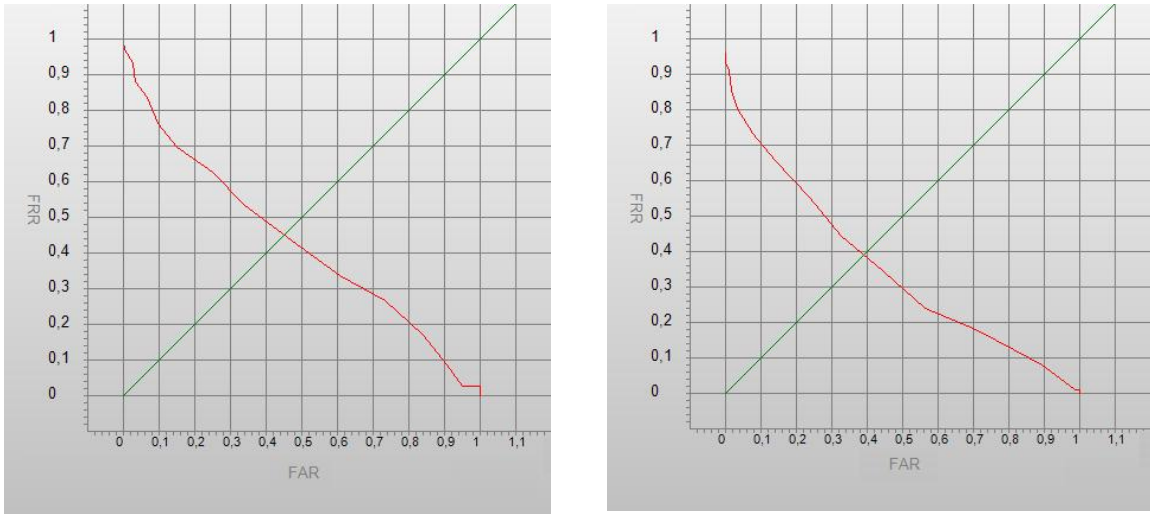


Figure IV.7. ROC de FVC2002(DB1, DB2) de l'algorithme BT.

