الجمهورية الجزائرية الدمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 20 Août 1955-Skikda
Faculté des Sciences
Département d'Informatique
D012114001M

جامعة 20 اوت 1955- سكيكدة
كلية العلوم
قسم الإعلام الآلي

MEMOIRE

Pour l'obtention du diplôme de :

Magister en Informatique

Option : Techniques Avancées pour les Systèmes Parallèles et Distribués

Thème

Systèmes multimodaux pour l'identification et l'authentification biométrique

Présenté par : Souheila Benkhaira

Soutenu publiquement le 01/10/2010

Devant le jury composé de :

Benmohammed Mohammed	Professeur	Université Constantine 2	Président
Tlili Yamina	Professeur	Université d'Annaba	Examinatrice
Mazouzi Smaine	MC Classe 'A'	Université de Skikda	Examinateur
Boucheham Bachir	MC Classe 'A'	Université de Skikda	Examinateur
Redjimi Mohammed	MC Classe 'A'	Université de Skikda	Encadreur

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

يَا أَيُّهَا الْإِنْسَانُ مَا غَرَّكَ بِرَبِّكَ الْكَرِيمِ (6) الَّذِي خَلَقَكَ فَسَوَّاكَ فَعَدَلَكَ (7) فِي أَيِّ صُورَةٍ مَا شَاءَ رَكَّبَكَ (8)

سورة الإنفطار

Je dédie ce mémoire à mes parents,

Remerciements

Mes louanges et gratitudes intarissable vont en premier à Dieu le Tout-Puissant qui m'a prodigué la santé, et le courage afin d'accomplir ce travail.

Mes plus vifs remerciements vont à mon encadreur Monsieur Redjimi Mohammed, maître de conférences à l'Université 20 Août 1955 de Skikda, qui avec son noble esprit scientifique, sa modestie, et sa patience m'a soutenu tout au long de la réalisation de ce travail.

Je tiens également à remercier tous les membres du jury qui m'ont fait l'honneur d'accepter de lire et de juger ce mémoire : Madame Tlili Yamina (Professeur, Université de Annaba), Messieurs Mazouzi Smaine (maître de conférences, Université de Skikda), Boucheham Bachir (maître de conférences, Université de Skikda) et Benmohammed Mohammed (Professeur, Constantine 2)

Je ne peux pas omettre d'exprimer ma gratitude envers ma famille et surtout mes parents, pour leur soutien indéfectible, leurs encouragements et leurs sacrifices, que ce mémoire leur soit dédié avec toute mon affection.

Résumé

La biométrie est la science qui consiste à établir l'identité d'une personne basée sur la reconnaissance de ses caractéristiques physiologiques et / ou comportementales. Les systèmes biométriques monomodaux utilisant une seule modalité présentent certaines limitations telles que la non-universalité, le manque d'individualité et la susceptibilité aux falsifications. Pour remédier aux ces problèmes tout en améliorant la performance de la reconnaissance, les informations issues de différentes sources biométriques sont combinées formant ainsi un système biométrique multimodal. Dans ce mémoire, nous nous sommes intéressés aux deux modalités les plus utilisées en termes de part de marché international de la biométrie à savoir le visage et l'empreinte digitale. Après avoir présenté l'état de l'art de chaque technologie, nous développons la méthode dite "Eigenfaces" pour la reconnaissance faciale et la méthode basée sur l'extraction des minuties des empreintes digitales, ensuite nous optons pour la fusion au niveau des scores afin d'établir un système multimodal qui permet d'atteindre de bons taux d'authentification.

Mots clés : Biométrie multimodale, visage, empreinte digitale, fusion de modalités, authentification, identification.

Abstract

Biometrics is the science that establishes the identity of a person based on his/her physiological and/or behavioral characteristics. Unimodal biometric systems that use only one modality suffer from several limitations such as non-universality, non-uniqueness and susceptibility to spoof attacks. To override these problems and enhance the performance of recognition application as well, informations from different biometric sources are combined to form multibiometric systems. In this thesis, we are interested in the two commonly used modalities in terms of international biometric market share, which are face and fingerprint. After describing a state of the art of each technology, we develop facial recognition method called "Eigenfaces", and the fingerprint method based on minutiae extraction, then we opt for scores fusion to establish a multimodal system, which gives good authentification rates.

Keywords: multimodal biometrics, face, fingerprint, modalities fusion, authentification, identification.

ملخص

القياس البيومتري (او القياس الحيوي) هو علم يهتم بتحديد هوية الأشخاص انطلاقا من سماتهم الخلقية و \ أو السلوكية. للنظام البيومتري المعتمد على سمة واحدة او ما يسمى "بالنظام البيومتري احادي الواسطة " بعض السلبيات من بينها امكانية تعرضه للتقليد كما ان بعض الافراد يشتركون في سمات معينة في حين ان آخرين يفتقدون بعضها. لتجاوز مثل هته العراقيل وتحسين مردودية القياس البيومتري, تدمج معلومات عدة انظمة بيومترية احادية الواسطة مشكلة ما يعرف " بالنظام متعدد الوسائط ". نهتم في هته المذكرة بالسمتين الاكثر استعمالا حسب مبيعات اسهم السوق الدولية للقياس البيومتري، وهما الوجه و بصمات الأصابع. نقوم اولا بعرض ما آلت اليه الابحاث في مجال القياس البيومتري, ثم نشرح مراحل التعرف على الوجه باستخدام تقنية ال "Eigenfaces", ومراحل تقنية استخراج الخواص الدقيقة لبصمات الأصابع, ثم ننشئ نظاما بيومتريا متعدد الوسائط انطلاقا من دمج نتائج التطابق للتقنيتين السالفة ذكرهما, ما يسمح بتحسين تحديد هوية الأشخاص.

كلمات جوهرية : القياس البيومتري متعدد الوسائط, الوجه, بصمات الأصابع, دمج الوسائط, اثبات الهوية, تحديد الهوية.

Table des matières

Introduction générale **1**

Chapitre 1 : Les systèmes biométriques **3**

1.1 Introduction 3
1.2 La structure d'un système biométrique 4
1.3 Les différentes modalités biométriques 6
1.4 Description des principales techniques biométriques 7
1.5 L'évaluation de performances des systèmes biométriques 9
1.5.1 L'évaluation d'authentification 10
1.5.2 L'évaluation d'identification 12
1.6 Le marché mondial de la biométrie 13
1.6.1 Chiffre d'affaire 13
1.6.2 Parts de marché 12
1.7 Quelle est la meilleure technique biométrique? 14
1.8 Pourquoi la multimodalité? 15
1.9 La multimodalité 17
1.9.1 Différentes formes de multimodalité 17
1.9.2 Architecture d'un système multimodal 18
1.10 Conclusion 20

Chapitre 2 : La reconnaissance faciale **21**

2.1 Introduction 21
2.2 Principales difficultés de la reconnaissance faciale 21
2.2.1 Changement d'illumination 22
2.2.2 Variation de pose 22
2.2.3 Expressions faciales 22
2.2.4 Occultations partielles 23
2.2.5 Autres difficultés 23
2.3 Le principe de fonctionnement d'un système de reconnaissance faciale 23
2.4 Détection de visage 25
2.5 La reconnaissance 2D de visage 25
2.5.1 Les méthodes globales 25
2.5.1.1 Les techniques linéaires 26
2.5.1.2 Les techniques non-linéaires 28
2.5.2 Les méthodes locales 29
2.5.2.1 Méthodes locales basées sur les caractéristiques d'intérêts 29
2.5.2.2 Méthodes locales basées sur l'apparence du visage 31
2.5.3 Les méthodes hybrides 33
2.6 L'Analyse en Composantes Principales (PCA) 34
2.7 Métriques de distances 38
2.8 Conclusion 39

Chapitre 3 : La reconnaissance des empreintes digitales **40**

3.1 Introduction 40
3.2 Caractéristiques d'une empreinte digitale 41
3.3 Structure d'un système de reconnaissance d'empreintes digitales 43
3.3.1 L'extraction des minuties 43
3.3.1.1 La méthode basée sur la binarisation 43
3.3.1.2 La méthode d'extraction directe 47
3.3.2 La comparaison d'empreintes digitales 49
a) Les approches basées sur la corrélation (correlation-based approches) 51
b) Les approches basées sur les minuties (minutiae-based approches) 51
c) Les approches basées sur les rides (ridge or texture-based approches) 52
3.3.2.1 La comparaison basée sur les minuties 52
3.4 Conclusion 54

Chapitre 4 : La fusion de modalités **55**

4.1 Introduction 55
4.2 Les niveaux de fusions 55
4.2.1 Fusion avant matching 55
4.2.2 Fusion après matching 56
4.3 La fusion au niveau des scores 57
4.4 La normalisation de scores 58
4.4.1 Les techniques de normalisation de scores 58
4.4.1.1 La normalisation Min-Max 59
4.4.1.2 La normalisation Decimal scaling 60
4.4.1.3 La normalisation Z-Score 60
4.4.1.4 La normalisation MAD (Median Absolute Deviation). 60
4.4.1.5 La normalisation QLQ 61
4.4.1.6 La normalisation double sigmoïde 61
4.3.1.7 La normalisation tanh (tangente hyperbolique) 62
4.5 Les méthodes de combinaison de scores. 64
4.6 Les méthodes de classification de scores 65
4.7 Conclusion… 66

Chapitre 5 : Résultats expérimentaux **67**

5.1 Introduction 67
5.2 La base de données multimodale 67
5.2.1 La base ORL 67
5.2.2 La base FVC 68
5.3 Interface graphique 68
5.4 Amélioration des images de test 72
5.4.1 Amélioration des images de visages 72
5.4.2 Amélioration des images des empreintes digitales 73
5.5 Evaluation du système biométrique. 77
5.6 Résultats et discussion 79
5.7 Conclusion 89

Conclusion générale **90**

Annexe A **91**

Annexe B.....**93**

Bibliographie **95**

Liste des figures

Chapitre 1

Figure 1.1. Enrôlement d'une personne dans un système biométrique [5] 5
Figure 1.2. Authentification d'une personne dans un système biométrique [5].....5
Figure 1.3. Identification d'une personne dans un système biométrique [5].....6
Figure 1.4. Les principales caractéristiques biométriques [3] 7
Figure 1.5. Exemples de différents capteurs biométriques.....9
Figure 1.6. Distribution des scores des personnes légitimes et des imposteurs.....11
Figure 1.7. La courbe DET.....11
Figure 1.8. La courbe CMC.....12
Figure 1.9. Evaluation du marché international de la biométrie [12].....13
Figure 1.10. Parts de marché des techniques biométrique [12] 14
Figure 1.11. Analyse Zephyr : comparaison de différentes modalités 15
Figure 1.12. Les différents systèmes multimodaux [20].....18
Figure 1.13. Architecture de fusion en série 19
Figure 1.14. Architecture de fusion en parallèle.....20

Chapitre 2

Figure 2.1. Exemple de variation d'éclairage.....22
Figure 2.2. Exemples de variation de poses 22
Figure 2.3. Exemples de variation d'expressions.....23
Figure 2.4. Variabilité intra-classe due à la présence d'occlusions partielles 23
Figure 2.5. Le principe de fonctionnement d'un système de reconnaissance faciale.....24
Figure 2.6. Phase d'apprentissage d'un système de reconnaissance faciale utilisant une méthode globale [40].....26
Figure 2.7. Illustration du principe de séparation optimale des classes par le LDA [45] 27
Figure 2.8. Comparaison entre les projections PCA et LDA de deux classes de points [49] 28
Figure 2.9. Localisation des caractéristiques géométriques utilisées dans [63] 30
Figure 2.10. Exemple de grille d'appariement [68] 31
Figure 2.11. Partitionnement de l'image visage en régions (ou patchs).....32
Figure 2.12. Processus de reconnaissance de visages basé sur l'MAA [65] 34
Figure 2.13. Passage d'une image vers un vecteur dans un espace vectoriel de grande dimension.....34
Figure 2.14. Exemple des cinq premiers Eigenfaces.....37
Figure 2.15. Illustration des cas possible de la projection d'une image sur l'espace des visages.....38

Chapitre 3

Figure 3.1. Vue en coupe de la peau au niveau du doigt.....41

Figure 3.2. Les principales classes d'empreintes digitales selon la classification de Galton-Henry [94] *42*
Figure 3.3. Les types de minutie *42*
Figure 3.4. Représentation des vecteurs de terminaison et de bifurcation [94] *43*
Figure 3.5. Représentation d'un système de reconnaissance d'empreinte digitale basée sur la binarisation *44*
Figure 3.6. Résultats des étapes de binarisation et de squelettisation *45*
Figure 3.7. Les cinq cas obtenus lors de processus d'extraction des minuties *45*
Figure 3.8. La détection des fausses minuties [98] *46*
Figure 3.9. Type des fausses minuties *47*
Figure 3.10. L'image directionnelle [103] *48*
Figure 3.11. Le suivi d'une strie [102] *49*
Figure 3.12. Les variations intra-classe d'une même empreinte digitale [94] *50*
Figure 3.13. La comparaison basée sur les minuties [108] *51*
Figure 3.14. L'extraction des informations de texture basées sur l'orientation locale de l'empreinte digitale [111] *52*
Figure 3.15. L'alignement de minuties de deux empreintes digitales [112] *53*

Chapitre 4

Figure 4.1. Les différents niveaux de fusion [20] *57*
Figure 4.2. La normalisation des scores par la technique Min-Max *59*
Figure 4.3. La Normalisation QLQ *61*
Figure 4.4. La Normalisation double sigmoïde *62*
Figure 4.5. Fonction d'influence de Hampel *63*

Chapitre 5

Figure 5.1. Interface graphique de la reconnaissance multimodale (visage et empreinte digitale) *71*
Figure 5.2. Égalisation de l'histogramme *73*
Figure 5.3. Processus d'amélioration de la qualité d'empreinte digitale [134] *74*
Figure 5.4. L'orientation de l'image sans et avec le lissage *76*
Figure 5.5. Amélioration de l'image de l'empreinte digitale *77*
Figure 5.6. La courbe FAR vs FRR *78*
Figure 5.7. La courbe FAR vs FRR et le taux EER *79*
Figure 5.8. La courbe DET du système de la reconnaissance faciale *80*
Figure 5.9. La courbe FAR vs FRR du système de la reconnaissance faciale *80*
Figure 5.10. La courbe DET du système de la reconnaissance d'empreintes digitales *81*
Figure 5.11. La courbe FAR vs FRR du système de la reconnaissance d'empreintes digitales *81*
Figure 5.12. Courbes DET pour la fusion MinMax-somme pondérée *82*
Figure 5.13. La courbe FAR vs FRR pour la fusion MinMax-somme pondérée *82*
Figure 5.14. Courbes DET pour la fusion Z-Score-somme pondérée *83*
Figure 5.15. La courbe FAR vs FRR pour la fusion Z-Score-somme pondérée *83*
Figure 5.16. Courbes DET pour la fusion MAD-somme pondérée *84*
Figure 5.17. La courbe FAR vs FRR pour la fusion MAD-somme pondérée *84*

Figure 5.18. Courbes DET pour la fusion QLQ-somme pondérée.....85
Figure 5.19. La courbe FAR vs FRR pour la fusion QLQ-somme pondérée.....85
Figure 5.20. Courbes DET pour la fusion Double Sigmoïde-somme pondérée 86
Figure 5.21. La courbe FAR vs FRR pour la fusion Double Sigmoïde-somme pondérée ...86
Figure 5.22. Comparaison des courbes ROC pour les différentes méthodes de normalisation.....88

Liste des tableaux

Tableau 1.1. Comparaison des modalités biométriques selon quelques propriétés [17]....16
Tableau 2.1. Comparaison des méthodes basées sur les caractéristiques locales et globales [80] 33
Tableau 2.2. Comparatif de quelques méthodes de reconnaissance de visage sur certaine base de données 39
Tableau 4.1. Résumé des techniques de normalisation de scores 64
Tableau 5.1. Les taux EER des différents systèmes d'authentification.....87
Tableau 5.2. Les temps d'exécution des différentes étapes du traitement 89

Introduction générale

De nos jours, la sécurité fait l'objet d'une attention particulière ; la nécessité de la protection civile et la lutte contre les fraudes d'une part et l'explosion de l'informatique et la croissance des moyens de communication d'autre part ont fait augmenter le besoin de s'assurer l'identité des individus. Les systèmes traditionnels d'authentification (mots de passe, codes PIN, badges, clefs, etc.) sont moins fiables à cause de leur inhabilité à différencier entre une personne autorisée et un imposteur (ils sont facilement falsifiables), problème qui provoque l'apparition des systèmes biométriques.

En fait la biométrie n'est pas vraiment récente ; l'utilisation de caractères anthropométriques comme le visage, la voix, et la démarche pour reconnaitre les individus est le moyen le plus naturel chez les êtres humains. L'utilisation des empreintes digitales remonte à la plus haute antiquité, les premières traces d'utilisation de cette caractéristique ont été découvertes en Egypte il y'a plus de 4000 ans. Les chinois ont utilisé très tôt ce trait à des fins de signature de documents. Depuis 1897, les services de police adoptent les empreintes digitales dans les enquêtes criminelles, cependant les procédures manuelles et la gestion des bases de données gigantesques rendent le travail long et couteux.

Aujourd'hui le développement des dispositifs informatiques favorise l'émergence des systèmes de reconnaissance automatiques de grande puissance, la biométrie n'est plus limitée aux empreintes digitales, beaucoup d'autres caractéristiques d'authentification (comme le visage, la voix, la signature, la rétine, l'iris etc.) sont utilisées. L'utilisation de la biométrie augmente de manière significative ; elle est demandée dans plusieurs domaines tels que le contrôle d'accès physique à des lieux protégés, et le contrôle d'accès virtuel à des réseaux d'ordinateurs.

Cependant les systèmes biométriques basées sur une seule modalité ont certaines limitations ; en fait ils ne peuvent pas garantir avec certitude une bonne identification, à titre d'exemple ; on ne peut pas trouver deux images strictement identiques du même visage, en outre les personnes handicapées sont privées de certaines caractéristiques biométriques. Pour réduire ces limitations la multimodalité est introduite.

L'identification par la biométrie multimodale consiste à combiner deux ou plusieurs systèmes unimodaux tout en améliorant la performance de la reconnaissance. Notre travail s'inscrit dans cette optique.

Dans ce mémoire ; nous utilisons deux principales technologies biométriques (en termes de part de marché international de la biométrie) à savoir l'empreinte digitale et la

reconnaissance faciale pour implémenter un système d'identification multimodal. Nous optons pour la fusion aux niveaux des scores pour combiner les informations issues de ces deux sources biométriques. Le manuscrit du mémoire est organisé autour de cinq chapitres, de la manière suivante :

- ✓ Dans le premier chapitre nous introduisons la biométrie, les mesures de performances couramment utilisées pour l'évaluation des systèmes biométriques, ainsi que l'état de l'art de la multimodalité.

- ✓ Dans le deuxième chapitre nous explorons les techniques de la reconnaissance faciale tout en détaillant la méthode de l'analyse en composantes principales.

- ✓ Le troisième chapitre concerne la reconnaissance d'empreintes digitales, deux méthodes d'extraction des minuties y sont présentées à savoir la méthode basée sur la binarisation et la méthode d'extraction directe.

- ✓ Dans le quatrième chapitre nous introduisons les différents niveaux de fusions en décrivant la fusion au niveau de scores.

- ✓ Le cinquième chapitre est dédié à l'évaluation des performances des systèmes de reconnaissance faciale et de reconnaissance d'empreintes digitales en fonction des méthodes mise en œuvre ainsi que les performances de système multimodal résultant de la fusion de deux modalités (visage et l'empreinte digitale).

- ✓ Nous terminons ce mémoire par une conclusion générale et les perspectives futures.

Chapitre 1

Les systèmes biométriques

1.1 Introduction

La biométrie se réfère à l'identification automatique d'un individu à partir d'une ou plusieurs caractéristiques physiologiques (visage, empreinte digitales, rétine, etc.), ou comportementales (démarche, écriture, etc.). Ces caractéristiques sont appelées « les modalités biométriques ». Etymologiquement, le mot « biométrie » vient de la concaténation de deux termes grecs « bio » qui signifie la vie, et « metron » qui se traduit par mesure [1], c'est ainsi John et al. [2] ont définit la biométrie comme « toutes caractéristiques physiques ou traits personnels automatiquement mesurable, robuste, et distinctives qui peuvent être utilisés pour identifier un individu ou pour vérifier l'identité prétendue d'un individu ». Donc la biométrie représente ce que l'on est contrairement à ce que l'on possède (carte, badge, clé, etc.), ou ce que l'on sait (mot de passe, code PIN, etc.) permettant de surmonter les problèmes liés à ces deux derniers systèmes à savoir : la duplication, le vol, l'oubli, et la perte.

Ainsi chaque caractéristique (physiologique et/ou comportementale) peut être qualifiée comme une modalité biométrique, si elle a les propriétés principales suivantes [3], [4]:

- ✓ *l'universalité*, signifie que chaque individu devrait posséder cette caractéristique.
- ✓ *l'unicité,* signifie que deux personnes doivent avoir des représentations différentes de leur trait biométrique.
- ✓ *la permanence,* ou *la stabilité,* c'est la constance d'une caractéristique au cours du temps.
- ✓ *la facilité de mesure,* se réfère à la facilité d'acquisition et de numérisation des données biométriques à l'aide d'un dispositif pertinent.

- ✓ *la performance,* se rapporte principalement à la précision de connaissance, à l'efficacité (vitesse d'exécution), et à la robustesse des ressources nécessaires pour atteindre la précision prévue.
- ✓ *l'acceptabilité,* signifie que les individus agréent de présenter leurs traits biométriques au système.
- ✓ *la non-reproductibilité*, reflète la difficulté avec laquelle le caractère d'un individu peut être falsifié par des méthodes frauduleuses.

1.2 La structure d'un système biométrique

Un système biométrique est essentiellement un système qui acquiert des données biométriques d'un individu, extrait un ensemble de caractéristiques à partir de ces données, puis les compare à un ensemble de données stockées au préalable pour pouvoir enfin exécuter une action ou prendre une décision à partir du résultat de cette comparaison [5]. Par conséquent, un système biométrique comprend quatre modules principaux:

- ***Le module de capture*** : Responsable de l'acquisition de certaines caractéristiques physiologiques, comportementales ou biologiques d'un individu (il peut être un appareil photo, un capteur d'empreintes digitales etc.). La qualité du capteur peut grandement influencer les performances de système.

- ***Le module d'extraction des caractéristiques*** : Extrait les traits fondamentaux et les caractéristiques des données acquises, en permettant d'obtenir une signature biométrique de l'individu, généralement sous forme d'un vecteur (appelé référence).

- ***Le module de comparaison (matching)*** : Compare l'ensemble des caractéristiques extraites avec le modèle préenregistré dans la base de données et détermine le degré de similitude.

- ***Le module de décision*** : Sert à prendre une décision sur le taux de correspondance de la signature pour la validation ou le rejet de l'identité de l'individu à connaitre.

Selon le contexte d'application, Le système biométrique peut fonctionner en trois modes à savoir l'enrôlement, l'authentification (ou la vérification) et l'identification

- ❖ ***Le mode d'enrôlement*** : (voir la figure 1.1) C'est la première phase de tout système biométrique, pendant laquelle les caractéristiques biométriques d'un individu sont enregistrées dans la base de données pour la première fois. Cet enregistrement est parfois accompagné de référence biographique correspond à cette personne comme le nom, le prénom, etc. qui sera utilisée plus tard dans la phase d'authentification. Pendant l'enrôlement on extrait des caractéristiques biométriques en utilisant des algorithmes adéquats, ces caractéristiques seront réduites par la suite pour minimiser la quantité de données à stockée en facilitant ainsi la vérification et l'identification.