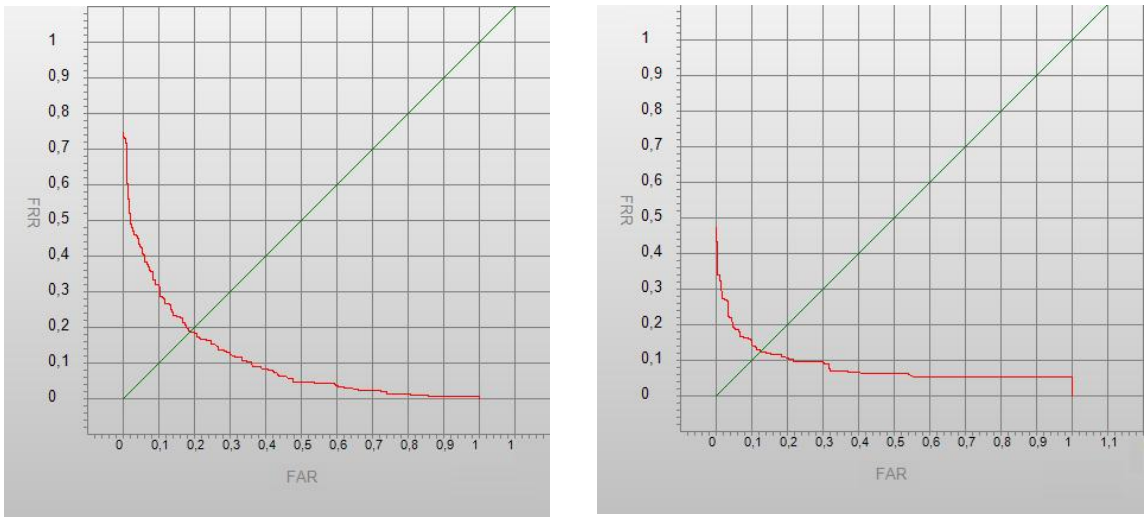


Figure IV.8. ROC de FVC2002(DB1, DB2) de l'algorithme MBT

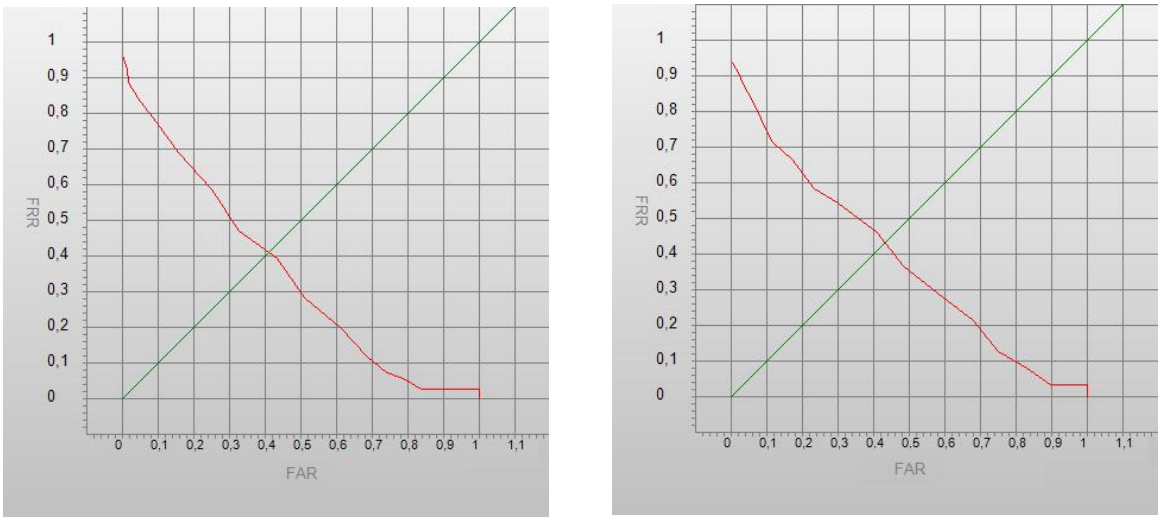


Figure IV.9. ROC de FVC2004(DB1, DB2) de l'algorithme BT.

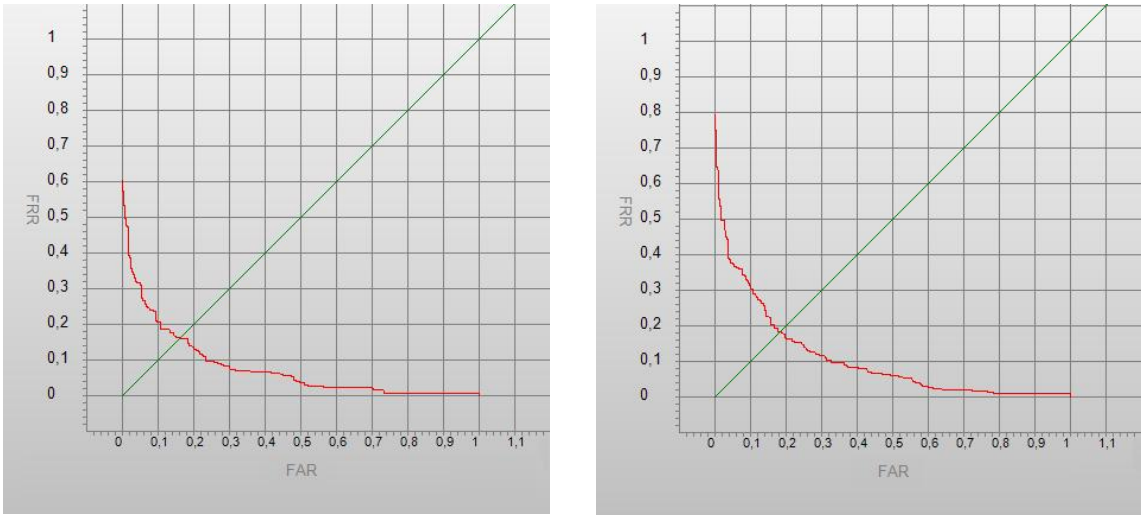


Figure IV.10. ROC de FVC2004(DB1, DB2) de l'algorithme MBT.