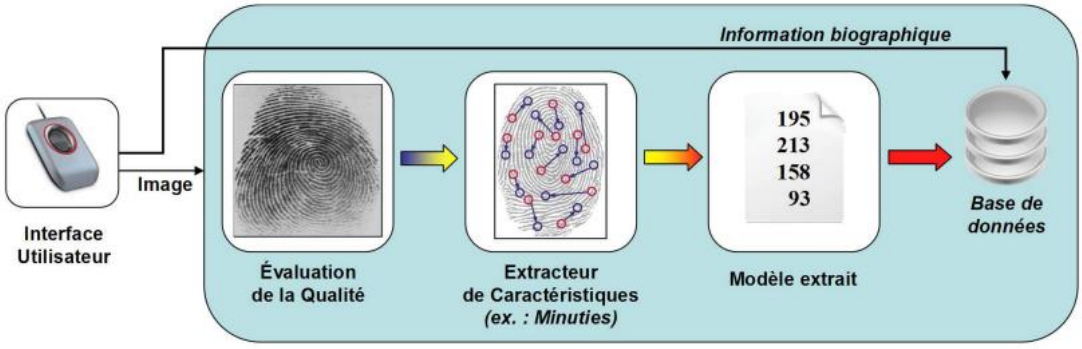


Figure 1.1. Enrôlement d'une personne dans un système biométrique [5].

- ***Le mode de vérification*** : (voir la figure 1.2) Le système vérifie l'authenticité d'une personne en effectuant une comparaison "un-à-un" (noté 1:1) de la signature biométrique capturée à ses propres modèles préenregistrés dans la base de données. Il répond alors à la question suivante : « ***Suis-je-bien la personne que je prétends être ?*** ». Dans cette phase le système effectue une mise à jour des signatures biométriques de modalité qui se changent légèrement à travers le temps (le visage, l'empreinte etc.).

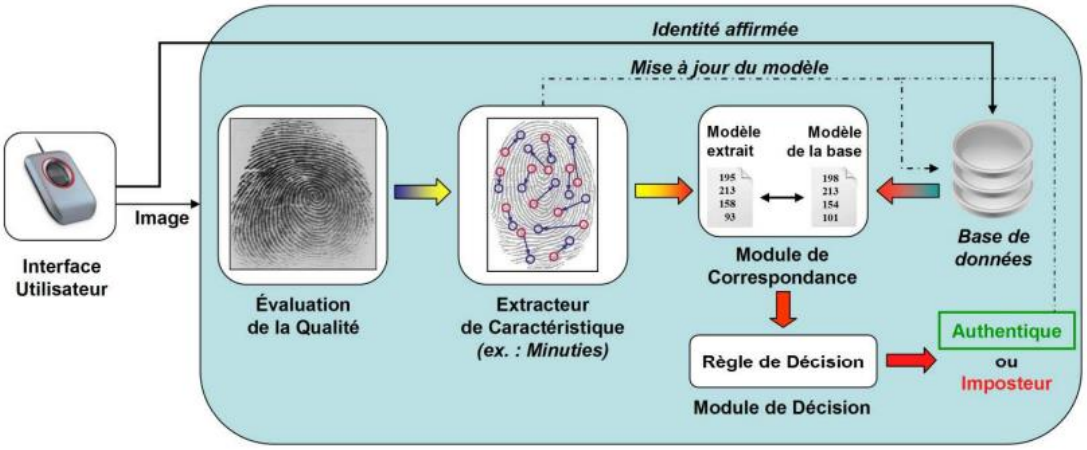


Figure 1.2. Authentification d'une personne dans un système biométrique [5].

- ***Le mode d'identification*** : (voir la figure 1.3) Le système identifie un individu par une recherche sur l'ensemble de modèles de toutes les personnes de la base de données en effectuant une comparaison "un-à-N" (noté 1:N). Ce mode consiste à associer une identité à une personne, le système répond donc à la question suivante : « ***Qui suis-je ?*** ».

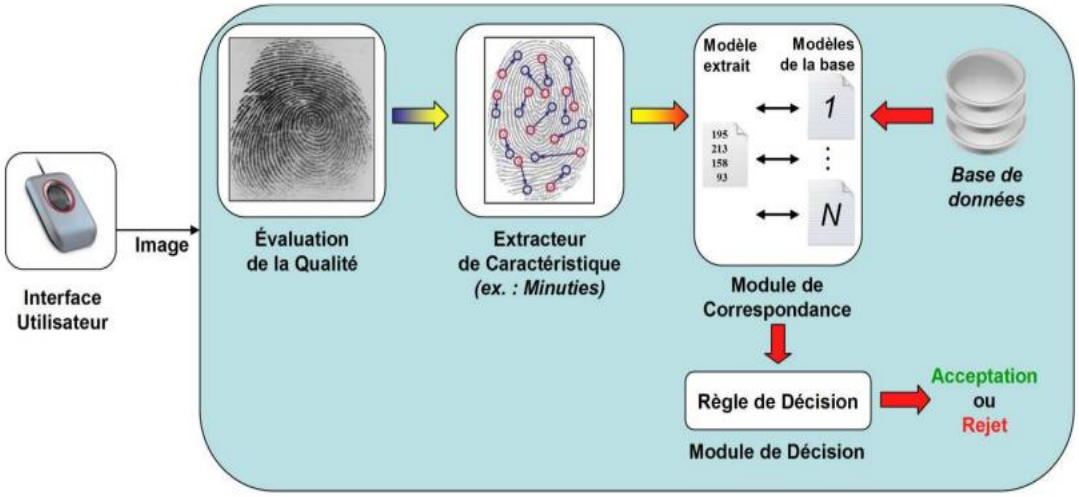


Figure 1.3. Identification d'une personne dans un système biométrique [5].

1.3 Les différentes modalités biométriques

Il existe un très grand nombre de modalités biométriques, qui peuvent se diviser en deux catégories :

- ✓ *L'analyse des traces biologiques* : utilise les caractéristiques biologiques des individus (ADN, salive, odeur etc.) qui sont très complexes à mettre en œuvre dans un système de reconnaissance.
- ✓ *L'analyse des traits physiques* : facile à mettre en œuvre, elle se représente en deux grandes classes [1] :
  - ➢ *La biométrie physiologique ou morphologique* : elle est basée sur l'identification des traits physiques particuliers, tel que la reconnaissance de la forme du visage, de la rétine, de l'empreinte digitale etc.
  - ➢ *La biométrie comportementale* : elle se base sur l'analyse de certains traits personnels du comportement de l'individu comme sa façon de taper sur un clavier, le tracé de sa signature, sa démarche, etc.

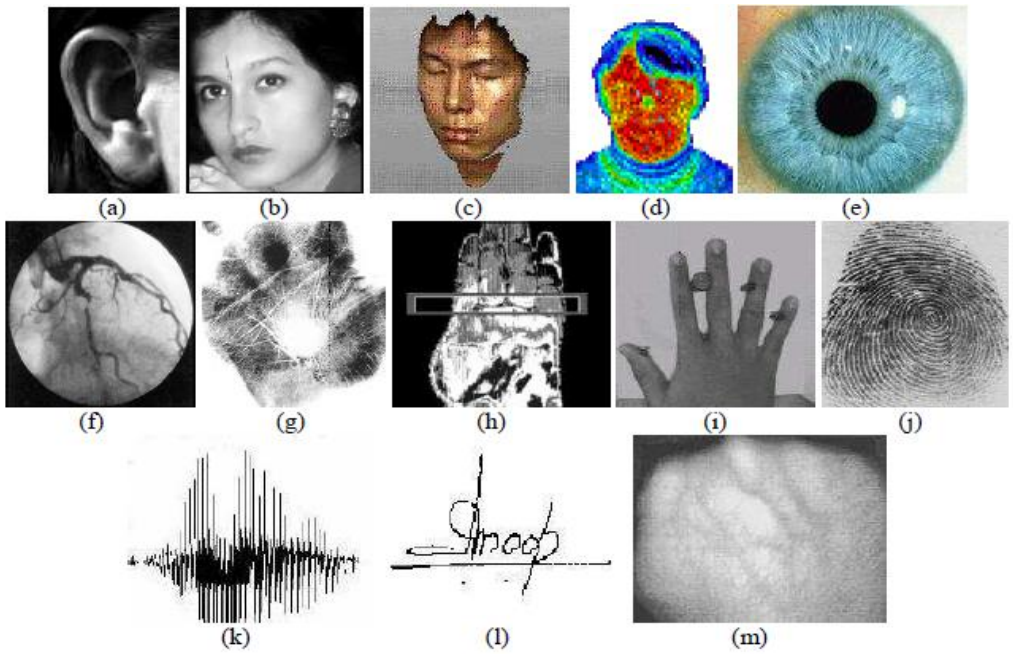


Figure 1.4. Les principales caractéristiques biométriques: a) forme de l'oreille, b) visage 2D, c) visage 3D, d) visage infrarouge, e) iris, f) rétine, g) empreinte de la main, h) thermogramme de la main, i) forme de la main, j) empreinte digitale, k) voix, l) signature et m) réseau veineux de la main [3].

1.4 Description des principales techniques biométriques

***L'empreinte digitale* :** l'empreinte digitale est le modèle de relief cutané des doigts. L'identification par cette caractéristique est la technique la plus anciennement utilisée. En fait, c'était toujours le choix biométrique évident pour les services de police depuis plus de 100 ans, c'est pour cela qu'elle est généralement mal acceptée par les utilisateurs en raison de l'alignement fort avec la criminologie. Il existe plusieurs types de système de capture d'empreinte digitale : optique, thermique, électromagnétique et ultrasons [6].

***Le visage* :** le visage est le moyen le plus naturel pour identifier les personnes, ce qui explique pourquoi cette caractéristique est bien acceptée par les utilisateurs. L'image de visage peut être captée par une caméra numérique, ou un appareil photo. Plusieurs recherches sont effectuées, pendant 25 ans [7], [8] pour améliorer la performance de ce genre de système cependant de nombreux problèmes se posent.

La géométrie de la main : cette modalité consiste à analyser la forme de la main sa longueur, sa largeur, son hauteur, la courbure des doigts etc. Cette technique est récente, simple, et bien acceptée par les utilisateurs qui suivent des guides des capteurs (LEDs infrarouge, des appareils photos numériques) pour qu'ils bien positionner leurs doigts, ce qui rende ainsi la détection / la segmentation plus aisée, cependant ce genre de système peut être trompé par de vrais jumeau ou même par des personnes ayant des formes de la main proches.

L'iris : l'iris est la région annulaire située entre la pupille et le blanc de l'œil. La biométrie par ce trait est la plus récente, et la plus fiable, selon les estimations de Daugmann[1] La probabilité de trouver 2 iris suffisamment identiques est 1 sur 10^{72} environ. L'image de l'iris est capturée par une caméra standard contraignantes (exemple la distance entre la camera et l'iris ne dépasse pas un mètre), ce qui limite l'utilisation de cette modalité.

La rétine : la rétine est la couche sensorielle de l'œil qui permet la vision, cette zone est parcouru par des vaisseaux sanguins dont leurs positions est inchangeable durant toute la vie de la personne. L'identification de la rétine n'est pas récente, elle remonte aux années 30. Cette technologie est la plus fiable toutefois elle est mal acceptée par les utilisateurs à cause des contraintes de l'acquisition.

La voix **:** la reconnaissance de la voix est une biométrie comportementale n'exige aucun contact physique avec le lecteur de système. En 1962 Lawrence Kersta [9] a prouvé que la voix de chaque personne est unique et qu'il est possible de la présenter graphiquement. Il existe deux principales méthodes de traitements de ce trait biométrique, la première dépend de texte prononcé, et la deuxième (la plus difficile) est indépendante de texte. Bien que cette modalité ne nécessite pas de matériel cher (microphone par exemple), cependant le bruit ambiant et les propriétés acoustiques telles que la réflexivité et l'absorption influencent la vérification de la voix en réduisant ainsi son utilisation.

La démarche **:** la démarche est l'une des biométries comportementale. Elle consiste à identifier les individus par leurs manières de marché, qui est supposée (presque) unique pour chaque individu. Deux types de techniques d'identification de ce trait peuvent être distingués, l'approche "*Model-based"* qui dépend de quelques paramètres comme la longueur des parties du corps, la longueur de pas, la jointure d'angle etc. et l'approche "*appearance-based"* qui analyse directement l'image en extrayant les caractéristiques. L'avantage de cette biométrie consiste dans le fait qu'on peut identifier la personne à distance, cependant l'exécution d'un tel système est particulièrement difficile.

La signature : la vérification par la signature est l'une des premières méthodes utilisées dans le domaine de la biométrie. Les systèmes de reconnaissance de l'écriture analysent soit la géométrie de la signature (mode statique), soit ses caractéristiques spécifiques comme la vitesse, la pression sur le crayon, ce mode qui s'appelle le mode dynamique est le plus discriminant. La capture se fait à l'aide d'une tablette graphique. bien que la signature soit bien acceptée par les utilisateurs, sa variabilité (à cause de l'état de santé ou l'état émotionnel de l'individu) pose un grand problème.

[1] http://www.cl.cam.ac.uk/users/jgd1000/

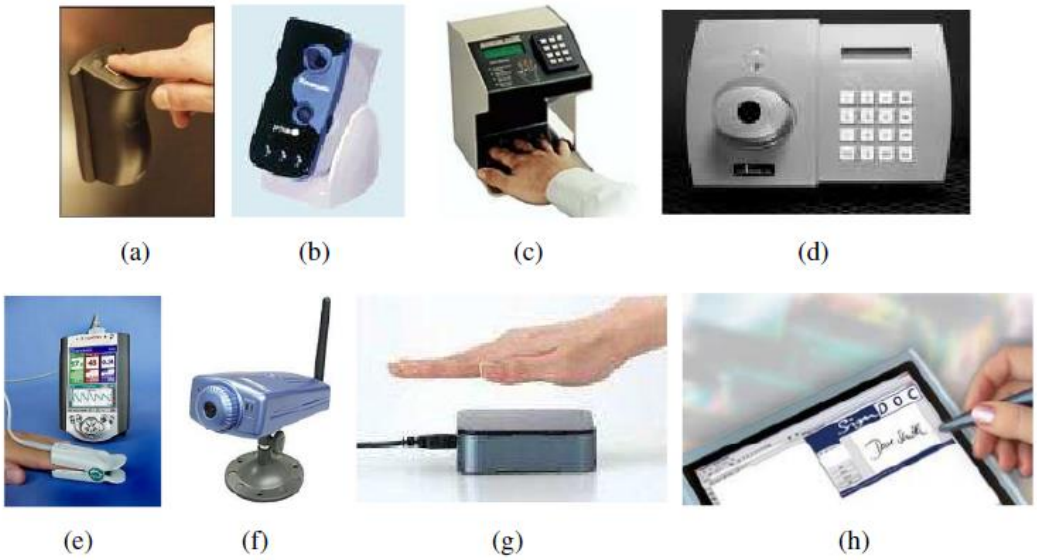


Figure 1.5. Exemples de différents capteurs biométriques concernant : (a) l'empreinte digitale, (b) l'iris, (c) la géométrie de la main, (d) la rétine, (e) la pulsation cardiaque, (f) le visage, (g) la thermographie de la main, et (h) la signature.

1.5 L'évaluation de performances des systèmes biométriques

L'évaluation de performance comprend de nombreux aspects, tels que la précision, l'efficacité, le volume de données stockées pour chaque individu, l'échec à l'acquisition ou à l'enrôlement, le coût des matériels et des logiciels, la facilité d'usage pour les utilisateurs, etc. Philips et al. [10] en distinguent trois types ; à savoir : *l'évaluation technologique* qui est liée à la partie algorithmique, *l'évaluation de scénario* y compris le test des capteurs et de l'environnement, et *l'évaluation opérationnelle.*

Nous nous intéressons dans ce qui suit à la précision, qui est considérée comme étant un point important pour estimer la performance d'un système biométrique. Tout d'abord, nous introduisons les trois critères principaux suivants :

- **FAR** *(False Acceptation Rate),* ou **F.M.R** *(False Match Rate),* représente le pourcentage des personnes déclarées par le système comme clients alors qu'ils sont des imposteurs. Il est égale au nombre de fausse acceptation divisé par le nombre total d'accès imposteurs.

- **FRR** *(False Reject Rate),* ou **F.N.M.R** *(False Non Match Rate),* représente le pourcentage des personnes censées être clients mais ils sont rejetés par le système. C'est le ratio entre le nombre de faux rejets et le nombre de tests des personnes légitimes.

- **EER** *(Equal Error Rate),* c'est le taux d'erreurs égales, un compromis entre les fausses acceptations et les faux rejets, autrement dit c'est le point de mesure sur lequel FAR = FRR.

Selon la nature de système biométrique que nous l'avons vu précédemment (authentification ou identification), deux façons de mesures de précision peuvent être distinguées :

1.5.1 L'évaluation d'authentification

L'authentification consiste à apparier le modèle biométrique stocké d'une personne (I) et les données acquises par le système biométrique, plus formellement :
Soient X_Q le vecteur de caractéristique (Feature vector), de la personne proclamée I, et X_I celui de la personne I stockée dans la base de données, le test de vérification est définit ainsi [11] :

$$(I, X_Q) = \begin{cases} w_1, & si\ S(X_Q, X_I) \geq t \\ w_2, & sinon \end{cases} \quad (1.1)$$

Où S est la fonction de similarité entre le vecteur X_Q et X_I, w_1 indique que la personne proclamée est un client et w_2 désigne qu'elle est imposteur, et t c'est le seuil.

La figure 1.6 illustre l'estimation des taux FAR et FRR suivant les distributions des imposteurs et des personnes légitimes données par les équations (1) et (2) ci-dessous [3]:

$$FAR = \int_0^t p\big(S(X_Q, X_I)\big|H_0\big)ds \quad (1.2)$$

$$FRR = \int_t^1 p\big(S(X_Q, X_I)\big|H_1\big)ds \quad (1.3)$$

Légende :

H_0 : Hypothèse indique que la signature X_Q et X_I, ne proviennent pas de la même personne.
H_1 : Hypothèse indique que la signature X_Q et X_I, proviennent de la même personne.
D_0 : la décision associée à H_0 qui indique que la personne n'est pas celle qu'elle prétend être.
D_1 : la décision associée à H_1 qui indique que la personne est celle qu'elle prétend être.
$P(s|H_0)$: la distribution des scores des personnes imposteurs.
$P(s|H_1)$: la distribution des scores des personnes légitimes.

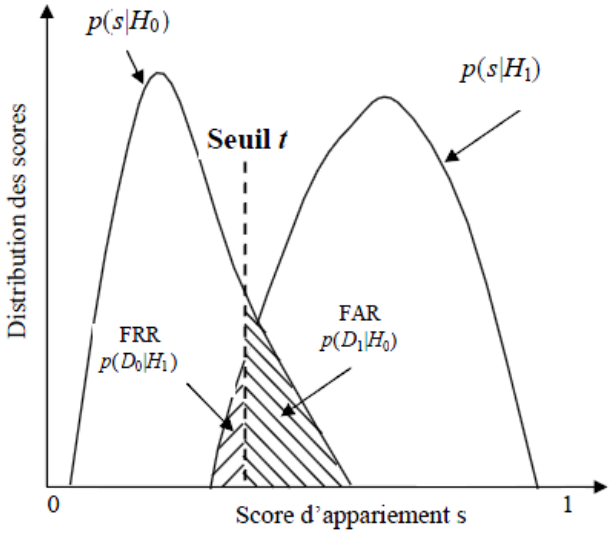


Figure 1.6. Distribution des scores des personnes légitimes et des imposteurs.

FAR et FRR sont influencés par un seuil de décision. Donc le choix de ce point de fonctionnement est important, il dépend essentiellement du type d'application et des performances souhaitées. En effet plus le seuil de décision est bas, plus le système acceptera de client, mais plus il acceptera aussi d'imposteur, par contre, plus ce seuil est grand plus le système rejettera d'imposteur mais plus il rejettera aussi de client.

L'utilisation de la courbe DET (Detection Error trade-off) est très fréquente pour l'évaluation de tels systèmes d'authentification, elle représente la variation de taux FRR en fonction du taux de FAR. L'allure de cette courbe est illustrée par la figure 1.7.

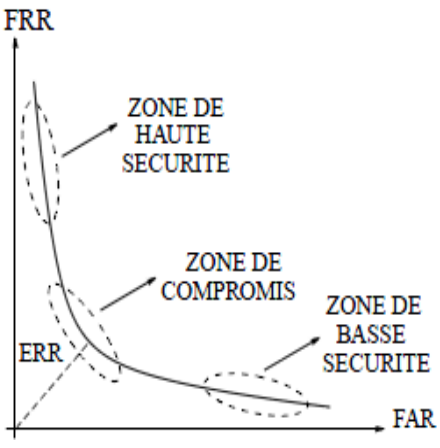


Figure 1.7. La courbe DET.

1.5.2 L'évaluation d'identification

L'identification peut se formuler par [11] :
Soit X_Q le vecteur de caractéristique d'entré, l'identification revient à déterminer l'identité I_k, $k \in \{1, 2,..,N, N+1\}$,où I_1, I_2,..,I_N sont les identités des personnes préalablement enrôlées dans le système, et I_{N+1} indique une identité rejetée. La fonction d'identification est définie ainsi:

$$(I, X_Q) = \begin{cases} I_k, & \text{si } \max\{S(X_Q, X_{I_k})\} \geq t, k = 1, \dots, N. \\ I_{N+1}, & \text{sinon} \end{cases} \quad (1.4)$$

Où X_{I_k} est le vecteur de caractéristique correspond à l'identité I_k, et le t est le seuil fixé.

Le teste d'identification se fait souvent en utilisant la courbe CMC (Cumulative Match Characteristics) qui représente le taux d'identification de système en fonction d'une variable que l'on appelle le *rang*, c'est-à-dire le taux d'erreur est mesuré sur k décision avec k variant de 1 à N, où N est le nombre d'identités enregistré dans la base de référence. On dit qu'un système reconnaît au rang 1 lorsqu'il choisit le plus proche individu (selon le module de similarité) correspond à l'identité proclamée. On dit qu'un système reconnaît au rang 2 lorsqu'il choisit parmi deux images celle qui correspond le mieux à l'image d'entrée, et ainsi de suite. La courbe CMC est une courbe croissante autrement dit, plus le degré du rang augmente, plus le taux de reconnaissance correspondant est lié à un niveau de sécurité faible.

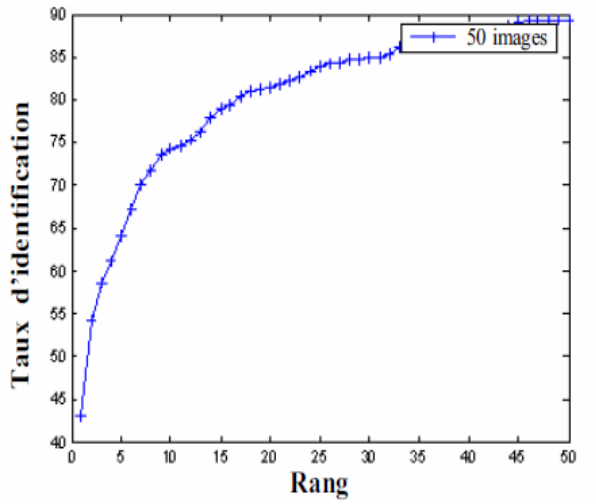


Figure 1.8. La courbe CMC.

1.6 Le marché mondial de la biométrie

1.6.1 Chiffre d'affaire

Le marché de la biométrie ne cesse d'évoluer depuis son apparition, d'après le rapport publié par IBG (*International Biometric Group*), le chiffre d'affaire de l'industrie biométrique (incluant les applications judiciaires et celles du secteur public) est en forte croissance jusqu'en 2014 (voir la figure 1.9). Une partie importante de cette progression s'attache au contrôle d'accès aux systèmes d'information (ordinateur / réseau) et au commerce électronique, quoique les applications du secteur public continuent à être une partie essentielle de l'industrie.

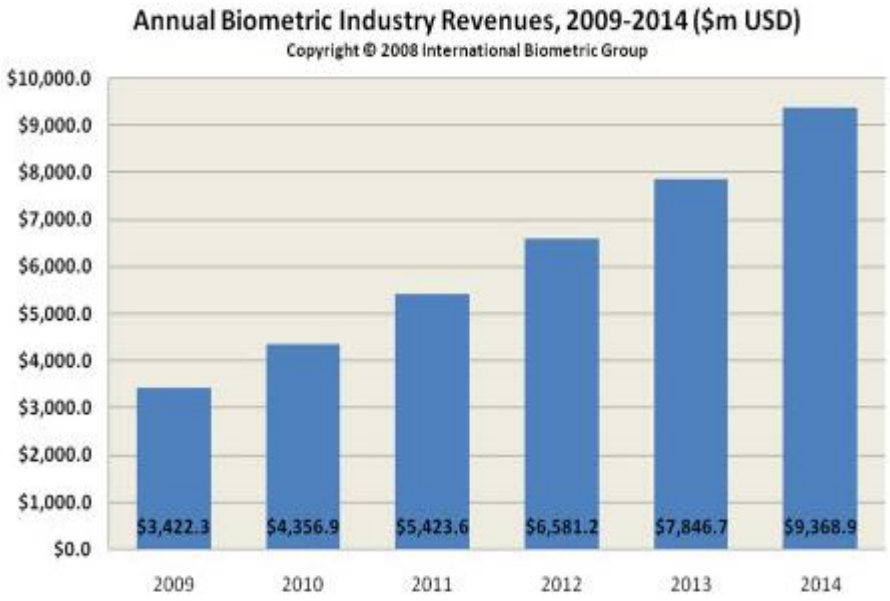


Figure 1.9. Evaluation du marché international de la biométrie [12].

1.6.2 Parts de marché

La figure 1.10 réalisée d'après les chiffres d'affaire de l'International Biometric Group montre les parts de marché des principales méthodes biométriques. On constate que l'empreinte digitale occupe la première place, plus de la moitié du marché mondial (hors application judiciaire). La reconnaissance de visage vient en deuxième position, près de 12% de ce même marché (hors application judiciaire), dépasse ainsi la reconnaissance de la main (qui avait avant la deuxième place après les empreintes digitales).

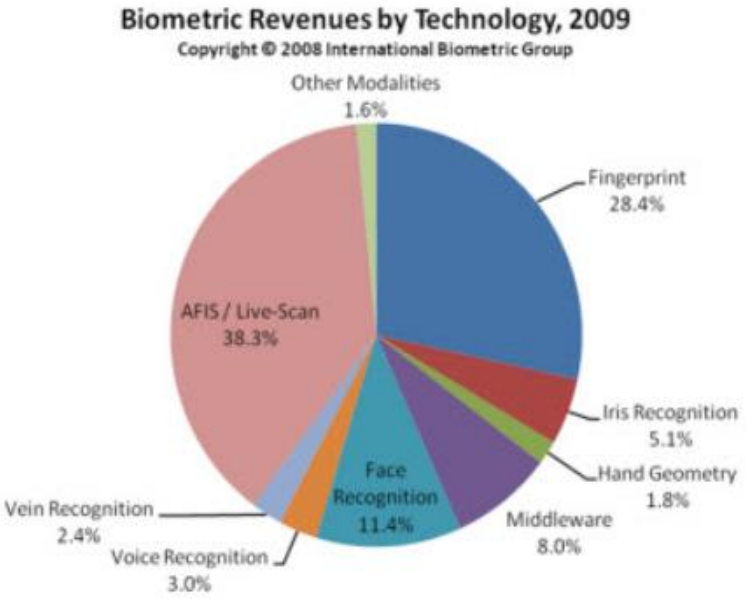


Figure 1.10. Parts de marché des techniques biométrique [12].

1.7 Quelle est la meilleure technique biométrique ?

La comparaison des différentes biométries est généralement effectuée en fonction de quatre critères à savoir *l'effort, l'intrusion, le coût,* et *la précision.* (Il y'a d'autres critères de comparaison des techniques biométriques (cf. Table1.1)).

- ✓ *l'effort (Effort)* : l'effort fourni par l'utilisateur lors de la vérification par le système biométrique.
- ✓ *l'intrusion (Intrusiveness)* : niveau de l'acceptation du test par l'utilisateur.
- ✓ *le coût (Cost)* : coût de la technologie (capteurs, lecteurs...).
- ✓ *la précision* (*Accuracy)* : efficacités de la méthode.

L'analyse Zephyr (figure 1.11) montre qu'il n'existe pas une méthode biométrique idéale, en effet chaque technique a ses forces et ses faiblesses. Le choix dépend essentiellement de la nature de l'application par exemple la voix et la signature sont des méthodes qui n'exigent pas un grand effort de l'utilisateur, sont peu intrusives, de coût modéré, cependant ils ne sont pas assez performantes. L'iris et la rétine sont fiables toutefois elles sont coûteuses et mal acceptées par le grand public. A noter que le choix de la modalité biométrique dépend aussi de la culture locale des utilisateurs, en Asie les méthodes nécessitant un contact physique comme l'empreinte digitale, sont rejetées pour des raisons d'hygiène alors que les méthodes qui n'exigent pas un contact sont bien acceptées.

AFIS: Automated Fingerptint Identification System.

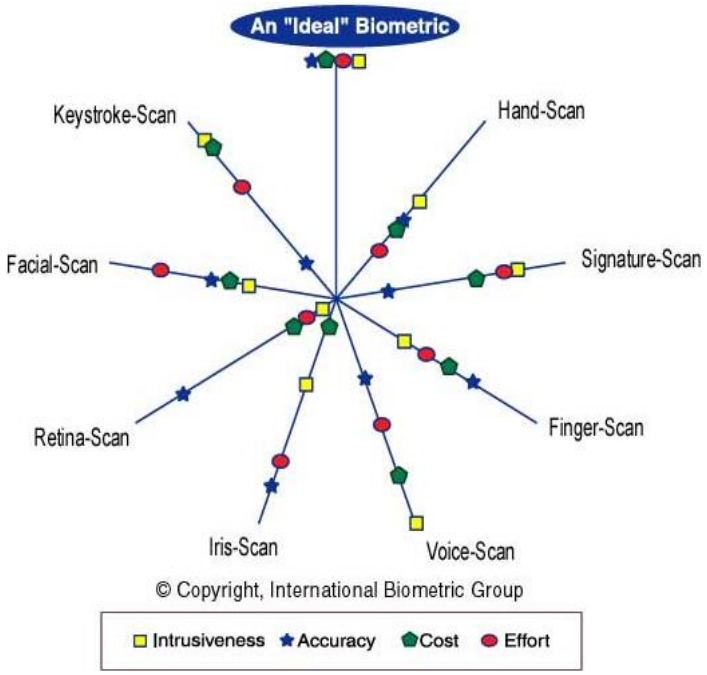


Figure 1.11. Analyse Zephyr : comparaison de différentes modalités selon quatre critères principaux : l'intrusivité, la précision, le coût et l'effort.

1.8 Pourquoi la multimodalité ?

Bien qu'il existe un grand nombre des techniques de reconnaissance biométrique unimodales (utilisant une seule modalité) améliorées, cependant celles-ci souffrent de plusieurs facteurs qui dégradent leurs performances [13] tels que :

- *Le bruit* : introduit par le capteur défaillant ou mal entretenu (exemple l'accumulation de poussière sur le capteur d'empreinte digitale), il peut sérieusement compromettre la précision du système [14].

- *Non-universalité* : en fait certaines modalités ne sont pas vraiment universelles par exemple certaines personnes peuvent avoir des empreintes non enregistrables à cause d'un accident ou d'un travail manuel prolongé. Selon l'estimation de NIST (*National Institute of Standards and Technologies)* il est impossible d'obtenir une bonne qualité d'empreinte digitale pour environ 2% de la population. C'est le cas d'une personne muette avec la modalité voix, ou une personne handicapée avec la modalité signature, ou les personnes qui ont une maladie oculaire ou ayant des long cils pour la modalité iris ou rétine. Ces problèmes entraînent des erreurs d'enrôlement "*Failure to Enroll*" et/ou des erreurs de capture "*Failure to Capture*".

- *Manque d'individualité* : nous pouvons avoir une similarité des caractéristiques extraites de deux personnes différentes cela dû à des facteurs génétiques (exemple les jumeaux, les membres de la même famille, ou même les membres de la même population), ce qui affecte significativement la performance de système biométrique en augmentant le taux de fausse acceptation.

- *Manque de représentation invariante* : les caractéristiques biométriques ne sont pas invariantes ni dans leur enrôlement ni dans leur nature, c'est ainsi nous pouvons avoir deux signatures différentes correspond à la même personne. Par exemple on ne peut jamais trouver deux images du même visage strictement identiques et cela est dû au changement de pose, des expressions faciales, des conditions d'éclairage, et même de l'utilisation des capteurs différents. Ces variations à l'intérieur d'une même modalité biométrique est appelées "variation intra-classe", elles augmentent le taux de faux rejet de système biométrique.

- *Sensibilité aux attaques* : le problème du fraude et du vol semble s'éliminer par les méthodes biométriques, cependant on ne fait que les réduire. Des études [15], [16] ont montré qu'il est possible de voler une empreinte digitale et de la produire (en utilisant par exemple la silicone), c'est le cas des modalités signature et voix qui sont faciles à falsifier.

La table ci-dessous presente une comparaison de quelques modalités biométriques selon les propriétés suivantes : universalité, unicité, permanence, facilité d'enregistrement, acceptabilité et performance [17].

Information	U	N	P	C	A	E
ADN	Oui	Oui	Oui	Faible	Faible	*****
Sang	Oui	Non	Oui	Faible	Non	*
Démarche	Oui	Non	Faible	Oui	Oui	***
Dynamique de frappe	Oui	Oui	Faible	Oui	Oui	****
Voix	Oui	Oui	Faible	Oui	Oui	****
Iris	Oui	Oui	Oui	Oui	Faible	*****
Rétine	Oui	Oui	Oui	Oui	Faible	*****
Visage	Oui	Non	Faible	Oui	Oui	****
Géométrie de la main	Oui	Non	Oui	Oui	Oui	****
Oreille	Oui	Oui	Oui	Oui	Oui	*****
Empreinte digitale	Oui	Oui	Oui	Oui	Moyenne	****

Table 1.1. Comparaison des modalités biométriques selon les propriétés suivantes : (U) universalité, (N) Unicité, (P) Permanence, (C) Collectabilité, (A) Acceptabilité et (E) Performance. Pour la performance, le nombre d'étoiles est relié à la valeur du taux d'égale erreur (EER) obtenue dans l'état de l'art [17].

1.9 La multimodalité

Afin d'améliorer la performance des systèmes biométriques susmentionnées, (dits système unimodaux) la multimodalité est introduite. Cette dernière consiste à combiner plusieurs techniques biométriques en réduisant ainsi le risque d'impossibilité d'enrôlement, et obtenant un système robuste aux fraudes.

1.9.1 Différentes formes de multimodalité

La multimodalité au sens large se réfère à cinq scenarios différents (figure 1.12) qui sont :

1. ***Systèmes multi-instances*** : il s'agit d'utiliser un seul capteur pour extraire des instances du même caractère biométrique, afin d'obtenir plusieurs variations de ce trait en enrichissant le modèle biométrique de l'individu. Par exemple l'acquisition de plusieurs images de visage en changeant la pose, l'expression, et/ou l'illumination.

2. ***Systèmes multi-capteurs*** : dans ce système nous utilisons plusieurs capteurs pour acquérir la même modalité, afin d'extraire plusieurs informations de même trait biométrique. Exemple la capture de la texture 2D, de la surface 3D, ainsi que l'image infrarouge de visage de l'individu avec différents gamme de capteurs.

3. ***Systèmes multi-algorithmes*** : dans ce genre de système plusieurs algorithmes sont utilisés, dans la phase d'extraction de caractéristiques et/ou dans la phase de la mise en comparaison pour traiter la même donnée. Exemple l'utilisation des algorithmes pour analyser la texture et les minuties de l'empreinte digitale afin d'extraire des caractéristiques pouvant améliorer la performance du système [18].

4. ***Systèmes multi-échantillons*** : ce type de système associent plusieurs échantillons de la même biométrie. C'est le cas par exemple de l'iris gauche et droit, ou deux empreintes digitales de doits différents. Ce genre de système ne nécessite ni plusieurs algorithmes, ni plusieurs capteurs, cependant il exige plusieurs références contrairement au système multi-instance qui n'utilise qu'une seule.

5. ***Systèmes multi-biométries*** : (ou système multimodale au sens strict) ici on combine différents traits biométriques d'un individu par exemple le visage et l'iris, le visage et l'empreinte digitale etc. Il faut noter ici que l'utilisation des biométries décorrélés (comme la rétine et la démarche) peut donner un système plus performant que celui obtenu en fusionnant deux biométries corrélés (comme la voix et le mouvement des lèvres) [19].

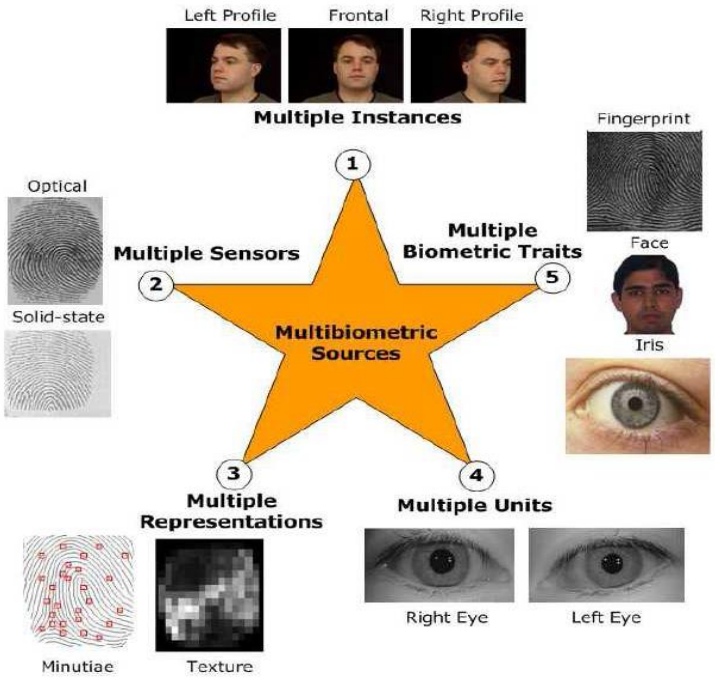


Figure 1.12. Les différents systèmes multimodaux [20].

Une combinaison de plusieurs scenarios ce qu'on appelle les ***systèmes hybrides*** [21] améliore évidemment la performance des systèmes unimodaux, par exemple l'utilisation des images de 2D et de 3D du visage et l'empreinte digitale d'un individu.

A noter, que les quatre premiers systèmes se basent sur une seule modalité, permettent de résoudre le problème de la variabilité intra-classe. Cependant ils ne peuvent pas pallier ni le problème de la non-universalité, ni celui de la fraude contrairement aux systèmes multi-biométries.

1.9.2 Architecture d'un système multimodal

Le système multimodal se réfère à la combinaison de deux ou plusieurs systèmes biométriques en acquérant les informations qui seront traitées par la suite. Ces deux procédures peuvent se faire simultanément (*architecture parallèle*) ou en série (*architecture série*). Cependant, l'architecture de système biométrique dépend essentiellement de la manière du traitement des données, car l'acquisition est souvent séquentielle. En fait il n'existe pas des capteurs capables de recevoir simultanément des traits biométriques de différentes modalités, à l'exception de quelques capteurs d'empreintes multi-doigts ou palmaires. Et même l'architecture de traitement est liée à la *décision*. Il s'agit, donc, de prendre la décision après le traitement de toutes les modalités biométriques, on parle alors de *la fusion en parallèle*. Ou bien décider après avoir un score de similarité de chaque modalité

ce qu'on appelle la *fusion en série*. Cette dernière est utilisée dans certaines applications, par exemple en cas de manque de trait biométrique (les personnes handicapées) on peut obtenir des caractéristiques issues d'une autre modalité. Néanmoins l'architecture en parallèle est la plus utilisée car elle est performante dans la mesure où elle utilise un grand nombre des informations disponibles, toutefois elle est coûteuse en temps et en matériel.

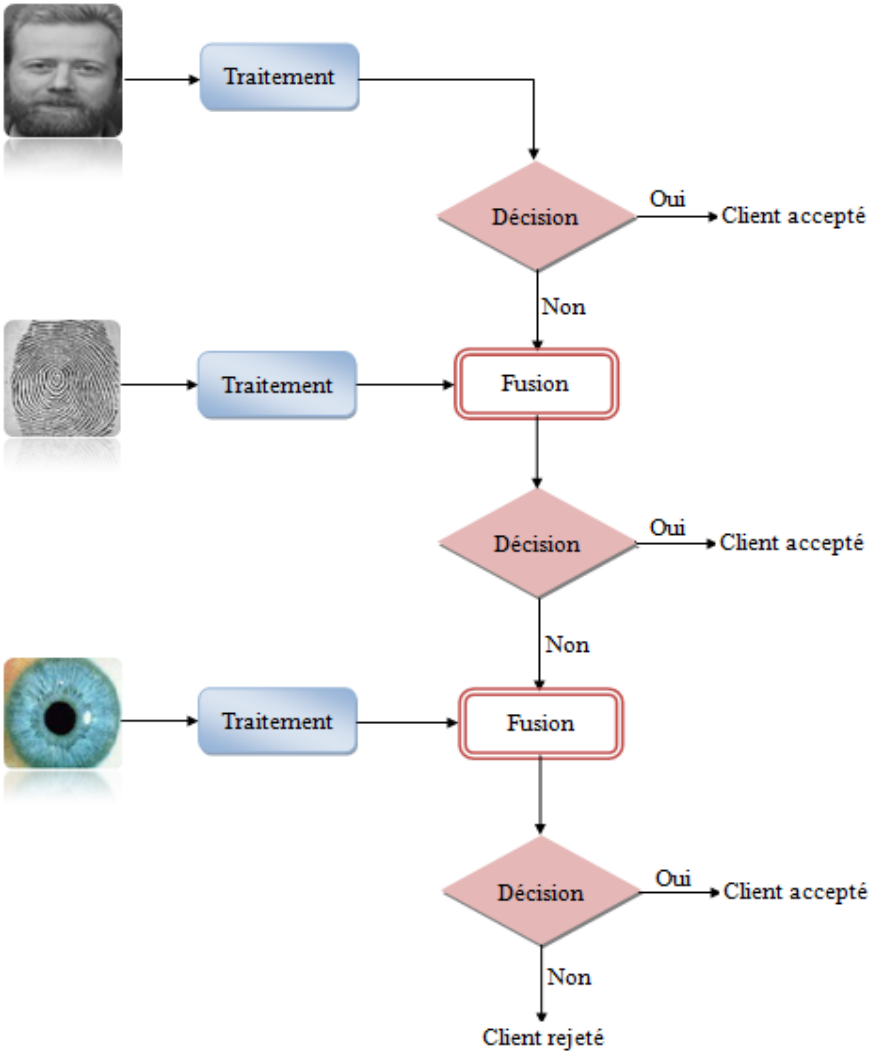


Figure 1.13. Architecture de fusion en série.

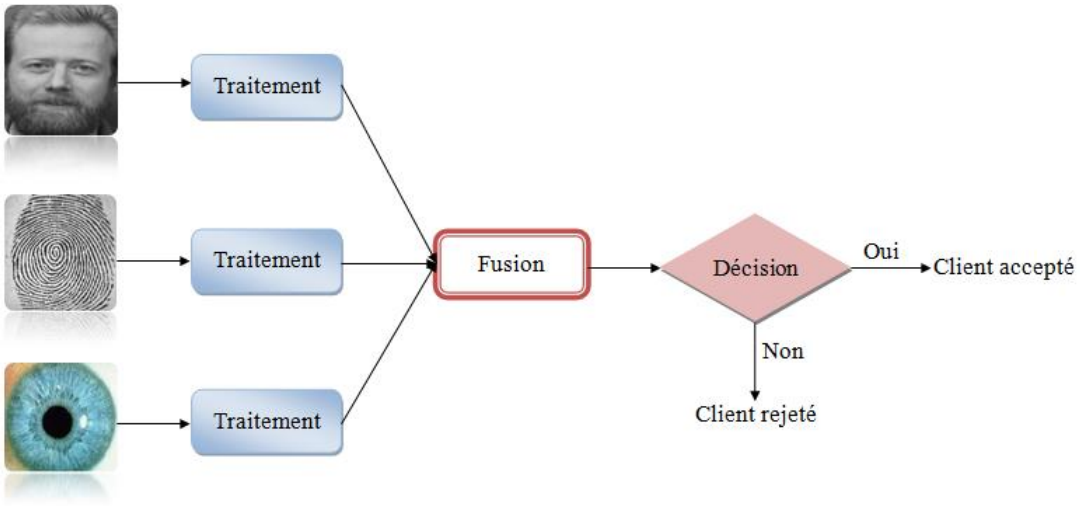


Figure 1.14. Architecture de fusion en parallèle.

1.10 Conclusion

Dans ce premier chapitre nous avons abordé l'état de l'art de la biométrie, qui fournit une alternative aux systèmes traditionnels. Nous avons déterminé ses propriétés, ses principes, et ses éléments. Nous avons définit la structure des systèmes biométriques, les critères utilisés pour les évaluer, et les limitations de chaque modalité, avant d'introduire la multimodalité (le cadre de ce mémoire), approche qui vise à réduire les inconvénients des systèmes unimodaux. Nous avons présenté ses formes, ainsi que les architectures qui peuvent être utilisées. Nous allons étudier la fusion multimodale en concentrant sur la fusion au niveau de scores dans le chapitre 4.

Chapitre 2

La reconnaissance faciale

2.1 Introduction

La reconnaissance faciale est la technique la plus acceptable parmi les techniques de reconnaissance biométriques, elle reçoit une attention accrue dans le domaine de recherche du fait de ses caractères. En fait ; elle est *naturelle, non-intrusive,* et *facile à utiliser*. Le développement des technologies telles que les appareils photos numériques, les ordinateurs équipés de webcam et les dispositifs mobiles a contribué son déploiement à large échelle. Elle est utilisée surtout dans les applications de sécurité tels que : la télésurveillance, le contrôle d'accès à des sites, l'accès à des bâtiments etc.