Commentaires sur les résultats

1. L'algorithme de **Mayur D Jain** donne des mauvais résultats en terme de (FAR, FRR, EER), au raison des problèmes que nous avons décrits dans **III.4.1.**
2. L'algorithme de **Mayur D Jain** est plus rapide puisque il fonctionne en temps linéaire.
3. Notre algorithme donne des meilleurs résultats que l'algorithme de **Mayur D Jain** en terme de (FAR, FRR, EER).
4. Le temps de matching de notre algorithme est moins rapide car il utilise un descripteur robuste mais prend un temps en matching.
5. Les taux d'erreurs de notre algorithme reste élevés par rapport des autres algorithmes come MTK de Miguel Angel Medina-Pérez [Mig-09], et l'algorithme de Xudong Jiang [Jia-00].

IV.5. Conclusion

Nous avons présenté dans ce chapitre l'étude conceptuelle de notre système en respectant la modélisation qui a été établie précédemment et ce, pour répondre, aux objectifs déclarés, ainsi nous avons détaillé l'implémentation en Visual C# en expliquant les principales classes que nous avons développées, et enfin nous avons établit les différents résultats de test du système.

Conclusion générale

Dans ce mémoire, nous avons donné un aperçu sur la biométrie en se focalisant sur l'empreinte digitale comme une modalité importante. Il était question dans ce travail de réaliser un système automatique d'identification par empreinte digitale.

- Une grande partie de ce travail a été dévolue à l'extraction des caractéristiques de l'empreinte digitale. Pour ce faire, le processus adopté se projette suivant le flux de travail suivant :
  - L'image subit une phase d'amélioration (Image Enhancement) :
  - L'image subit une phase de binarisation.
  - Détection de point core.
  - L'image subit une phase de squelettisation.
  - Les minuties sont extraites grâce aux nombre de connexions.
  - Les minuties résultats sont filtrées par un algorithme de post-traitement pour éliminer les fausses détections.
- Une deuxième partie de ce travail a été dédiée à l'étape de matching qui est originalement un problème d'optimisation. Nous avons étudié l'algorithme de **Mayur D Jain** dont l'idée primordiale est d'organiser les minuties sous forme d'un arbre binaire. Ayant une complexité linéaire, cet algorithme présente l'inconvénient de taux élevés en terme de FAR et FRR. Nous avons proposé quelques modifications pour améliorer ses résultats. L'évaluation des deux algorithmes a été menée sur les bases de données «FVC2002, FVC2004 ».

A la lumière de ce qui a été fait, l'objectif préliminaire a été en substance atteint. Mais en outre, des idées futures peuvent se dessiner.

Bibliographie

A

[Arc-84]: C.Arcelli, D.Baja. "A Width Independent Fast Thinning Algorithm". IEEE Trans on PAMI, Vol. 4, No. 7, pp.463-473.

B

[Bar-88]: O.Baruch. "Line Thining by Line Folowing", Pattern Recognition Letters, Vol.8, No.4, pp. 271-276, 1984

[Bel-06]: Belguechi Rima, "Contribution à la Reconnaissance D'empreinte Digitale par une Approche Hybride", présentée pour obtenir le grade de Magister de l'institut Nationale de Formation en Informatique (I.N.I), Alger 2009.

C

[Coe-93]: L.Coetzee. "*Fingerprint Recognition In Low Quality Images*", Pattern Recognition, Vol. 26, No. 10, pp. 1661-1671, Novembre 1993.

D

[Dav-03]: Davide Maltoni et al: " *Handbook of Fingerprint Recognition ",* 2003.

H

[Hun-93]: D.C.D.Hung. "*Enhancement And Feature Purification Of Fingerprint Images*", Pattern Recognition, Vol. 26, pp. 1661-1671, 1993.

J

[Jai-00] : Xudong Jiang and Wei-Yun Yau, "Fingerprint Minutiae Matching Based on the Local And Global Structures ",0-7695-0750-6/00, 2000 IEEE.