De nombreuses méthodes de reconnaissance faciale ont été proposées depuis plus de trois décennies, cependant les systèmes de reconnaissance du visage restent complexes et offrent un grand challenge pour les chercheurs. L'amélioration de taux de reconnaissance, la diminution de taux de fausse acceptation, et l'accélération de temps de recherche dans les grandes bases de données sont des défis auxquels faire face les algorithmes proposés.

2.2 Principales difficultés de la reconnaissance faciale

Les êtres humains peuvent détecter et identifier facilement les visages des gens familiers à partir d'images de mauvaise qualité ou de faible résolution, ce qui n'est pas le cas par un système automatique de reconnaissance du visage. Le principal problème est dû à la variation d'acquisition des images, notamment la variance *intra-classe,* c.à.d. la variabilité que peut prendre le visage d'une même personne à cause des facteurs que nous détaillons ci-dessous.

2.2.1 Changement d'illumination

L'intensité et la direction d'éclairage influent énormément sur l'aspect d'un visage dans l'image acquise (voir la figure 2.1). Les changements d'éclairage peuvent affecter tout le visage, ou entrainer la création d'ombres et de zones éclairées, masquant certaines caractéristiques faciales. En fait le changement de l'image du visage dû à l'illumination se révèle parfois plus critique que la variation inter-classe (la variabilité que prennent les visages de différentes personnes), ceci est expérimentalement constaté par Adini et al [22] où les auteurs ont utilisé une base de données de 25 personnes. Ainsi ; l'extraction de caractéristiques faciales invariantes dues aux changements de luminosité dans un environnement non contrôlé reste un domaine de recherche ouvert.

Figure 2.1. Exemple de variation d'éclairage.

2.2.2 Variation de pose

C'est la variation de rotation qu'a pris le visage lors de sa capture (frontal, 45 degrés, profil), elle est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. En effet ; des tests d'évaluations élaborés sur la base FRVT2000 [23] ont démontré que la rotation de la tête n'entraine pas de baisse significatives des taux de reconnaissance jusqu'à ±25°. Cependant lorsque cette rotation dépasse ce seuil elle engendre une chute de performances. La figure 2.2 présente l'exemple d'un visage d'une même personne subissant des variations de pose.

Figure 2.2. Exemples de variation de poses.

2.2.3 Expressions faciales

Le facteur expression faciale de l'émotion combinée avec la parole affecte l'apparence du visage (figure 2.3) en entrainant une diminution significative du taux de reconnaissance, surtout celui des systèmes fondés sur des points d'intérêt. En fait ; l'information faciale, notamment celle qui se situe dans la partie inférieure comme la bouche, affecte la forme géométrique et la position des caractéristiques faciales. L'utilisation de plusieurs modèles de

visages, un par catégorie d'expression faciale [24], [25] peut être une résolution de ce problème en comparant ensuite le visage-test à la base des visages arborant la même expression.

Figure 2.3. Exemples de variation d'expressions.

2.2.4 Occultations partielles

Le visage peut être partiellement occulté par d'autres objets dans la scène (par exemple occulté par une autre personne), ou par le port des accessoires tels que les lunettes de vue ou de soleil, un chapeau, les cheveux longs, etc. Gross et al. [26] montrent que les performances des algorithmes testant les visages des individus portant de lunette de soleil et de cache nez sont relativement faibles par rapport à ceux testant des visages entièrement découverts. La figure 2.4 présente la variabilité intra-classe due à la présence d'occlusions partielles.

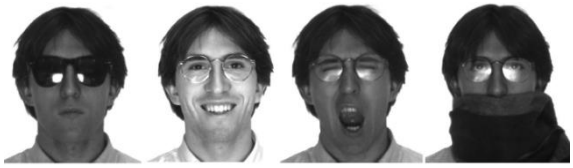


Figure 2.4. Variabilité intra-classe due à la présence d'occlusions partielles.

2.2.5 Autres difficultés

D'autres difficultés peuvent influencer le système automatique de reconnaissance de visage. La présence des composants structurels tels que la barbe, la moustache, la présence de maquillage, d'opérations chirurgicales etc. peuvent modifier énormément les caractéristiques faciales. Nous ajoutons un autre point très important qui est l'âge de capture (moment à laquelle les captures ont été réalisées). Gross et al. [26] utilisent la base AR pour déterminer l'impact de ces facteurs où le délai entre deux sessions de prises de vue est deux semaines, ils ont obtenu une baisse des taux de reconnaissance estimée à 20%.

2.3 Le principe de fonctionnement d'un système de reconnaissance faciale

La reconnaissance automatique de visages se décompose principalement en deux modules (la figure 2.5) :

- ✓ *Le module de détection/normalisation* : il se charge de détecter et/ou localiser le visage dans une image ou une vidéo en effectuant une éventuelle normalisation pour ramener le visage à une taille standard.

- ✓ *Le module de reconnaissance* : ce module comporte trois phases à savoir, le *prétraitement* qui vise à segmenter l'image du visage en éliminant les artefacts avant de passer à la seconde étape qui est *l'extraction de caractéristiques*, la dernière étape consiste à *comparer* l'empreinte biométrique testée à l'ensemble des empreintes biométriques stockées dans la base de données (qui s'appelle la galerie), pour prendre enfin une décision sur l'appartenance d'un individu à l'ensemble des visages ou pas.

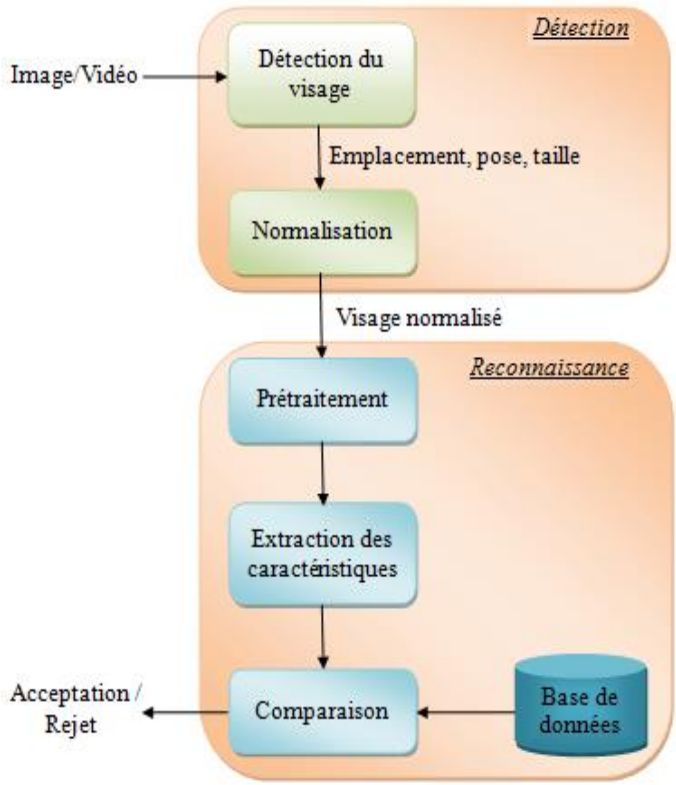


Figure 2.5. Le principe de fonctionnement d'un système de reconnaissance faciale.

Notons que la détection du visage et l'extraction de caractéristiques peuvent être exécutées simultanément. Dans le reste de ce chapitre nous nous intéressons au module de la "reconnaissance faciale" en décrivant quelques techniques les plus utilisées. Nous présentons la méthode *Eigenfaces* étant donné que nous allons l'utiliser dans notre travail. D'abord nous parlons brièvement de l'étape de détection (voir l'état de Yang et al. [27] pour plus de détails) pour que le discours soit complet et cohérent.

2.4 Détection de visage

La détection de visage dans l'image brute d'entrée est l'étape la plus complexe du processus de reconnaissance faciale. Cette étape comporte deux sous étapes : *la localisation* et *la normalisation*. Pour détecter les visages, de nombreuses méthodes peuvent être appliquées Yang et al. [27] les divisent en quatre catégories.

- ✓ ***Approches basées sur les reconnaissances acquises (top-down)*** : ces approches se basent sur les caractéristiques du visage (nez, bouche et yeux) en définissant des règles qui décrivent les relations entre ces caractéristiques. Exemple de techniques : le calcul des distances entre les yeux, le nez et la bouche [28], l'application des projections horizontales et verticales de l'image [29].

- ✓ ***Approches basées sur prototypes (Template-matching)*** : ces approches consistent à comparer tout ou partie de l'image candidate avec des modèles de visages enregistrés dans la base de connaissances. Ces modèles peuvent être définis manuellement ou à l'aide des fonctions, Yuille et al. [30] par exemple ont utilisé un prototype déformable pour la modélisation des caractéristiques faciales.

- ✓ ***Approches basées sur l'apparence*** : il s'agit des techniques d'apprentissage automatique qui créent des modèles à partir d'un ensemble représentatif d'image. Le problème de détection de visage est considéré comme un problème de classification sépare l'image en *visage* et *non-visage*. Parmi les techniques proposées dans cette catégorie on trouve celle de réseaux de neurones artificiels [31], [32], [33].

- ✓ ***Approches basées sur les traits invariants (bottom-up)*** : ces approches visent à trouver les traits structurels invariants d'un visage même si les conditions de prise d'image (la pose, la vue, la lumière) varient. Les caractéristiques les plus utilisées sont la forme du visage, la texture et la couleur de peau [34].

2.5 La reconnaissance 2D de visage

Il existe de nombreuses méthodes de reconnaissance de visage que nous pouvons classer en trois groupes : les ***méthodes locales***, les ***méthodes globales*** et les ***méthodes hybrides***.

2.5.1 Les méthodes globales

Ces méthodes de reconnaissances prennent le visage dans son entier comme entrée du système et utilisent des techniques d'analyses statistiques bien connues. Le principe de cette approche est de représenter l'image de visage de dimension (m, n) par un seul vecteur de

dimension m×n obtenu par la concaténation des valeurs du niveau de gris des pixels de l'image faciale, puis on projette l'image d'entrée dans un espace de plus faible dimension, en sélectionnant les caractéristiques nécessaires et discriminantes. L'avantage des méthodes globales est qu'elles conservent implicitement les informations de texture et de forme utiles pour la reconnaissance de visage, ainsi qu'elles permettent une meilleure capture de l'aspect global du visage que les présentations locales [35], de plus elles sont rapides à mettre en œuvre. Cependant leurs inconvénients résident dans leurs sensibilités aux conditions de la luminosité, de pose et d'expression faciale, de même, la taille des données à traiter est importante [36], [37], [38], par exemple une image 100×100 est représentée par un vecteur de dimension 10^4 [39], donc l'utilisation des technique de réduction de la dimension est indispensable.

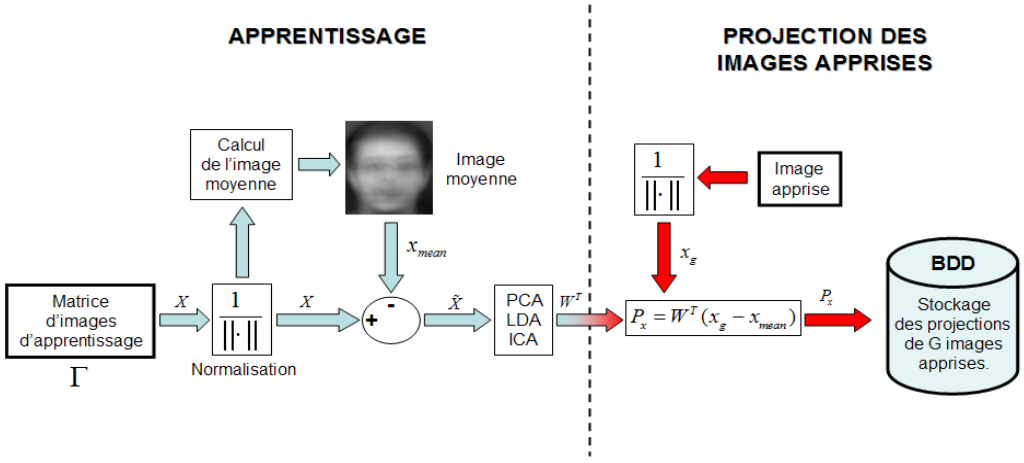


Figure 2.6. Phase d'apprentissage d'un système de reconnaissance faciale utilisant une méthode globale [40].

Les méthodes globales peuvent se décomposer en deux types de techniques à savoir : les *techniques linéaires* et les *techniques non linéaires*.

2.5.1.1 Les techniques linéaires

Ce sont des techniques qui projettent linéairement les données des visages d'espace de grande dimension sur un espace de plus faible dimension. Cependant elles sont sensibles au luminosité et surtout aux variations non convexe , et même l'utilisation des distances classiques dans l'espace projeté comme la distance Euclidienne et plus généralement les distances de Mahalanobis ne permettent pas de réaliser une bonne classification entre les classes "visages" et " non-visages" et entre les individus eux-mêmes.

La technique linéaire la plus connue est Eigenfaces [41], qui est basée sur *l'Analyse en Composantes Principales* (PCA pour *"Principal Component Analysis"*). Son principe consiste à chercher les directions de l'espace qui représentent le mieux les corrélations entre les variables aléatoires. Il s'agit d'une transformation orthogonale linéaire des données dans un nouvel espace représenté par des vecteurs propres associés aux plus grandes valeurs propres, les images des visages seront ensuite projetées sur cet espace.
Cette technique est à la base de nombreuses méthodes globales qui apportent des améliorations ou des variations. Par exemple pour les travaux réalisés sur le choix des vecteurs propres nous pouvons citer :

- ✓ Les travaux de Kirby et al. [42] qui proposent de retenir les vecteurs propres associés aux plus grandes valeurs propres jusqu'à ce que la somme de ces dernières dépasse 90% de leur énergie totale.
- ✓ Martinez et al. dans [43] obtiennent un meilleur taux de reconnaissance en ignorant les premiers vecteurs propres (encodant souvent les variations d'illumination).

Une autre méthode très connue présentée par Belhumeur et al. [44] basée sur *l'Analyse Discriminante Linéaire* (LDA pour *"Linear Discriminant Analysis"*), appelée *Fisherfaces* qui divise les visages en classes selon le critère de Fisher. Cette technique est une variante de PCA, elle vise à trouver les directions de projection les plus discriminantes dans l'espace propre en maximisant le ratio entre la variance inter-classe et la variance intra-classe. Cependant lorsque les variations intra-personnelles sont petites (c.-à-d. quand il n'y a pas beaucoup d'image par individu) la matrice de variance intra-personnelles peut être singulière et son inversion pose donc un problème qui est connu sous le nom de *small sample size problem.* Une des méthodes la plus utilisée pour le contourner consiste à utiliser d'abord le PCA pour diminuer la dimension des échantillons et ensuite nous réalisons le LDA. L'annexe A comporte une description plus complète de l'analyse discriminante linéaire.

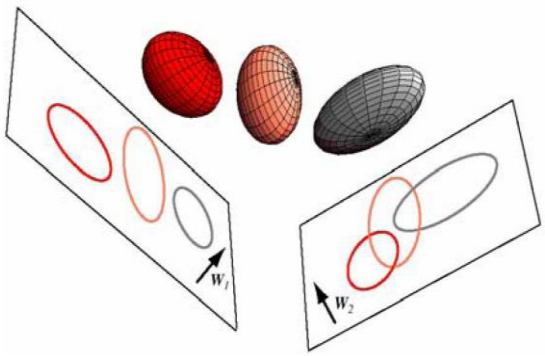


Figure 2.7. Illustration du principe de séparation optimale des classes par le LDA. Trois distributions 3D sont projetées sur deux sous-espaces 2D décris par les vecteurs W_1 et W_2. Puisque le LDA essaye de trouver la plus grande séparation parmi les classes, on voit bien que W_1 est ici le vecteur optimal [45].

Donc chaque visage se compose d'un grand nombre de pixels sera réduit à un plus petit ensemble de combinaisons linéaires avant la classification. Les combinaisons linéaires obtenues en utilisant PCA construisent un sous-espace pour représenter de manière "optimale" seulement "l'objet visage" tandis que LDA cherche à réaliser un sous-espace discriminant pour distinguer de façon "optimale" les visages de différentes personnes. Cependant quand il n'y a pas beaucoup d'images par personne cette méthode marche moins bien que celle basée sur l'PCA [46]. Une comparaison de ces méthodes est effectuée par Socolinsky et Selinger dans [47], et par Wu et al. dans [48].

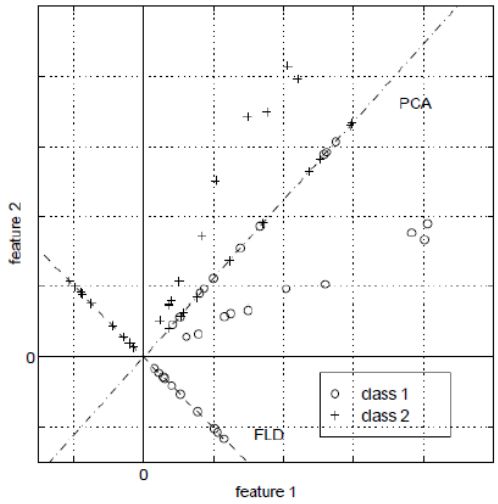


Figure 2.8. Comparaison entre les projections PCA et LDA de deux classes de points [49].

D'autres techniques linéaires ont également été proposées nous citons par exemple :

- L'analyse en composantes indépendantes (ICA) [50].
- La factorisation de matrices non négatives (NMF) [51], [52].
- L'analyse discriminante bilinéaire (BDA) [53].
- La technique de "Vecteurs communs discriminants" (DCV) [54].

Bien que les méthodes globales linéaires soient efficaces, et aient eu beaucoup de succès, elles ne sont pas assez précises. Ceci est dû à des transformations non-linéaires, une simple modification de la luminosité déforme les images de visages de façon non linéaire.

2.5.1.2 Les techniques non-linéaires

Afin de pouvoir traiter le problème de la non-linéarité, des techniques globales non-linéaires ont été développées. Parmi ces méthodes on trouve :

Kernel–PCA (*l'Analyse en Composantes Principales à Noyaux*) [55], et Kernel–LDA (*l'Analyse Discriminante Linéaire à Noyaux*) [56], qui se basent sur la notion mathématique "*noyau*" en étendant les techniques PCA et LDA.

La technique de MDS (*Multi Dimensional Scaling*) [57], [58] vise à préserver autant que possible les propriétés globales des données d'apprentissage dans le nouvel espace de faible dimension en réalisant un mapping. La qualité de ce dernier est exprimée à travers une fonction de stress (ex. la fonction de stress brut, la fonction de coût de Sammon), qui est une mesure de l'erreur des distances entre les pairs des échantillons dans les espaces de faible et de grande dimension.

Il existe d'autres techniques comme l'Isomap [59], les diffusion maps [60], les approches neuronales [61], [62] etc.

Bien que les méthodes non-linéaires peuvent atteindre une amélioration de taux de reconnaissance sur les données d'entraînement en effectuant une projection non-linéaire de l'espace des images sur l'espace de caractéristiques *"feature space"* permettant une réduction de dimension de meilleure façon, elles ne peuvent pas être robustes pour de nouvelles données en raison de leur grande flexibilité contrairement aux méthodes linéaires.

2.5.2 Les méthodes locales

Les méthodes locales reposent sur la reconnaissance *à priori* que l'on possède sur les objets caractéristiques pour pouvoir les localiser. Ces méthodes peuvent être classées en deux catégories, la première concerne les méthodes *basées sur les caractéristiques locales*, il s'agit des méthodes d'extraction et de localisation des points d'intérêts, la deuxième catégorie est celle des méthodes basées sur *l'apparence faciale* qui divisent l'image de visage en petites régions (ou patchs) de caractéristiques locales.

2.5.2.1 Méthodes locales basées sur les caractéristiques d'intérêts

Ces méthodes se basent sur l'extraction des caractéristiques géométriques du visage telles que les distances entre les yeux, la largeur de la tête, etc. Cependant ces approches présentent deux inconvénients à savoir :

- La difficulté d'extraire les caractéristiques géométriques dans certains cas complexes (exemple occultations, la variation de pose, d'illumination, etc.).
- Les caractéristiques géométriques seules ne sont pas suffisantes pour présenter entièrement le visage, tandis que d'autres informations utiles telles que les niveaux de gris sont écartées.

Pour remédier à ces deux limites, deux directions de recherche sont produites. La première se base sur la performance des détecteurs de points caractéristiques du visage. C'est ainsi que Brunelli et Poggio [63] proposent une technique qui extraie automatiquement un ensemble de 35 caractéristiques géométriques d'une image de visage (figure 2.9). Ils comparent ensuite ces ensembles de caractéristiques deux à deux via la distance de Mahalanobis. Ils ont obtenu un taux de reconnaissance estimé à 90% sur une base de données de 47 sujets.

Rowley et al. [64] utilisent plusieurs détecteurs de traits correspondant à chaque partie du visage. Lanitis et al. [65] proposent de construire des modèles statistiques de la forme du visage. Malgré ces travaux, les chercheurs n'arrivent pas à trouver un détecteur de points caractéristiques suffisamment fiable.

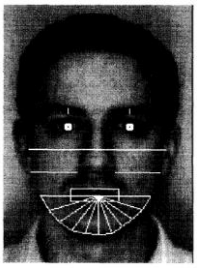


Figure 2.9. Localisation des caractéristiques géométriques utilisées dans [63].

Plutôt que d'utiliser des méthodes purement géométriques, la deuxième direction de recherche s'appuie sur les méthodes basées sur des représentations fournissant des informations portées par les points caractéristiques du visage. Ainsi Manjunath et al. [66] ont proposé une méthode pour détecter et extraire les caractéristiques locales du visage en utilisant les ondelettes de Gabor [67]. Un graphe topographique est construit pour modéliser la relation entre les points caractéristiques. Ils ont obtenu un taux de reconnaissance estimé à 90% sur un ensemble de données de visages de 86 sujets, toutefois le graphe topographique ne peut pas être modifié par la suite pourtant l'image de visage se change en différentes variations (illumination, expression, pose, etc.).

C'est ainsi que Lades et al. [68] ont proposé un graphe topologique élastique déformable qui varie en fonction des variations d'apparence du visage connu sous le nom de *"Elastic Graph Matching"* au lieu d'un graphe topologique fixe. Ce graphe est une grille rectangulaire, placée sur l'image de visage (figure 2.10), où les nœuds sont des points *labélisés* auxquels on associe un jeu de coefficients d'ondelettes complexes de Gabor, appelés *Jets*. La comparaison entre deux graphes de visage s'effectue en mettant en correspondance le graphe de l'image de visage à reconnaitre et ceux des visages de base de données.

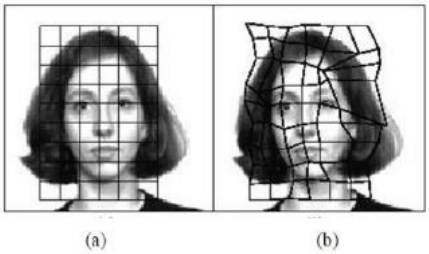


Figure 2.10. Exemple de grille d'appariement. (a) grille de référence, (b) grille correspondante [68].

Wiskott et al. [69] ont étendu la méthode d'*Elastic Graph Matching* à une méthode très connue appellée *"Elastic Bunch Graph Matching"* (EBGM), où chaque nœud comporte un ensemble de 40 coefficients complexes d'ondelette de Gabor, incluant la phase et l'amplitude, codant la variation des niveaux de gris, toutefois cette technique nécessite un temps de calcul important.

Beaucoup de travaux apportent des améliorations à cette méthode, nous citons par exemple la méthode proposée par Kepenekci et al. [70]. Les chercheurs utilisent, ici, un ensemble de matrices de filtres de Gabor pour parcourir les régions faciales locales, et au lieu de fixer le nombre de points caractéristiques du visage comme dans la méthode EBGM, ils choisissent seulement ceux obtenus avec la réponse fréquentielle la plus haute du filtre de Gabor pour représenter le visage, ce qui minimise le temps de calcul.

Les méthodes basées sur l'extraction des caractéristiques d'intérêts sont efficaces dans le cas où une seule image d'apprentissage par personne est disponible. Cependant leurs performances dépendent de la précision des algorithmes de localisation des points caractéristiques. En pratique cette tâche reste difficile notamment dans le cas où la forme et l'apparence du visage peuvent grandement changer.

2.5.2.2 Méthodes locales basées sur l'apparence du visage

Ces techniques sont appliquées de manière modulaire sur les différentes régions de visage. En fait les méthodes basées sur l'apparence faciale comportent généralement quatre étapes : *le découpage en régions de la zone du visage, l'extraction des caractéristiques locales, la sélection des caractéristiques* et *la classification.*

- ✓ ***Le découpage en régions* :** cette étape consiste à localiser les régions d'intérêt à l'intérieur de visage. Deux paramètres définissent ces régions : *la forme* et *la taille*. La forme peut être rectangulaire, elliptique, etc. (voir la figure 2.11), mais ce qui est le plus largement utilisé c'est le découpage rectangulaire. La taille des régions influence le nombre des caractéristiques donc elle est cruciale pour les performances de la méthode choisie.

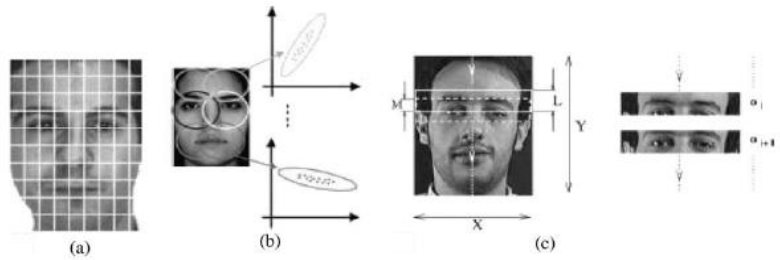


Figure 2.11. Partitionnement de l'image visage en régions (ou patchs).

- ✓ ***L'extraction des caractéristiques locales* :** dans cette étape nous déterminons les caractéristiques de chaque région en utilisant quelques techniques comme, les Ondelettes de Gabor [70], [71], les Ondelettes de Harr [72], les transformées de Fourier, les caractéristiques basées sur les indices LBP (*Local Binary Pattern*) [73], SIFT (*Scale Invariant Feature Transform*) [74] ou l'analyse des valeurs de niveau de gris [75], [76].

- ✓ ***La sélection des caractéristiques* :** une fois les caractéristiques locales ont été extraites, nous en sélectionnons les plus pertinentes. Cette étape permet d'accélérer le traitement, le PCA, le LDA, et l'Adaboost [77] sont des méthodes couramment utilisées pour la sélection des caractéristiques les plus discriminantes.

- ✓ ***La classification* :** dans cette étape, nous identifions le visage en utilisant par exemple la stratégie de fusion par vote à la majorité ou par somme pondérée (*Matcher Weighting* [78]) de l'ensemble des scores de chaque classificateur.

Notons que les quatre étapes ne sont pas obligatoires pour toutes les méthodes. Par exemple la sélection des caractéristiques peut être éliminée ou combinée avec d'autres étapes.

Martinez et al [79] présentent une approche probabiliste locale pour la reconnaissance de visages occultés partiellement et avec des variations d'expression, En utilisant une seule image d'apprentissage par classe, ils divisent le visage en six régions elliptiques. Les tests effectués sur un ensemble de 2600 images montrent que l'approche probabiliste locale ne réduit pas la précision même lorsque 1/6 du visage est occulté. Toutefois les coûts de calcul et de stockage et la procédure de génération des échantillons virtuels sont très élevés (6615 exemples par individu dans [79]).

Bien que les méthodes locales de reconnaissance de visage soient efficaces dans le cas d'une seule image d'apprentissage, de plus elles peuvent modéliser les variations d'acquisition, plusieurs problèmes restent non résolus. En fait, les méthodes locales sont

robustes vis-à-vis d'une certaine variation ; par exemple la méthode EBGM est efficace pour modéliser le changement d'expression, d'illumination et de pose, mais elle n'est pas robuste vis-à-vis d'occultation, alors que la méthode probabiliste locale est robuste aux variations d'expression et aux occultations mais pas aux changements de pose. De même, les méthodes locales nécessitent souvent le placement manuel des points caractéristiques, ce qui les rend lourdes à mettre en œuvre.

Nous avons abordé ci-dessus deux catégories concernant les méthodes de reconnaissance faciale ; les méthodes globales et les méthodes locales. Nous avons vu que les méthodes globales présentent certains avantages, elles traitent le problème de reconnaissance faciale comme un problème d'analyse de sous-espace de visages, pour lequel plusieurs méthodes statiques peuvent être utilisées, ainsi que les méthodes ne nécessitant pas des images de bonne qualité. Cependant ces méthodes n'ont pas d'a priori sur le physique des visages, et ne donnent pas de bons résultats qu'avec des variations limitées d'illumination et d'expression, elles sont trop sensibles aux variations de pose et d'occultation. Tandis que les méthodes locales sont robustes dans certaines fortes variations (voir la table 2.1). De même, la connaissance a priori sur le visage intégrée au modèle de ces méthodes améliorent leur efficacité, toutefois la construction de ces modèles est une tâche difficile et exige des images de bonne qualité et/ou de résolution suffisante. Nous pouvons exploiter les avantages des approches locales et globales en évitant leurs inconvénients par l'utilisation d'une autre catégorie des méthodes, il s'agit des méthodes hybrides.

Facteurs de variations	Caractéristiques locales	Caractéristiques globales
Illuminations [81]	Très sensible	Sensible
Expressions [76], [79]	Pas sensible	Sensible
Pose [82]	Sensible	Très sensible
Bruit [83]	Très sensible	Sensible
Occlusion [76], [79]	Pas sensible	Très sensible

Tableau 2.1. Comparaison des méthodes basées sur les caractéristiques locales et globales [80].

2.5.3 Les méthodes hybrides

Les méthodes hybrides sont des approches qui résultent de la combinaison des méthodes globales et locales, afin d'améliorer les performances de la reconnaissance faciale. Ainsi Latinis et al. [65] ont utilisé une méthode dite "Modèle Actif d'Apparence" (MAA) proposée par Cootes et al. [84] pour la reconnaissance de visages. Cette méthode modélise indépendamment la forme et la texture d'un visage en appliquant le PCA (voir la figure

2.12), les vecteurs de paramètres de forme et de texture obtenus sont ensuite utilisés pour la reconnaissance, toutefois cette méthode est coûteuse et dépend de nombreux paramètres.

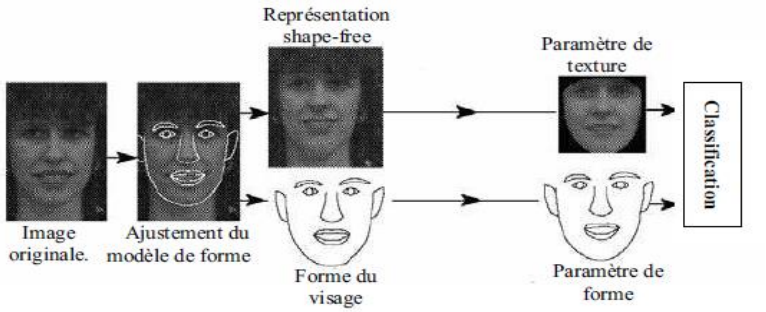


Figure 2.12. Processus de reconnaissance de visages basé sur l'MAA [65].

Notons que, d'un certain point de vue, les méthodes locales peuvent être considérées comme des méthodes hybrides, puisque les informations globales sont généralement prises en compte dans l'algorithme de reconnaissance. Par exemple, dans la méthode probabiliste locale [79] de nouveaux échantillons d'apprentissage pour chaque personne sont d'abord produits avec la méthode globale, une méthode locale est ensuite utilisée pour l'identification.

2.6 L'Analyse en Composantes Principales (PCA)

L'algorithme PCA, aussi connu sous le nom de *transformée de Karhunen-Loeve* [85], ou sous le nom *Eigenfaces* car il utilise des vecteurs propres (Eigenvectors) et des valeurs propres (Eigenvalues). C'est une méthode proposée par *M. A. Turk* et *A. P. Pentland* au *MIT Media Lab*, en 1991. L'idée principale de cette méthode consiste à extraire les caractéristiques discriminantes d'une image de visage, grâce à une projection orthogonale qui maximise la variance dans l'espace projeté, en minimisant ainsi l'erreur de la reconstruction.

Soit l'image $I_{i_{(m,n)}}$,et $\Gamma_{i_{(m\times n,\,1)}}$ le vecteur dans un espace vectoriel de grande dimension (N=m×n) obtenu en concaténant les colonnes de cette image (figure 2.13).

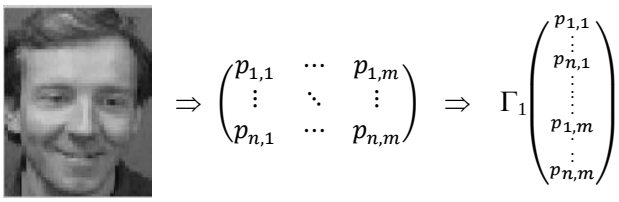


Figure 2.13. Passage d'une image vers un vecteur dans un espace vectoriel de grande dimension.

Où $p_{i,j}$ sont des coefficients représentant les valeurs des pixels en niveau de gris, codés de 0 à 255. Nous rassemblons ensuite les M images de la base de données dans une matrice où chaque colonne représente le vecteur Γ_i .

$$S = \{\Gamma_1, \Gamma_2, ..., \Gamma_m\} \tag{2.1}$$

Donc

$$S = \begin{pmatrix} p_{1,1} & b_{1,1} & \cdots & z_{1,1} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & b_{n,1} & \cdots & z_{n,1} \\ \vdots & \vdots & & \vdots \\ p_{1,m} & b_{1,m} & \cdots & z_{1,m} \\ \vdots & \vdots & & \vdots \\ p_{n,m} & b_{n,m} & \cdots & z_{n,m} \end{pmatrix}$$

La méthode Eigenface comporte deux processus, le processus d'apprentissage et le processus de reconnaissance :

Le processus d'apprentissage

- ✓ Nous calculons l'image moyenne ψ des vecteurs images tel que :

$$\psi = \frac{1}{M}\sum_{i=1}^{M} \Gamma_i \tag{2.2}$$

- ✓ Nous soustrayons de chaque image Γ_i l'image moyenne :

$$\phi_i = \Gamma_i - \psi, \qquad 1 \leq i \leq M \tag{2.3}$$

- ✓ Nous calculons la matrice de covariance C définie par :

$$C = \sum_{i=1}^{M} \phi_i \phi_i^T = A\,A^T, \qquad A = [\phi_1 \phi_2 \ldots \phi_M] \tag{2.4}$$

- ✓ Nous déterminons ensuite les vecteurs propres et les valeurs propres de C de taille (N×N). Cependant cette matrice est très grande et le temps de calcul sera de l'ordre du nombre de pixels dans l'image, autrement dit si nous avons des images de résolution de 300×200, nous allons résoudre une matrice C de 60000×60000. Plusieurs travaux [43], [86] ont démontré que le nombre M'< N d'Eigenfaces est généralement suffisant pour identifier efficacement les visages, ainsi les M' vecteurs correspondent aux plus grandes valeurs propres encodant les directions dans lesquelles les variations sont les plus

marquées, en fait ; ils présentent la majeure partie des informations telles que les différences d'éclairage, les visages des individus portant des lunettes ou une barbe.

Comment faire pour accélérer les calculs ?

- ✓ Soit e_i les vecteurs propres de $C = AA^T$ associés aux valeurs propres λ_i nous avons :

$$C\, e_i = \lambda_i e_i \tag{2.5}$$

- ✓ Et les vecteurs propres v_i de $L = A^TA$, associés aux valeurs propres μ_i tels que :

$$L\, v_i = \mu_i v_i$$

- ✓ Nous avons : $$A^TA\, v_i = \mu_i v_i$$

- ✓ En multipliant les deux côtés de l'égalité par A nous obtenons :

$$AA^TAv_i = A\mu_i v_i$$

- ✓ Et comme $C = AA^T$ nous pouvons simplifier :

$$C(Av_i) = \mu_i(Av_i) \tag{2.6}$$

- ✓ De (2.5) et (2.6), nous voyons que Av_i et μ_i sont respectivement les vecteurs propres et les valeurs propres de C :

$$\begin{cases} e_i = Av_i \\ \lambda_i = \mu_i \end{cases} \tag{2.7}$$

- ✓ Donc au lieu de calculer les valeurs propres de l'énorme matrice C, il suffit de calculer celles de la matrice L beaucoup plus petite. C'est ainsi pour trouver les vecteurs propres de C nous pré-multiplions les vecteurs propres de L par la matrice A. Nous ordonnons ensuite les vecteurs propres trouvés selon leurs valeurs propres correspondantes, de manière décroissante. Puis nous sélectionnons les k "meilleurs" vecteurs propres correspondant aux k plus grandes valeurs propres, en créant ainsi un espace vectoriel, que l'on appelle *l'espace des visages* E_v "*Face Space*". Les représentations graphiques de ces vecteurs donnent des images fantômes, chacune met en avant une partie du visage, on les appelle *Eigenfaces* (figure 2.14). Les images originales peuvent être reconstituées par combinaison linéaire de ces k vecteurs propres.

Figure 2.14. Exemple des cinq premiers Eigenfaces.

- ✓ Une fois les vecteurs de base trouvés, nous projetons nos images de départ sur le nouvel espace E_v. Une image Γ_i est alors transformée en ses composants Eigenfaces comme suit :

$$\omega_k = e_k^T(\Gamma_i - \psi), \qquad k = 1, \ldots, M'. \tag{2.8}$$

Où ω_k sont les poids, ils forment une matrice $\Omega^T = [\omega_1, \omega_2, \ldots, \omega_{M'}]$ qui décrit la contribution de chaque Eigenface dans la représentation des images d'entrée, c'est ainsi nous obtenons les poids de la base d'apprentissage $\Omega_{\Gamma_k}^T$, k = 1, ..., M′.

<u>Le processus de reconnaissance</u>

- ✓ Nous transformons l'image d'entrée $T_{(m\times n)}$ en un vecteur X de dimension (m×n). Nous déterminons ensuite ϕ la différence entre le vecteur X et le vecteur moyen ψ. Puis nous multiplions ϕ par l'ensemble des vecteurs propres e_k pour obtenir le vecteur de poids Ω_x.

- ✓ Nous cherchons la distance (ex. la distance Euclidienne) entre le visage à reconnaitre et les visages de la base d'apprentissage en prenant la plus petite valeur :

$$\varepsilon^2 = \| \Omega_\Gamma - \Omega_x \|^2 \tag{2.9}$$

- ✓ On dit qu'un visage appartient à une classe g, si le minimum ε_g est en dessous d'un certain seuil θ_ε, sinon le visage est considéré inconnu et peut éventuellement être utilisé pour créer une nouvelle classe de visage. Nous avons ainsi quatre possibilités pour une image de visage d'être reconnue ou non (figure 2.15) :

  - Cas 1 : un individu est reconnu et identifié.
  - Cas 2 : un individu inconnu du système est présent.
  - Cas (3 et 4) : l'image n'est pas une image de visage.
    - le cas 3, l'image est éloigné de E_v mais la projection est proche d'une classe connue, on a alors une *fausse acceptation.*

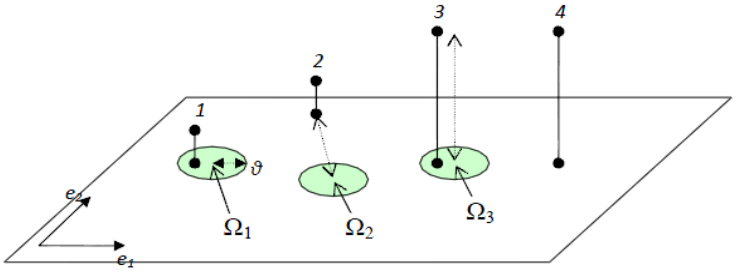


Figure 2.15. Illustration des cas possible de la projection d'une image sur E_v. Ici, nous avons deux vecteurs propres (e_1 et e_2) et trois classes d'individus connus (Ω_1,Ω_2, Ω_3).

2.7 Métriques de distances

La comparaison entre deux images est effectuée par une mesure métrique entre les projetés des images dans l'espace réduit. Perlibakas [87] a testé 14 métriques de distance, il a constaté que les performances de la reconnaissance faciale basée sur le PCA dépendent de la fonction des métriques utilisées. Parmi les mesures proposées dans la littérature nous citons :

- La distance de Minkowski d'ordre p notée L_p est définie par :

$$\mathrm{Lp} = \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{\frac{1}{p}} \tag{2.10}$$

- Pour p=1, nous obtenons la distance *City-Block* (ou distance de *Manhattan*) notée L_1 définie par :

$$\mathrm{L1} = \sum_{i=1}^{N} |x_i - y_i| \tag{2.11}$$

- Pour p=2, nous obtenons la distance *E*uclidienne notée L_2 définie par :

$$\mathrm{L2} = \sqrt{\sum_{i=1}^{N} |x_i - y_i|}^{\,2} \tag{2.12}$$

- La distance de Mahalanobis définie par :

$$d_{\mathrm{maha}} = -\sum_{i=1}^{N} \frac{1}{\sqrt{\lambda_i}} x_i y_i \tag{2.13}$$

Où λ_i sont les vecteurs propres associés aux vecteurs propres.

- La distance cosinus définie par :

$$d_{\cos} = -\frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N}(x_i)^2 \sum_{i=1}^{N}(y_i)^2}} \quad (2.14)$$

Des travaux [88], [89] faits sur la base FERET ont prouvé que la distance de Mahalanobis est la plus efficace. Cependant la distance cosinus est la mesure de similarité la plus largement utilisée, du fait qu'elle est rapide à calculer et qu'elle donne des performances très proches de celles de la distance de Mahalanobis.

2.8 Conclusion

Dans ce chapitre, nous avons présenté les techniques les plus populaires utilisées en reconnaissance faciale automatique. Ces méthodes peuvent être classées en trois catégories : les méthodes globales, les méthodes locales et les méthodes hybrides. Nous avons exposé leurs avantages et leurs inconvénients, en effet de nombreux paramètres doivent être pris en compte afin de comparer les algorithmes : la base de données utilisée, la taille des images, le nombre d'images de l'enrôlement et de test et les variations d'acquisition (le tableau 2.2 résume les performances de quelques approches de reconnaissance de visage sur certaines bases de données de la littérature). Nous avons ensuite opté pour la méthode "Eigenface" qui se base sur l'Analyse en composantes principales (PCA) ; méthodes mathématique permet de représenter efficacement les images de visages en réduisant la dimensionnalité de l'espace.

Réf.	Méthode	Base de données	Taille des images	Nb. Images	Taux (%)	Expr.	Ill.	Pose
[43]	PCA LDA	AR AR	85*60 85*60	100–250 100–250	70 88		N N	N N
[54]	DCV	Yale AR	126*152 229*299	15–150 350–350	97.33 99.35		O	N
[50]	ICA	FERET	60*50	425–421	89	O	N	N
[69]	EBGM	FERET	256*384	250–250	80	O		O
[90]	PCA	UND		166–166	98	O	O	N
[47]	PCA	Equinox	99*132	770–2310	93	O	O	N
[91]	ICA	AR,Yale, ORL, Bern, FERET	46*56	1685–1490	98		O	O

Tableau 2.2. Comparatif de quelques méthodes de reconnaissance de visage sur certaine base de données (Nb. Images, indique le nombre d'images utilisées pour l'enrôlement et le nombre d'images utilisées pour les tests, les colonnes Expr., Ill. et Pose, indiquent si les images possèdent des variations d'expression faciale, d'illumination ou de pose (O pour Oui, N pour Non).

Chapitre 3

La reconnaissance des empreintes digitales

3.1 Introduction

L'utilisation de l'empreinte digitale comme moyen d'identification est très ancienne, son histoire remonte au moins à 6000 av. JC. Trois siècles av. JC les Chinois apposaient des empreintes sur les documents officiels. Depuis 1897 la dactyloscopie (étude des empreintes digitales sans aide d'un ordinateur) connaît une avancée majeure dans les enquêtes criminelles. Au début de 1960, le premier système automatique de reconnaissance des empreintes digitales est mis en place. De nos jours, cette technique est largement utilisées et reconnue comme méthode d'authentification fiable.

L'empreinte digitale est le modèle du relief cutané des doigts formée à environ sept mois du développement du fœtus, elle est unique pour chaque individu. *Francis Galton* a démontré (en 1888) la permanence de sa forme tout au long de la vie de la personne, ainsi que son inaltérabilité. Depuis longtemps le seul moyen de l'acquisition de cette modalité consiste à utiliser l'encre, les avancées technologiques ont permis l'apparition des appareils numériques basés sur la capture optique, thermique, électromagnétique ou sur les ultrasons [6].

3.2 Caractéristiques d'une empreinte digitale

L'empreinte digitale se compose d'un ensemble de lignes localement parallèles dessinées sur l'épiderme appelées "les stries" (ou les crêtes qui sont les reliefs en contacts avec la surface au toucher) et des creux entre les stries appelées "les vallées" (ou les sillons). La largeur des stries et des vallées peut être estimées entre 200 et 800 μm selon les individus (voir la figure 3.1).

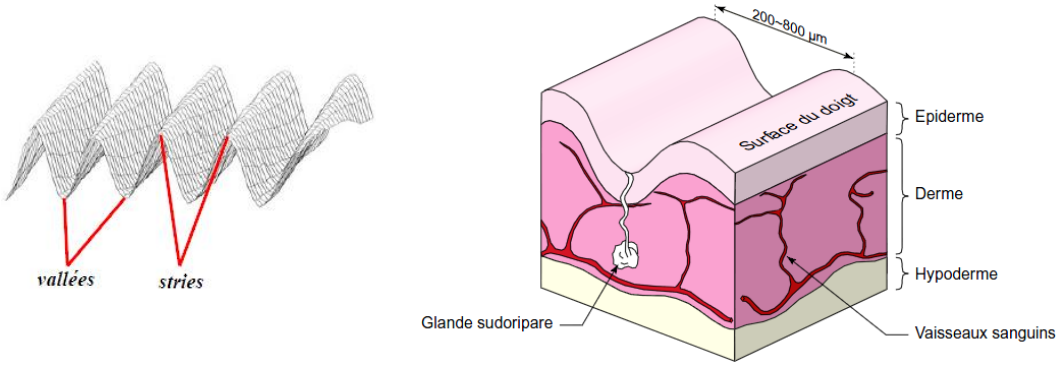


Figure 3.1. Vue en coupe de la peau au niveau du doigt

L'empreinte digitale est caractérisée par ses propriétés globales et locales. Typiquement, les représentations globales sont utilisées pour la classification d'empreinte digitale, alors que les représentations locales sont destinées à la comparaison de cette modalité.