[Jan-97] : A.Jain, L.Hong, S.Pankanti, R.Bolle. "An Identity-Authentification system Using Fingerprints", Proceedings of The IEEE, Vol.85, September 1997.

[Joh-06]: John R. Vacca : "*Biometric Technologies and Verification Systems* "2006.

L

[Lin-98]: Lin Hong, Yifei Wan, Anil Jain, "*Fingerprint image enhancement: algorithm and performence evaluation*", 1998

M

[Mai-98]: D.Maio,D.Maltoni. "*Neural Network Based Minutiae Filtering in Fingerprint Images*", In14th International Conference on Pattern Recognition, pages 1654–1658, 1998.

[May-06]: Mayur D Jain , Nalin Pradeep, Prakash et Balasubramanian Raman. "*Binary Tree Based Linear Time Fingerprint Matching*",1-4244-0481-9/06/2006 IEEE.

[Med-09] : Miguel Angel Medina-Pérez, Andrés Gutiérrez-Rodríguez,and Milton García-Borroto, "*Improving Fingerprint Matching Using an Orientation-Based Minutia Descr*", CIARP 2009, LNCS 5856, pp. 121–128, 2009.

N

[Nan-04]: K.Nandakumar, A.K.Jain. : "*Local correlation-based fingerprint matching*", Indian Conference on Computer Vision, Graphics and Image Processing, pp. 503–508, 2004.

[Nic-09]: Nicolas MORIZET : "*Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris*"présentée pour obtenir le grade de Docteur de l'Ecole Nationale Supérieure des Télécommunications paris 2009.

O

[Ots-79]: N.Otsu. *"A Threshold Selection Method from Gray Level Histograms"*, IEEE Transactions on Systems, Man and Cybernetics, Vol.9, pp.62–66, 1979.

P

[Pas-07]: Pascal Roques, Franck Vallée, *"UML 2 en action de l'analyse des besoins à la conception"*, 4éme edition EROLLES 2007.

[Pra-99]: S.Prabhakar, A.Jain, L.Hong, S.Pankanti. *"FingerCode: A Filterbank for Fingerprint Representation and Matching"*,1999.

R

[Rat-95] : N.Ratha, S.Chen, A.Jain. *"Adaptive Flow Orientation Based Texture Extraction in Fingerprint Images"*, Pattern Recognition, Vol.28, No; 11, pp.1657-1672, Nov 1995.

[Ros-02]: A.K.Jain, A.Ross. *"Fingerprint Mosaicking"*, IEEE Intentional Conference on ICASSP,Vol. 4, May 2002.

S

[Syl-04]: Sylvie Craipeau et al : *"La Biometrie : usages et rerésentation"* 2004.

T

[Tic-01]: Marios Tico, *"On Design and Implementation of Fingerprint-Based Biometric Systems",* 2001.

V

[Var-00]: G. T. B Verwaaijen, A. M. Baze, S. H. Garez, L. P. J. Veelunturf, et B. J. van der Zwaag : *"A Correlation-Based Fingerprint Verification System"*, IEEE ProRISC2000 Workshops on Circuits, Systems and Signal Processing, Nov 2000.**[Vér-06]**: Véronique Messéant et al : *" Les empreintes digitales"*Master Didactique des Mathématiques - Université Paris VII -Juin 2006.

W

[Wu-03]: C.Wu, J.Zhou, Z.Bian, G.Rong. "*Robust Crease Detection in Fingerprint Images*", InConference on Computer Vision and Pattern Recognition, 2003.

X

[Xia-91]: Q.Xiao, H.Raafat. "Combining Statistical and Structural Information For Fingerprint Image Processing", In Pattern Recognition Architectures, Algorithms and Applications, 1991.

Webographie

[1] : http://www.Biometrie-online.net, 2011

[2] : http://jeff.cs.mcgill.ca/~godfried/teaching/projects97/azar/, 2011

[3] : http://www.omg.org/spec/UML/2.3/Superstructure , 2011.

[4] : http://www.sourceforge.net/projects/sourceafis/files/SourceAFIS-1.4.0.msi, 2011.

[5] : http://bias.csr.unibo.it/, 2011