- La représentation globale : chaque empreinte digitale a un ensemble des *points singuliers globaux* qui sont les *centres* et les *deltas*, symbolisés respectivement par O et Δ. Le *centre* est le lieu de convergence des stries (il est aussi appelé le *core*), alors que le *delta* correspond au lieu de divergence. La position et le nombre de ces points permettent la classification des empreintes digitales, c'est ainsi que *Francis Galton* les a subdivisées en trois grandes familles [92] :

  - Les *Boucles* (*Loops*) : une empreinte est de classe boucle si ses stries rentrent d'un côté et ressortent du même côté et si elle possède un point singulier de type boucle et un point singulier de type delta. Les boucles représentent 65% des empreintes des doigts humains.

  - Les *Spires* (*Whorls*) : une empreinte appartient à la classe spire si elle possède au moins une strie qui fait 360°. Elle peut aussi contenir jusqu'à deux régions singulières de type boucles et deux régions singulières de type deltas. Les spires représentent 30% des empreintes des doigts humains.

  - Les *Arches* (*Archs*) : une empreinte est de classe arche si elle possède des stries qui rentrent d'un côté et ressortent du côté opposé et si elle ne contient ni boucle ni delta comme points singuliers. Les arches ne représentent que 5% des empreintes des doigts humains.

Edward Henry les a classées [93] en six sous-classes principales : *arche*, *boucle à gauche* (*left loop*), *boucle à droite* (*right loop*), *arche penchée* (*tented arch*), *spires* et *spires imbriquées* ou *boucles jumelles* (voir la figure 3.2).

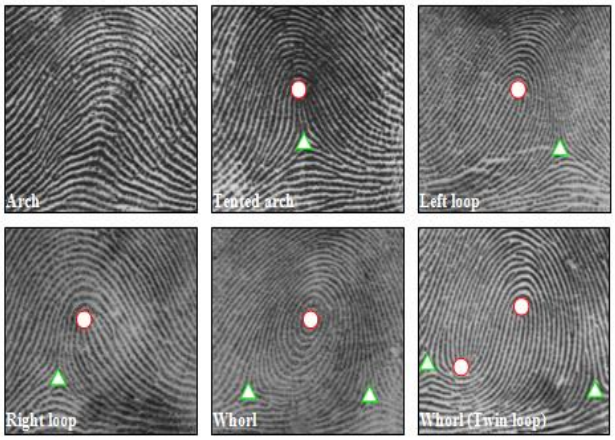
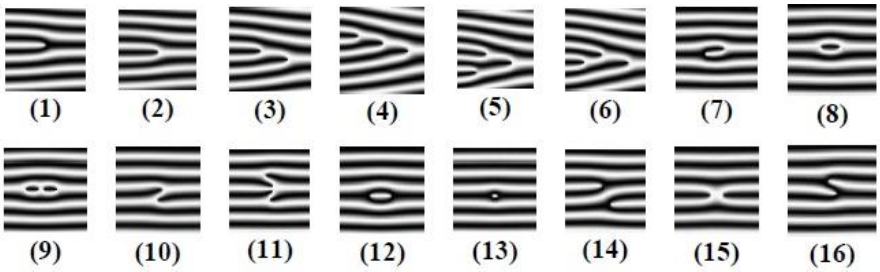


Figure 3.2. Les principales classes d'empreintes digitales selon la classification de Galton-Henry [94].

- La représentation locale : il s'agit des caractéristiques les plus utilisées "*les minuties*" (littéralement : petits détails), qui sont en fait les points d'irrégularités qui se trouvent sur les lignes capillaires. Nous pouvons distinguer jusqu'à seize types de minuties différentes (voir la figure 3.3), mais dans les algorithmes on n'en s'intéresse qu'aux deux types suivants parce qu'ils sont facilement détectables :

  - La bifurcation : c'est le point où la strie se divise en deux.
  - La terminaison : c'est le point où la strie s'arrête.

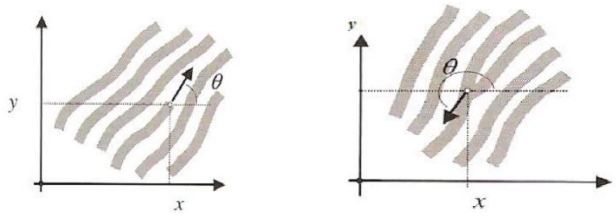


1	Terminaison	9	Boucle double
2	Bifurcation simple	10	Pont simple
3	Bifurcation double	11	Pont jumeau
4	Bifurcation triple I	12	Intervalle
5	Bifurcation triple II	13	Point isolé
6	Bifurcation triple III	14	traversée
7	crochet	15	Croisement
8	Boucle simple	16	Tête bêche

Figure 3.3. Les types de minutie.

En fait les autres types de minuties ne sont que les résultats des combinaisons des minuties de terminaison et de bifurcation. Par exemple, les boucles peuvent être visualisées en tant que deux bifurcations.

Chaque minutie est représentée par les cordonnées (x, y) de sa position dans l'image, et l'angle θ qui est la direction associé à la strie (voir la figure 3.4). Donc chaque minutie de l'empreinte est repérée par un vecteur de la forme : (Type de minutie, x, y,θ).



(Terminaison, x, y,θ). (Bifurcation, x, y,θ).

Figure 3.4. Représentation des vecteurs de terminaison et de bifurcation [94].

3.3 Structure d'un système de reconnaissance d'empreintes digitales

Comme tout système biométrique la reconnaissance de l'empreinte digitale se base sur l'architecture présentée dans le chapitre I. la figure 1.1 illustre le principe de son fonctionnement : après l'acquisition de l'image de l'empreinte digitale cette dernière va subir d'un prétraitement pour améliorer sa qualité en supprimant les zones de bruit, beaucoup d'algorithmes ont été proposés dans ce cadre. Ensuite une extraction des caractéristiques est effectuée suivie d'un stockage dans la base de données après une éventuelle classification.

3.3.1 L'extraction des minuties

Bien qu'il existe plusieurs méthodes d'extraction des minuties nous pouvons les classer en deux catégories : celles qui se basent sur la binarisation de l'image de l'empreinte digitale (c.à.d. la conversion de l'image en une image noir et blanc) dites les méthodes classiques, et celles qui travaillent directement sur l'image filtrée [94]. Dans notre travail nous avons choisis d'utiliser la première approche.

3.3.1.1 La méthode basée sur la binarisation

La méthode basée sur la binarisation (ou *Binarization-based methods*) est la méthode la plus utilisée, elle consiste à extraire les minuties à partir du squelette binaire de l'image filtrée (figure 3.5).

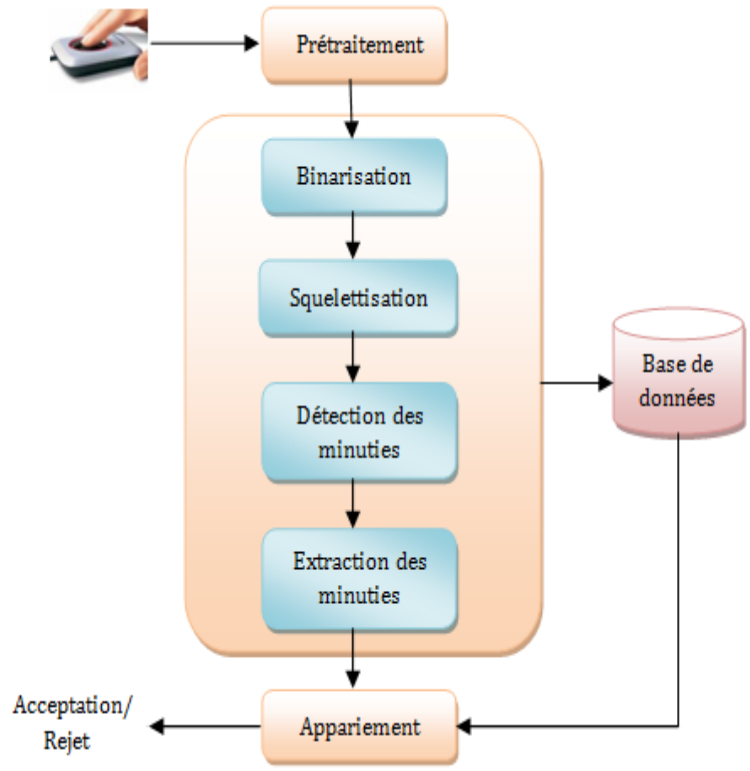


Figure 3.5. Représentation d'un système de reconnaissance d'empreinte digitale basée sur la binarisation.

A. la binarisation

Pour augmenter la visibilité des minuties l'image doit d'abord être binarisée, il s'agit de donner une intensité aux stries et une autre intensité différente aux vallées. Plusieurs techniques de binarisation ont été proposées [95], la plus utilisée consiste à comparer chaque pixel à un seuil S, en lui assignant la valeur un (blanc) si son intensité est supérieure à S sinon il prend la valeur zéro (noir). Le seuil S peut être global fixé dès le départ ou local adaptatif obtenu en calculant la moyenne des pixels de chaque bloc de l'image de l'empreinte digitale.

B. la squelettisation

Pour pouvoir détecter rapidement les minuties l'image doit être squelettisée : c'est une sorte d'amincissement basée sur des opérations morphologiques itératives, dans le but d'obtenir une image plus schématique de l'empreinte dans laquelle toutes les stries ont un pixel de largeur tout en conservant la connexité (c.à.d. respecter la continuité des stries). Nous distinguons plusieurs algorithmes de squelettisation dans la littérature [96].

Figure 3.6. Résultats des étapes de binarisation et de squelettisation, (a) Empreinte originale, (b) Empreinte binarisée, (c) Empreinte squelettisée.

Une fois l'image squelettisée, nous passons à l'extraction de signature. Cette dernière doit être la plus représentative de l'empreinte digitale c.à.d. le nombre de minuties extraites doivent être *suffisant*, 12 au minimum (pour pouvoir effectuer une comparaison fiable entre les empreintes), et *fiable* (c.à.d. les minuties retenues ne font pas partie des zones bruitées ni des zones altérées temporairement de l'empreinte digitale comme les blessures par exemple).

C. la détection des minuties

Il existe plusieurs méthodes de détection des minuties, la plus répandue est celle initiée par Arcelli [97] qui traite une fenêtre de 3×3 autour chaque pixel I, I (i, j) entouré ainsi de huit pixels (N_i, $i \in [1..8]$). L'idée de ce processus consiste à identifier comme bifurcation le pixel avec trois pixels voisins c.à.d. $\sum_{i=1}^{8} N_i = 3$ et identifier comme terminaison le pixel avec un pixel voisin c.à.d. $\sum_{i=1}^{8} N_i = 1$. Le nombre $(\sum_{i=1}^{8} N_i)$ s'appelle le *nombre de connexions (ou crossing number CN)*. Nous obtenons donc cinq cas différents (voir la figure 3.7) :

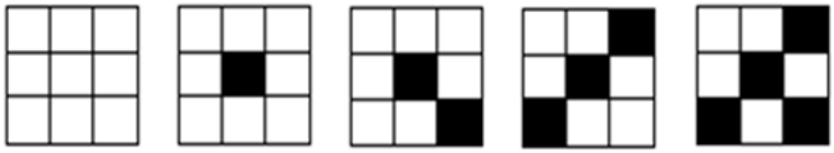


Figure 3.7. Les cinq cas obtenus lors de processus d'extraction des minuties. Les pixels de valeur 1 et ceux de valeur 0 sont respectivement symbolisés par les carrés noirs et les carrés blancs.

- I (i, j) = 0 : c'est une vallée de l'empreinte digitale.
- I (i, j) = 1 et aucune stries voisine (c.à.d. $\sum_{i=1}^{8} N_i = 0$) : c'est un point isolé (voir la figure 3.3), cependant nous n'en tenons pas compte, car il existe rarement dans une empreinte, il est probablement dû à un bruit.
- I (i, j) = 1 et une seule strie voisine (c.à.d. $\sum_{i=1}^{8} N_i = 1$) : c'est une terminaison.
- I (i, j) = 1 et deux stries voisines (c.à.d. $\sum_{i=1}^{8} N_i = 2$) : il s'agit ici d'une ligne continue c.à.d. il n'y a pas de minutie.

- I (i, j) = 1 et trois stries voisines (c.à.d. $\sum_{i=1}^{8} N_i = 3$) : c'est une bifurcation.

Le cas de I (i, j) = 1 et plus de trois stries voisines (c.à.d. $\sum_{i=1}^{8} N_i > 3$) : Ne compte pas parce que les stries de l'empreinte digitale ne sont pas transversales.

À la fin de ce processus, nous acquérons un grand nombre des minuties, 100 minuties en moyenne. Parmi lesquelles environ 60% sont des minuties érronées, y compris celles introduites lors des étapes de binarisation et de squelettisation (voir la figure 3.8). Donc un post-traitement éliminant les fausses minuties est nécessaire avant de passer à l'étape suivante.

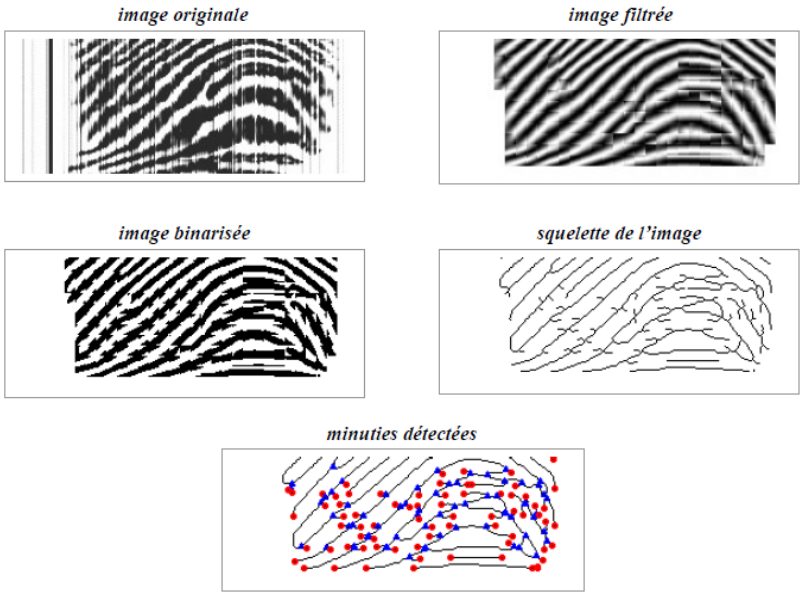


Figure 3.8. La détection des fausses minuties [98].

D. L'élimination des fausses minuties

Dans cette étape nous allons éliminer le maximum des fausses minuties détectées au cours des étapes de binarisation et de squelettisation. Pour cela nous utilisons des considérations heuristiques [99], [100] basées sur le fait que la distance entre deux minuties voisines ne doit pas dépasser un certain seuil. Pratiquement, si on trouve plusieurs minuties dans une petite région cela indique la présence de bruit, deux terminaisons plus proches indiquent une coupure dans la strie [94], ainsi que, beaucoup de fausses minuties se situent généralement au bord de l'image, d'autres considérations se présentent dans [99] [101]. La figure 3.9 illustre les fausses minuties qui peuvent être rencontrées.

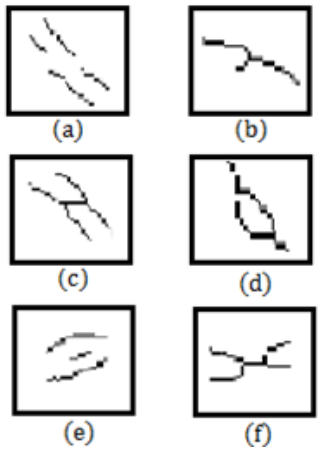


Figure 3.9. Type des fausses minuties, (a) coupure dans les stries, (b) segment court, (c) pont, (d) ilot, (e) segment court, (f) segment court.

Bien que les méthodes basées sur la binarisation soient faciles à mettre en place et moins coûteuses, elles ont quelques défauts qui résident principalement dans la possibilité de perdre beaucoup d'informations due au bruit introduit lors de binarisation. De plus la binarisation et la squelettisation exigent un temps de calcul important vient de plusieurs balayage de l'image, ce qui diminue la performance des méthodes classiques, cela amène à introduire les méthodes d'extraction directe.

3.3.1.2 La méthode d'extraction directe

La méthode d'extraction directe (ou *Direct gray-scale methods*) est proposée par D. Maio et D. Maltoni [102]. Elle consiste à extraire l'ensemble des minuties directement sur l'image filtrée en niveau de gris sans passer par les étapes de binarisation et de squelettisation. Le principe de cette approche se base sur le suivi des stries de l'image directionnelle selon le maximum local.

- **L'image directionnelle**

C'est l'estimation de l'orientation de lignes de l'empreinte digitale. Soit I l'image de l'empreinte en noir et blanc, $I_{(i,\ j)}$ représente le pixel (i, j). L'image directionnelle (ou l'orientation de l'empreinte) c'est la matrice notée D dont chaque case (i, j) contient l'orientation locale θ (i, j) entre l'axe horizontal et l'orientation (approximative) des lignes de l'empreinte dans le voisinage du pixel (i, j) (figure 3.10).

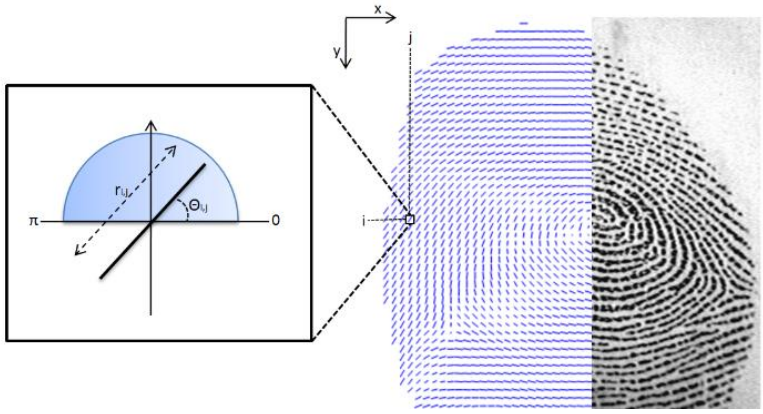


Figure 3.10. L'image directionnelle [103].

L'approche la plus simple et la plus courante pour calculer la matrice D consiste à calculer les gradients aux différents points de l'image I. RAO [104] a proposé une méthode efficace pour estimer le champ d'orientation de l'empreinte. Il divise l'image en blocs de taille W×W pixels, puis il calcule les gradients G_x et G_y dans les directions de x et de y respectivement pour tous les pixels (i, j) dans chaque bloc, ensuite il estime l'orientation locale de ces pixels en utilisant les formules ci-dessous :

$$V_x(i,j) = \sum_{u=i-\frac{W}{2}}^{i+\frac{W}{2}} \sum_{u=j-\frac{W}{2}}^{j+\frac{W}{2}} \left(2G_x(u,v)\, G_y(u,v)\right) \tag{3.1}$$

$$V_y(i,j) = \sum_{u=i-\frac{W}{2}}^{i+\frac{W}{2}} \sum_{u=j-\frac{W}{2}}^{j+\frac{W}{2}} \left(G_x^2(u,v) - G_y^2(u,v)\right) \tag{3.2}$$

$$\theta(i,j) = \frac{1}{2}\tan^{-1}\left(\frac{V_x(i,j)}{V_y(i,j)}\right) \tag{3.3}$$

- **Le suivi des stries**

Le suivi des stries nécessite l'estimation du champ d'orientation de l'empreinte. Soit le point (i_a, j_a) appartenant à une strie, pour la parcourir nous avançons de μ pixels dans la direction θ_a afin d'atteindre un point soit x_a. Ensuite nous cherchons le maximum local le plus proche de x_a sur la section orthogonale à la direction θ_a centrée en x_a de longueur σ en obtenant un autre point (i_b, j_b) de la strie (voir la figure 3.11). Nous recommençons le même processus en prenant comme nouveau point de départ le point (i_b, j_b), on s'arrête lorsqu'on détecte une bifurcation ou une terminaison. On refait les mêmes opérations avec toutes les

stries de l'empreinte digitale. Pour assurer qu'une crête n'est parcourue qu'une seule fois, nous adoptons la stratégie d'étiquetage, c'est ainsi nous détectons toutes les minuties de l'empreinte.

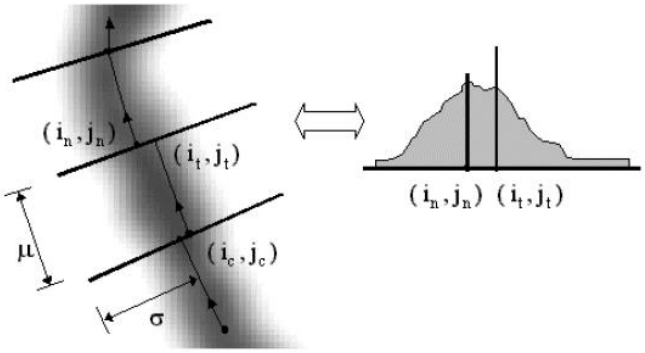


Figure 3.11. Le suivi d'une strie [102].

Le choix des paramètres μ et σ est crucial dans le déroulement du suivi. Une valeur faible de σ peut empêcher la détection du maximum local. Ainsi que l'utilisation d'une valeur faible de μ augmente le temps de calcul relatif au nombre considérable des opérations nécessaires au suivi d'une strie, tandis que la valeur trop élevé de μ risque d'entraîner le saut d'une crête (on peut ainsi s'écarter de la strie qu'on veut parcourir). D. Maio et Maltoni [102] choisissent μ en fonction des caractéristiques globales (en effectuant le test sur un ensemble de 150 images d'empreintes digitales, ils obtiennent un taux moyen d'erreur élevé), alors que X. Jiang et al. [105] adaptent localement μ en fonction des variations locales de courbure et de contraste, ce qui accélère les calculs.

Bien que l'extraction des minuties directement à partir de l'image filtrée soit plus rapide et produit très peu de fausses minuties par rapport à celles générées par les méthodes basées sur la binarisation, cependant l'ajustement des paramètres μ et σ est très sensible aux fortes variations locales de la distance inter-strie.

3.3.2 La comparaison d'empreintes digitales

La comparaison des empreintes digitales consiste à réaliser un accord entre la signature à identifier et les signatures stockées dans la base de données. Cependant cette tâche n'est pas facile, notamment à cause de la variabilité dans les différentes impressions d'une même empreinte (*variation intra-classe*). Les facteurs principaux responsables des variations intra-classe sont : l'état de la peau (exemple le doigt n'est pas propre entaché de graisse ou de sueur), la variation de la pression du doigt sur le capteur d'empreinte, le déplacement, la rotation c'est ainsi qu'un déplacement du doigt de 2mm qui est imperceptible à l'œil humain provoque une translation d'à peu près 40 pixels dans une image scannée à une résolution de 500 dpi [94]. Donc les algorithmes de comparaison doivent prendre en compte tous ces

paramètres, en fait parfois on doit comparer un petit bout d'image de l'empreinte avec l'image entière de la même empreinte [107], comme c'est le cas dans les enquêtes criminelles. La figure 3.12 illustre quelques exemples des images d'une même empreinte digitale extraites de la base de données FVC2002.

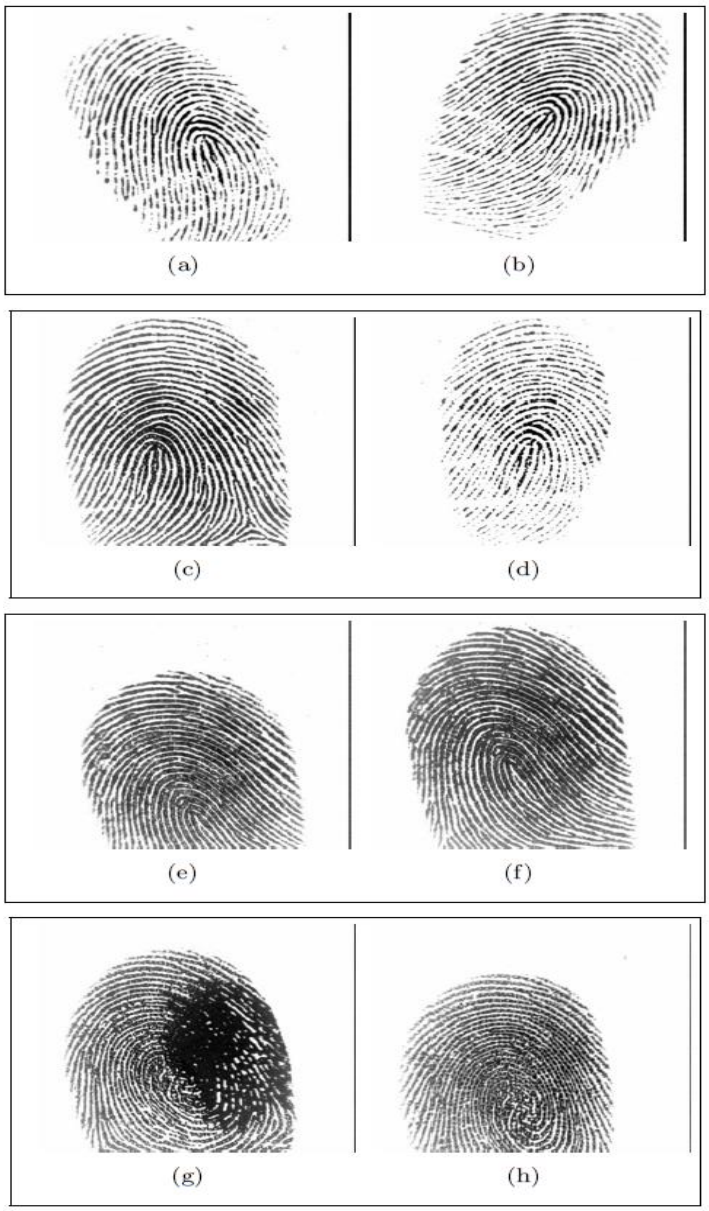


Figure 3.12. Les variations intra-classe d'une même empreinte digitale, (a) le doigt est incliné à gauche sur le capteur d'empreinte, (b) le doigt est incliné à droite, (c) le doigt est plus pressé, (d) le doigt est légèrement pressé, (e) une partie du doigt n'est pas apposée sur le capteur d'empreinte, (f) le doigt est correctement apposé sur le capteur d'empreinte, (g) le doigt est sale, (h) le doigt est propre [94].

Dans la littérature, beaucoup d'algorithmes de comparaison peuvent être distinguées, ils sont classifiés en trois grandes familles :

a) Les approches basées sur la corrélation (correlation-based approches) : dans ces approches, deux images d'empreinte digitale sont superposées et la corrélation entre les pixels correspondants sera calculée pour différents alignements (ex. rotation, déplacement). Les techniques basées sur la corrélation de pixels ne sont pas efficaces car elles sont sensibles à la variation intra-classe. En fait les différentes impressions d'une même empreinte peuvent donner des images très différentes, ce qui rendent les valeurs de leurs pixels différents (voir la figure 3.12). De plus ces approches sont coûteuses en temps de calcul, ce problème peut être résolu en calculant seulement la corrélation de certaines régions locales déterminées suivant quelques critères, toutefois les comparaisons basées sur la corrélation de pixels ne sont pas assez rigoureuses [106].

b) Les approches basées sur les minuties (minutiae-based approches) : ce sont les approches les plus utilisées. Les minuties sont extraites de deux empreintes digitales et représentées sous forme d'un ensemble de points dans un plan à deux dimensions selon le modèle des coordonnées (figure 3.4). L'assortiment basé sur les minuties consiste à trouver un bon alignement entre les minuties de deux empreintes à comparer qui donne un maximum de paires de minuties semblables. Une marge de tolérance [94] afin de compenser les erreurs introduite par le bruit et la distorsion sera définie (figure 3.13). Nous allons développer ces concepts au point 3.3.2.1.

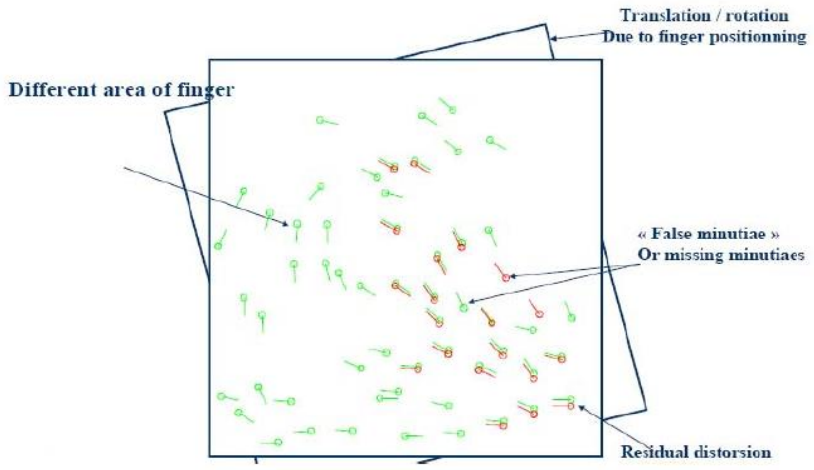


Figure 3.13. La comparaison basée sur les minuties [108].

c) *Les approches basées sur les rides (ridge or texture-based approches)* : lorsque la qualité de l'image de l'empreinte digitale est mauvaise l'extraction des minuties est très difficile voire même impossible [109], [110]. D'autre particularités des rides d'empreinte telles que l'orientation locale, la fréquence, la texture, les pores de respirations, la forme des rides etc. peuvent être extraites plus facilement que les minuties, cependant la distinction de ces singularités est faible, pour cela ces techniques sont très peu utilisées. La figure 3.14 illustre l'extraction des informations de texture basées sur l'orientation locale d'une image d'empreinte digitale.

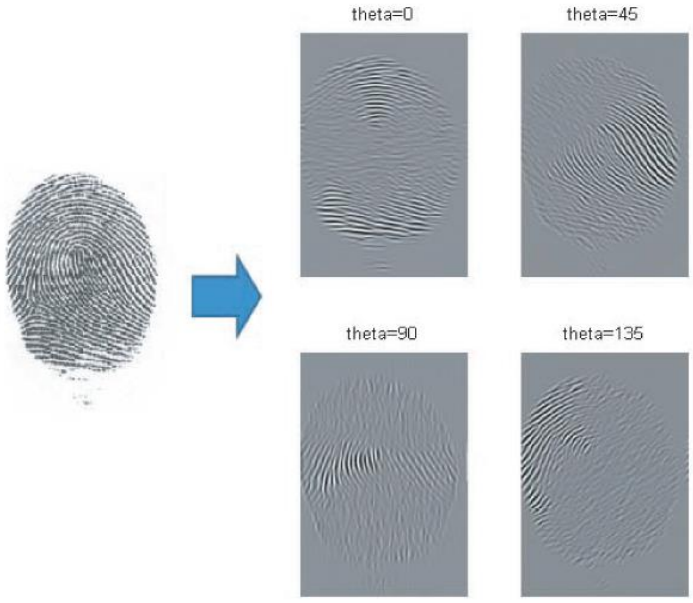


Figure 3.14. L'extraction des informations de texture basées sur l'orientation locale de l'empreinte digitale [111].

3.3.2.1 La comparaison basée sur les minuties

La reconnaissance d'empreintes basée sur la comparaison des minuties consiste à comparer les minuties extraites de deux empreintes dans un plan bidirectionnel et de retourner les paires de minuties ''matchées'' c.à.d. les paires possédant le même emplacement et la même orientation.

Soit I et T deux représentations de deux empreintes digitales. La représentation est un vecteur de minuties extraites, où chaque minutie m est représentée à son tour par un vecteur (x, y,θ) où x et y spécifient les coordonnées de la position de la minutie dans l'image de l'empreinte et l'angle θ son orientation.

Considérons les deux ensembles de minuties suivants :

$$I = \{m_1, m_2, \dots, m_3\} \qquad m_i = \{x_i, y_i, \theta_i\} \qquad i=1..m.$$

$$T = \{m'_1, m'_2, \dots, m'_3\} \qquad m'_j = \{x'_j, y'_j, \theta'_j\} \qquad j=1..n.$$

Où m et n représentent respectivement le nombre des minuties de I et T.

Une minutie m'_j de T est considérée appariée avec une minutie m_i de I si la distance *sd* (pour *spatial distance*) entre elles est inférieure ou égale à une certaine tolérance r_0 et la différence d'angle entre leurs orientations *dd* (pour *direction difference*) est inférieure ou égale aussi à une certaine tolérance θ_0 :

$$sd(m'_j, m_i) = \sqrt{(x'_j - x_i)^2 + (y'_j - y_i)^2} \quad \leq r_0 \tag{3.4}$$

$$dd(m'_j, m_i) = \min(|\theta'_j - \theta_i|, 360 - |\theta'_j - \theta_i|) \quad \leq \theta_0 \tag{3.5}$$

Où *sd* et *dd* sont nécessaires pour compenser les erreurs inévitables introduites par les algorithmes d'extraction des minuties ainsi que les petites déformations dues aux conditions d'acquisition de l'empreinte digitale. Cependant ces marges de tolérance sont insuffisantes pour faire une comparaison optimale c'est pour cela qu'une étape d'alignement est obligatoire. Cette dernière consiste à donner une tolérance supplémentaire (Δx,Δy,Δθ) ajoutée à chaque minutie m'_j de T pour s'adapter à l'alignement de minuties m_i de I (figure 3.15).

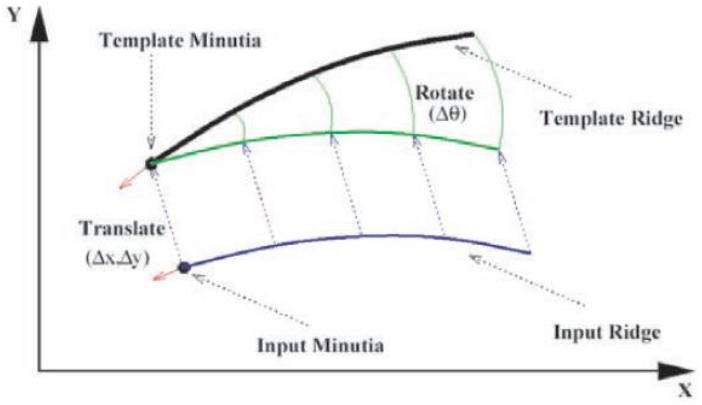


Figure 3.15. L'alignement de minuties de deux empreintes digitales [112].

En fait, on effectue plusieurs alignements pour avoir un nombre maximum de minuties assorties. Deux empreintes sont considérées identiques si un certain nombre de minuties identiques est atteint.

Soit une fonction *mm* (pour *munitiae matcher*) de comparaison de minuties :

$$mm(m_j, m_i) = \begin{cases} 1 & si\ sd(m_j, m_i) \leq r_0\ et\ dd(m_j, m_i) \leq \theta_0 \\ 0 & sinon \end{cases} \tag{3.6}$$

Le nombre de minuties matchées est donné par :

$$max \sum_{i=1}^{m} mm\,(m'_{p(i)}, m_i) \tag{3.7}$$

Où p(i) est une fonction telle que :

- ✓ p(i) = j si m'_j est la minutie la plus semblable à m_i parmi toutes les minuties de T qui satisfont les équations (3.4) et (3.5).
- ✓ p(i) = null si $\forall\, k = 1..n, mm(m_k, m_i) = 0$

3.4 Conclusion

Au cours de ce chapitre consacré à la reconnaissance de l'empreinte digitale ; nous avons vu les caractéristiques de l'empreinte, sa classification ainsi que la structure globale d'un système de reconnaissance d'empreintes. Nous avons décrit deux techniques d'extraction des minuties, une méthode classique qui exige un squelette binaire de l'image filtrée de l'empreinte et une autre méthode directe basée sur le suivi des stries. Nous avons également représenté les différentes approches de comparaison des empreintes digitales existant dans la littérature en focalisant sur la comparaison basée sur les minuties. Le chapitre suivant sera dédié à la fusion de modalités biométriques.

Chapitre 4

La fusion de modalités

4.1 Introduction

Comme nous l'avons vu au premier chapitre, les systèmes biométriques peuvent être monomodaux (basés sur une seule modalité biométrique) ou multimodaux (basés sur plusieurs modalités biométriques). Cette dernière technique permet d'améliorer les performances de la reconnaissance en combinant les informations de plusieurs systèmes. Nous allons maintenant traiter la question de la fusion. Nous allons parler des niveaux de fusion dont le plus courant est la fusion au niveau des scores (l'objet de ce chapitre). Nous abordons ses types, en rappelant les méthodes de normalisation des scores.

4.2 Les niveaux de fusions

En biométrie monomodale nous suivons une chaine de traitement qui comporte quatre modules (module de capture, d'extraction des caractéristiques, de comparaison et le module de décision). L'introduction de la multimodalité consiste à fusionner des informations disponibles dans n'importe quel de ces modules (nivaux). Deux familles de fusions peuvent être distinguées, la fusion avant comparaison (matching) et la fusion après comparaison [113].

4.2.1 Fusion avant matching

La fusion avant matching (*pré-classification*) correspond à la fusion des informations au niveau du module de capture ou au niveau du module d'extraction des caractéristiques.

Niveau Capteur (*Sensor Level*)

Il s'agit ici de fusionner les données brutes ("*raw data*") qui proviennent de plusieurs capteurs [114] pour obtenir une nouvelle donnée de même trait biométrique. Ce type de fusion nécessite une homogénéité des informations acquises. Il est impossible, par exemple ; à ce niveau de combiner l'image de visage avec celle du démarche, mais plutôt on peut fusionner des images 2D de visage pour obtenir une image 3D. Un autre exemple consiste à réaliser une mosaïque de plusieurs images d'empreinte digitale afin de créer une image plus complexe de cette empreinte [115], [116]. Parmi les méthodes proposées pour la fusion au niveau capteur on trouve celle de Bebis et al. [117] qui utilisent des algorithmes génétiques pour calculer les poids de la fusion.

Niveau Caractéristiques (*Feature Level*)

La fusion au niveau des caractéristiques consiste à combiner plusieurs caractéristiques issues de différentes modalités biométriques. Contrairement à la fusion au niveau capteur, cette fusion ne nécessite pas une homogénéité entre les données. Donc si les vecteurs de caractéristiques en entrée sont homogènes, le vecteur de caractéristique résultant est calculé comme une somme pondéré de ces vecteurs. Lorsque les vecteurs de caractéristiques à fusionner sont hétérogènes, leur concaténation est souvent le moyen utilisé pour créer un nouveau vecteur, c'est ainsi que Jing et al. [118] ont proposé une méthode de fusion des caractéristiques de modalité visage et d'empreinte palmaire en effectuant une concaténation des images par la transformée de Gabor. Toutefois la concaténation augmente l'espace de classification, ce qui pose le problème de la malédiction de la dimension [119].

Malgré que la fusion avant matching est plus efficace dans la mesure où elle intègre des informations riches que celles fournies par les systèmes de fusion après matching, elle a un certain nombre de contraintes, tels que le coût de l'application, la dimension gigantesque du vecteur de caractéristiques obtenu par la concaténation des vecteurs entrants, ainsi que la relation entre les espaces de caractéristiques (*feature spaces*) des différents systèmes biométriques n'est pas toujours connue, ce qui conduit les chercheurs à s'intéresser à la fusion après matching.

4.2.2 Fusion après matching

La fusion après matching (*post-classification*) combine l'information au niveau décision ou au niveau score.

Niveau Décision (*Decision Level*)

La fusion au niveau des décisions est simple et abstraite. En effet chaque système décide individuellement la correspondance possible "accepté" ou "rejeté" que l'on peut représenter

par 1 et 0, puis le système de fusion prend une décision finale en fonction de la série de réponses (1 et 0). Les méthodes les plus utilisées sont *le majority voting* [120] (si la majorité des systèmes ont décidé 1 alors la décision finale est "accepté"), *les règles ET* (si tous les systèmes ont décidé 1 alors la décision finale est "accepté") *et OU* (si un système a décidé 1 alors la décision finale est "accepté") [121], *le behavior knowledge space* [122], et *le weighted voting* [123].

Niveau Score (*Score Level*)

La fusion au niveau des scores est l'approche la plus utilisée, parce qu'elle est simple et facile à implémenter, de plus les données en sortie des modules de reconnaissance biométriques (*Matchers*) possèdent une grande richesse d'information. Ce niveau de fusion peut être appliqué à tous les types de biométries dans un espace de dimension limité (en fait, le vecteur de scores résultant a une dimension égale au nombre de sous-systèmes). Dans le reste de ce chapitre nous allons détailler la fusion au niveau des scores.

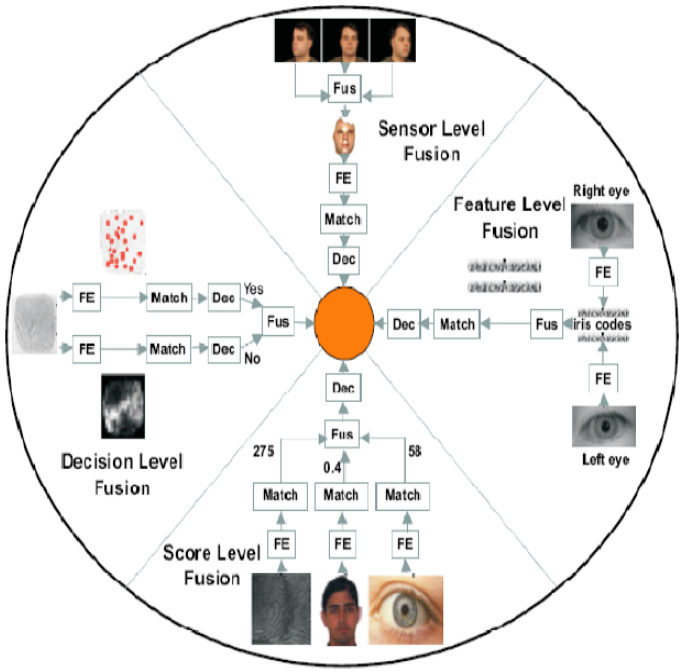


Figure 4.1. Les différents niveaux de fusion [20].

4.3 La fusion au niveau des scores

Il existe plusieurs méthodes pour combiner les scores obtenus par les modules de reconnaissances biométriques, nous pouvons les classifier en deux grandes familles : les *méthodes de combinaison* et les *méthodes de classification* des scores.

Dans les méthodes de combinaison, les scores sont traités séparément avant de les combiner pour construire un unique score qui sert à prendre une décision finale. Cependant, les scores provenant de différents matchers doivent être cohérents entre eux, pour cela, une étape de normalisation qui transforme les scores dans un domaine commun est nécessaire avant de les combiner.

Dans les méthodes de classification de scores, on construit un vecteur de caractéristiques en utilisant les scores de correspondance en sorties de chaque matcher, puis on l'affecte à l'une des deux classe : client ou imposteur. Le classifieur adopté, afin de réaliser ce processus, est apte à prendre la décision sans tenir compte de la manière dont le vecteur de caractéristiques a été créé, autrement dit ; les scores en sortie de différents sous-systèmes peuvent être non homogènes, et ils n'exigent aucun traitement avant de les envoyer dans le classifieur.

Jain et al. ont montré que les méthodes basées sur la combinaison des scores sont plus performantes que la plupart des méthodes à base de classifieurs [18], [124]. Nous allons aborder les approches par combinaison dans le paragraphe 4.5, tandis que les approches par classification des scores seront introduites dans le paragraphe 4.6.

4.4 La normalisation de scores

La normalisation de scores est une étape essentielle, elle consiste à rendre les scores de correspondance en sortie des *matcher*s homogènes avant de les combiner et cela s'effectue par le changement de leurs paramètres de position (moyenne) et d'échelle (écart-type). Trois problèmes de fusion de scores imposent l'utilisation de la normalisation :

- Les scores de chaque matcher peuvent être de nature différente, par exemple un matcher peut produit en sortie une mesure de proximité (similarité) tandis que l'autre matcher peut produit en sortie une mesure de distance (dissimilarité).
- Les sorties de chaque matcher ne sont pas nécessairement inclues dans le même intervalle de variation des scores, par exemple les scores d'un matcher varient entre 0 et 1 pendant que les scores d'un autre matcher varient entre 0 et 100.
- Les scores retournés par les matchers peuvent suivre différentes distributions statistiques.

4.4.1 Les techniques de normalisation de scores

Nous pouvons distinguer plusieurs techniques de normalisation de scores. Deux caractéristiques doivent être prises en compte lors du choix de la technique à adopter :

- ✓ *La robustesse* : se réfère à l'insensibilité à la présence de valeurs aberrantes (*outliers*).

- ✓ *L'efficacité* : se réfère à la proximité de la distribution des transformées par rapport à la distribution originale des scores.

Une explication de ces deux concepts (*robustesse* et *efficacité)* est effectuée par *Huber* dans [125].

4.4.1.1 La normalisation Min-Max

La normalisation Min-Max est la technique de normalisation la plus simple, et la plus adaptée lorsque les bornes (valeurs minimales et valeurs maximales) des scores sont connues. Dans ce cas on translate les scores minimum et maximum respectivement vers 0 et 1. Si les scores minimum et maximum ne sont pas connus nous pouvons les estimer à partir d'un jeu d'entraînement de score donné, cependant la technique reste valable mais pas robuste (sensible aux valeurs aberrantes dans les données utilisées pour l'estimation). Le score normalisé Min-Max est donné par :

$$s'_i = \frac{s_i - \min_i(s_i)}{\max_i(s_i) - \min_i(s_i)} \tag{4.1}$$

Où $\min_i$ et $\max_i$ sont déterminés pour chaque matcher, s_i représente les scores de correspondance de sortie de chaque sous système. La normalisation Min-Max conserve la distribution des scores originale à un facteur d'échelle prés. La figure 4.2 Illustre la normalisation des scores par la méthode Min-Max.

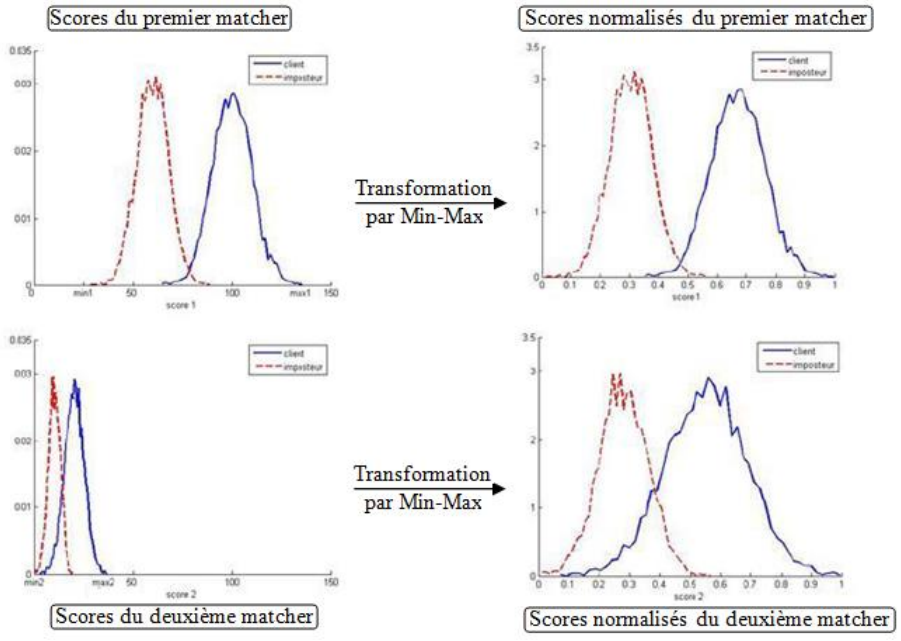


Figure 4.2. La normalisation des scores par la technique Min-Max.

Nous constatons que les scores de chaque matcher se répartissent dans un intervalle [0, 1], les clients sont proches de 1 et les imposteurs sont proches de 0.

4.4.1.2 La normalisation Decimal scaling

La méthode *decimal scaling* peut être appliquée lorsque les scores de différents *matchers* évoluent selon une loi logarithmique. Par exemple si un *matcher* produit des scores appartiennent à l'intervalle [0, 1] et un autre matcher produit des scores appartiennent à l'intervalle [0, 100], la normalisation ci-dessous peut être utilisée.

$$s'_i = \frac{s_i}{10^n} \tag{4.2}$$

Où $n = \log_{10} \max(s_i)$. Les problèmes de cette technique sont le manque de robustesse ainsi qu'elle suppose que les scores de différents matchers évoluent selon une échelle logarithmique.

4.4.1.3 La normalisation Z-Score

La technique de normalisation Z-Score est la plus employée, elle utilise la moyenne et l'écart-type de la distribution des scores de chaque *matcher*. Ces deux paramètres peuvent être déduits de l'algorithme de connaissance. Si on n'a pas une information a priori sur la nature de ce dernier, on estime la moyenne et l'écart-type à partir d'un jeu de scores de correspondance donné. La normalisation Z-Score s'effectue par :

$$s'_i = \frac{s_i - \mu}{\sigma} \tag{4.3}$$

Où μ est la moyenne de la distribution calculée ou estimée et σ est l'écart-type des données. La normalisation Z-Score n'est pas robuste, en fait la moyenne et l'écart-type sont sensibles aux valeurs aberrantes. De plus elle ne conserve pas la distribution originale sauf si cette dernière est gaussienne, donc pour une distribution arbitraire les paramètres de position et d'échelle sont raisonnables mais pas optimaux.

4.4.1.4 La normalisation MAD (*Median Absolute Deviation*)

Les méthodes basées sur la médiane et l'écart-type absolu ne sont pas sensibles aux valeurs aberrantes et aux points aux extrémités d'une distribution. La normalisation des scores s'effectue par :

$$s'_i = \frac{s_i - \text{median}}{\text{MAD}} \tag{4.4}$$

Où MAD = median ($|s_i - \text{median}(s_i)|$). Cependant lorsque la distribution de scores n'est pas gaussienne la médiane et l'écart-type absolu médian sont des estimateurs pauvres en

paramètres de position et d'échelle. Ainsi cette technique ne conserve pas la distribution originale.

4.4.1.5 La normalisation QLQ

Snelick et al. [78] utilisent la fonction quadratique-linéaire-quadratique (QLQ). Ils effectuent d'abord une normalisation Min-Max pour ramener les scores entre 0 et 1 (s_{MM}), puis ils appliquent la normalisation QLQ. Cette dernière prend comme paramètres le centre c et la largeur w de la zone de recouvrement des distributions des scores imposteurs et clients (figure 4.3).

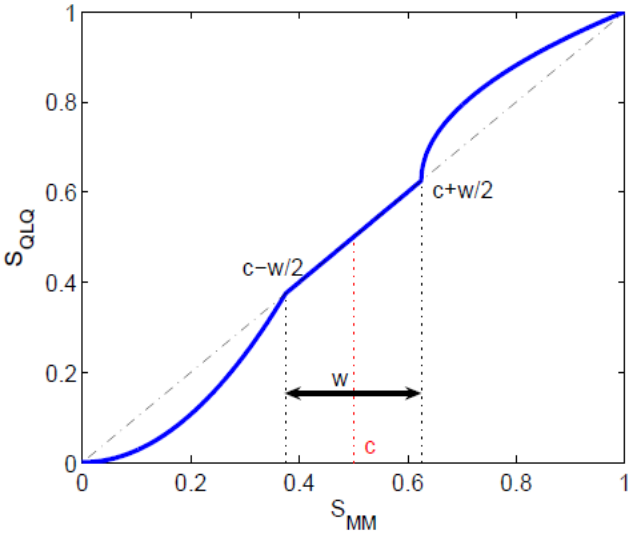


Figure 4.3. La Normalisation QLQ.

La zone de recouvrement des distributions des scores imposteurs et clients reste inchangée (fonction linéaires) tandis que les deux autres régions sont transformées à l'aide de fonctions quadratiques par :

$$s_{QLQ} = \begin{cases} \dfrac{1}{\left(c - \frac{w}{2}\right)} s_{MM}^2, & si\ s_{MM} \leq \left(c - \dfrac{w}{2}\right) \\ s, & si\ \left(c - \dfrac{w}{2}\right) < s_{MM} \leq \left(c + \dfrac{w}{2}\right) \\ \left(c + \dfrac{w}{2}\right) + \sqrt{\left(1 - c - \dfrac{w}{2}\right)\left(s_{MM} - c - \dfrac{w}{2}\right)}, & sinon \end{cases} \quad (4.5)$$

4.4.1.6 La normalisation double sigmoïde

Cappelli et al. [126] utilisent une fonction double sigmoïde pour normaliser les scores de différents matchers. Le score normalisé est donné par :

$$s'_i = \begin{cases} \dfrac{1}{1+\exp\left(-2\left(\frac{s_i - t}{r1}\right)\right)} & si\ s_i < t, \\ \dfrac{1}{1+\exp\left(-2\left(\frac{s_i - t}{r2}\right)\right)} & sinon \end{cases} \quad (4.6)$$

Où t est le point de référence et r_1 et r_2 sont les paramètres permettant de définir deux fonctions sigmoïdes. Ces fonctions montrent des caractéristiques linéaires dans les régions [t-r_1, t] et [t, t+r_2]. La figure 4.4 présente un exemple d'une normalisation double sigmoïde, où les scores inclus dans l'intervalle [0, 300] sont transformés dans l'intervalle [0, 1], avec t = 200, r_1 = 20 et r_2 = 30.

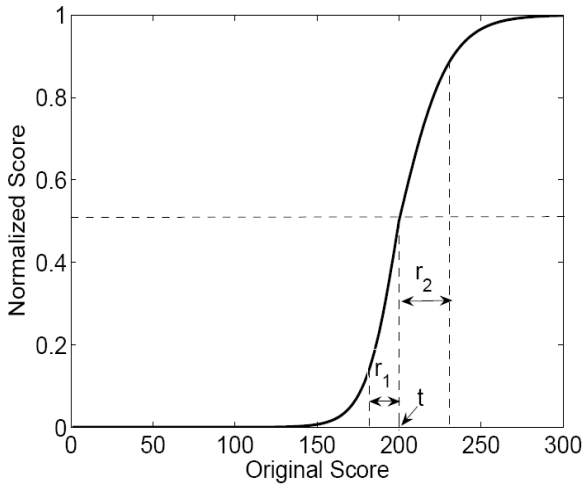


Figure 4.4. La Normalisation double sigmoïde.

Les paramètres t, r_1 et r_2 doivent être soigneusement choisis pour obtenir une bonne efficacité. En effet t est choisi de façon à être une valeur appartenant à la zone de recouvrement des imposteurs et des clients, et les paramètres r_1 et r_2 sont choisis comme bords gauche et droit des distributions des imposteurs et des clients (autrement dit les valeurs de r_1 et r_2 correspondent respectivement au minimum des scores de similarité des clients et au maximum des scores de similarité des imposteurs). Ainsi les scores appartenant à la région de recouvrement sont transformés de manière linéaire, tandis que les scores à l'extérieur de cette région sont transformés de manière non-linéaire. La normalisation double sigmoïde ne peut être utilisée dans le cas où il y'a plusieurs intervalles de zones de recouvrement des imposteurs et des clients.

4.4.1.7 La normalisation tanh (*tangente hyperbolique*)

La méthode tanh introduit par *Hampel* est robuste et très efficace [127]. La normalisation des scores est donnée par :

$$s'_i = \frac{1}{2}\left\{\tanh\left(0.01\left(\frac{s_i - \mu_{GH}}{\sigma_{GH}}\right)\right) + 1\right\} \tag{4.7}$$

Où μ_{GH} et σ_{GH} sont respectivement la moyenne et l'écart-type de la distribution des scores clients, tels qu'ils sont donnés par les estimateurs de Hampel. Ces derniers se basent sur la fonction d'influence (ψ) ci-dessous :

$$\psi(u) = \begin{cases} u & 0 \le |u| < a, \\ a * sign(u) & a \le |u| < b, \\ a * sign(u) * \left(\dfrac{c - |u|}{c - b}\right) & b \le |u| < c, \\ 0 & |u| \ge c, \end{cases} \tag{4.8}$$

La méthode tanh est peu sensible aux valeurs aberrantes car la fonction de Hampel réduit l'influence des points aux bords d'une distribution (identifiés par a, b et c) pendant l'estimation des paramètres de position et d'échelle. En fait si plusieurs points constituant un bord d'une distribution ne sont pas pris en compte la méthode sera robuste mais pas efficace (optimal). Si tous les points sont considérés la méthode sera efficace mais pas robuste, ainsi le choix des paramètres a, b, et c dépend de la quantité de robustesse exigée. La figure 4.5 illustre la fonction d'influence de Hampel.

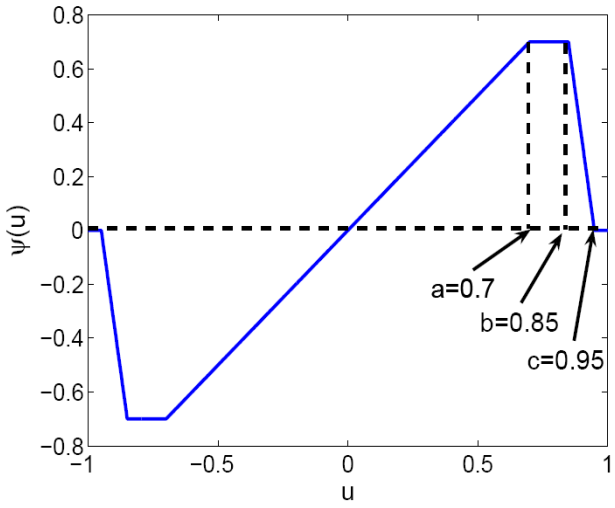


Figure 4.5. Fonction d'influence de Hampel (a = 0.7, b = 0.85 et c = 0.95).

Nous constatons que les scores clients s_a doivent être d'abord transformés dans l'intervalle [−1, 1], cela peut être effectué par la normalisation ci-dessous (proche de la normalisation Min-Max :

$$s'_a = \frac{2 * s_a - (\max(s_a) + \min(s_a))}{\max(s_a) - \min(s_a)} \tag{4.9}$$

La table 4.1 répertorie les différentes méthodes de normalisation que nous avons présenté ci-dessus ainsi que leurs deux principales caractéristiques : "*la robustesse*" et "*l'efficacité*".

Technique de Normalisation	Robustesse	Efficacités
Min-Max	Non	Élevée
Decimal Scaling	Non	Élevée
Z-Score	Non	Elevée (optimale pour des données gaussiennes)
Médiane et MAD	Oui	Modérée
QLQ	Oui	Élevée
Double Sigmoïde	Oui	Élevée
Estimateurs tanh	Oui	Élevée

Table 4.1. Résumé des techniques de normalisation de scores.

4.5 Les méthodes de combinaison de scores

Les méthodes de combinaison de scores sont des méthodes simples permettant de produire un score final c à partir des scores s_i de R matchers. Les méthodes les plus utilisés sont le produit, la somme, le maximum, le minimum et la médiane [128].

✓ ***La règle produit*** ("*product rule*") : La règle produit consiste à calculer un nouveau score en faisant le produit des scores de chaque matcher :

$$c = \prod_{i=1}^{R} s_i \tag{4.10}$$

La méthode produit est susceptible aux valeurs aberrantes si la normalisation adoptée n'est pas robuste. Ainsi que la normalisation des scores entre les valeurs 0 et 1 peut poser un problème, en effet nous pouvons obtenir un score final égale à 0 si un des scores normalisés en entrée égale à 0.

✓ ***La règle* somme** ("*sum rule*") : La règle somme consiste à calculer un nouveau score en faisant la somme des scores de chaque matcher :

$$c = \sum_{i=1}^{R} s_i \tag{4.11}$$

La méthode somme est plus efficace que la méthode produit car elle est robuste aux valeurs aberrantes et au bruit. De plus le score normalisé à 0 n'annule pas les autres scores comme c'est le cas de la méthode produit.

✓ ***La règle* maximum** ("*max rule*") : La règle maximum consiste à définir un nouveau score comme étant le score maximal des scores de chaque matcher :

$$c = \max_i (s_i) \tag{4.12}$$

La méthode maximum n'est pas robuste aux valeurs aberrantes.

- ✓ ***La règle* minimum** ("*min rule*") : La règle minimum consiste à définir un nouveau score comme étant le score minimal des scores de chaque matcher :

$$c = \min_i (s_i) \tag{4.13}$$

La méthode minimum n'est pas robuste aux valeurs aberrantes.

- ✓ ***La règle* somme pondérée** : À la différence des méthodes de combinaison ci-dessus la méthode somme pondérée nécessite un réglage de paramètres. Jain et Ross [129] ont proposé l'utilisation de poids w_i spécifiques à chaque utilisateur pour réaliser la somme pondérée de scores provenant de différentes modalités :

$$c = \sum_{i=1}^{R} w_i \, s_i \tag{4.14}$$

L'idée de cette méthode est que certains individus ont certains traits biométriques de moins bonne qualité que ceux d'autres individus. Par exemple, certaines personnes peuvent avoir des empreintes digitales endommagées à cause d'un travail manuel prolongé. Pour de tels utilisateurs on affecte un faible poids au score d'empreinte digitale et un poids plus élevé aux scores d'autres modalités. Cela va réduire les probabilités de faux rejet, cependant cette méthode exige un apprentissage de poids spécifiques pour chaque utilisateur.

4.6 Les méthodes de classification de scores

Les méthodes de fusion à base de classifieurs cherchent à prendre une décision à partir du vecteur de scores des matchers disponibles. Plusieurs classifieurs peuvent être utilisés.

Sanderson et al. [113] ont utilisé un classifieur basé sur une Machine à Vecteurs de Support (SVM) afin de combiner les scores provenant de système de reconnaissance faciale et de système de reconnaissance de la parole. Ils montrent que la performance de ce classifieur se dégrade en la présence de bruit. Afin de surmonter ce problème ils implémentent des classifieurs résistants au bruit tel que le classifieur linéaire défini par morceau (*piece-wise linear classifier*).

Wang et al. [130] utilisent l'analyse discriminante linéaire de Fisher (LDA) et un classifieur de réseau de neurones associé à une fonction de base radiale (RBF) pour combiner les scores issus de système de reconnaissance faciale et de système de reconnaissance de l'iris.

Ross et Jain [18] ont utilisé un arbre décisionnel et des classifieurs discriminants linéaires pour combiner les scores issus des systèmes de reconnaissance faciale, d'empreinte digitale et de géométrie de la main.

4.7 Conclusion

Dans ce chapitre, nous avons introduit la fusion de modalités, nous avons présenté quatre niveaux de fusion : la fusion au niveau du *capteur*, la fusion au niveau des *caractéristique*s, la fusion au niveau des *décisions* et la fusion au niveau des *scores*. Cette dernière est la technique la plus courante étant donnée sa simplicité d'implémentation et sa grande flexibilité. Nous avons présenté ensuite deux approches de fusions des scores, l'approche à base de classifieurs, et l'approche de combinaison de scores qui est plus performante que la première, cependant elle nécessite une homogénéité des scores provenant de chaque matcher, cela nous amène à étudier les méthodes de normalisation des scores les plus utilisées.

Chapitre 5

Résultats expérimentaux

5.1 Introduction

Dans ce chapitre nous allons évaluer les performances des systèmes unimodaux de la reconnaissance faciale et de la reconnaissance d'empreintes digitales en fonction des algorithmes que nous avons détaillés durant ce mémoire, ainsi que la performance de système multimodal obtenu en associant les informations issues de deux sources biométriques au niveau des scores.

5.2 La base de données multimodale

L'évaluation des algorithmes que nous allons utiliser dans nos expérimentations nécessite la création d'une base de donnée multimodale de N individus virtuels ayant deux modalités différentes en associant de manière aléatoire à chaque individus des images de chaque modalité c.-à-d. pour construire notre base multimodale combinant visages et empreintes digitales nous utilisons une base de données A contient les visages des individus et autre base B contient les empreintes digitales. Nous créons ensuite les identités "virtuelles" en associant par exemple le visage A_1 de la première base avec l'empreinte B_1 de la deuxième base, et le visage A_2 de la première base avec l'empreinte B_2 de la deuxième base, etc. [131], [132].

5.2.1 La base ORL

Pour la base de visage nous avons choisis la base ORL, base qui a été recueillie entre avril 1992 et avril 1994 par un laboratoire de AT&T de Cambridge. Cette base contient 400 images, représentant les visages de 40 personnes, chacune dispose 10 vues différentes, avec

des changements de pose, d'éclairage, d'expressions faciales (expression neutre, sourire et yeux fermés) et des occultations partielles par les lunettes mais toujours sur un fond foncé. Ces images ont été collectées à des dates différentes, elles sont de taille 112 × 92 pixels. L'annexe B.1 présente des échantillons de la base ORL.

5.2.2 La base FVC

Pour la base des empreintes digitales nous avons utilisé la base FVC2002 (Fingerprint Vérification Competition), base qui a été collectée par l'université de Bologne. Elle comporte 800 images, représentant les empreintes de 100 personnes, en faisant 8 occurrences du même doigt. Nous avons choisis la base BD1 recueillie par un capteur capacitif à bas prix. Les images sont de taille 388 × 374 pixels et de résolution 500 DPI (pour Dots Per Inch ou Points Par Pouce). L'annexe B.2 représente des échantillons de la base FVC.

À partir des bases de données ORL et FVC nous avons élaboré un corpus de 640 images (2 modalités × 40 personnes × 8 images pour chaque personnes). Pour chaque modalité on a pris 90 *images d'apprentissage* (30 individus avec 3 photos par individus), 150 *images de test* (30 individus avec 5 photos par individus), et 80 *images d'inconnus* (10 individus avec 8 photos par individus).

5.3 Interface graphique

Nous avons conçu notre interface graphique utilisateur (GUI) "Face And Fingerprint Recognition System" (voir la figure 5.1) sous le langage Matlab 7.5.0.342 (R2007b) avec l'outil GUIDE. Nous avons utilisé un ordinateur SONY-VAIO ayant les caractéristiques suivantes :

- Processeur : Intel(R) Core(TM) i3-2330M CPU @ 2.20GH_Z 2.20 GH_Z.
- RAM : 4,00 Go.
- OS : Microsoft Windows Seven.

Notre interface est simple, elle permet d'illustrer le processus de la reconnaissance faciale et celui de la reconnaissance de l'empreintes digitales dès la phase de l'enrôlement de trait biométrique jusqu'à l'affichage des scores de similarité de l'individu identifié avec la base de donnée. Cette interface présente aussi la fusion multimodale et l'évaluation des trois systèmes biométriques, c'est ainsi qu'elle est divisée en trois volets à savoir :

- Face Process : volet consacré à la reconnaissance faciale, il comporte :
  - ✓ Base Enrollment : ce bouton permet d'acquérir la base de visages, d'extraire les caractéristiques de chaque image de cette base et de stocker les résultats.
  - ✓ Mean Face : calcule le visage moyen de la base de données.

- ✓ First Eigenfaces : affiche les vingt premiers vecteurs propres correspondant aux vingt plus grandes valeurs propres.
- ✓ Eigenfaces Variance : ce bouton permet de connaitre le nombre des vecteurs propres à retenir, il représente le pourcentage de la variance de l'espace des vecteurs propres.
- ✓ Recognition panel : cette partie est dédiée à la reconnaissance des visages en utilisant l'analyse en composantes principales. Un prétraitement par l'égalisation d'histogramme est d'abord effectué, cela permet de répartir uniformément les niveaux de gris des images de visages en améliorant leurs contrastes. Cette partie regroupe aussi les boutons suivant :
  - Reconstructed Face : bouton qui visualise la qualité du visage reconstruit à partir du visage moyen et des vecteurs propres de la base de données.
  - Save : bouton qui permet de stocker les caractéristiques du visage identifié.
  - Match : ce bouton visualise le visage de la base de données qui ressemble le plus à l'individu traité (pour effectuer la comparaison nous avons opté pour la distance cosinus décrite au point 2.7).
  - Cosine Distance : ce bouton affiche les différentes distances cosinus entre la projection du visage à identifier et les projections des visages de la base de données.

➢ Fingerprint Process : partie dédiée à la reconnaissance des empreintes digitales, elle comporte :

- ✓ Base Enrollment : ce bouton permet d'extraire et de stocker les caractéristiques des images de la base des empreintes digitales (la base FVC).
- ✓ Recognition panel : cette partie illustre la reconnaissance des empreintes digitales. Pour l'extraction des minuties, nous avons choisi la méthode basée sur la binarisation. Pour le prétraitement des images, nous avons utilisé la transformée de Fourier ainsi que le filtrage de Gabor, cela améliore le taux d'identification de notre système. Cette partie comporte, en outre, les boutons suivants :
  - Binarization : ce bouton convertie l'image de l'empreinte digitale en une image binaire.
  - Thining : bouton qui permet d'amincir l'empreinte digitale binarisée.
  - Munitia Extraction : ce bouton permet d'extraire les minuties de l'image binaire squelettisée de l'empreinte digitale.
  - Remove False Minutia : élimine le maximum des fausses minuties détectées au cours des étapes de binarisation et de squelettisation.

- Save : bouton qui permet de stocker les caractéristiques de l'empreinte identifiée.
- Match : ce bouton visualise l'empreinte de la base de données qui ressemble le plus à l'empreinte identifiée.
- Percent_Match : affiche les scores de similarité (pourcentages des minuties matchées) de l'empreinte identifiée avec la base de données.

➢ Score Fusion : ce volet est voué à la fusion de deux modalités biométriques et plus précisément à la fusion au niveau des scores. Nous avons opté pour la règle "somme pondérée", c'est ainsi que *Face Weight* et *Finger Weight* représentent respectivement les poids spécifiques au visage et à l'empreinte digitales de la personne à identifier choisis par l'utilisateur. Ce volet comporte :

✓ Fusion : ce bouton permet de visualiser le résultat de la fusion (image de visage et celle de l'empreinte digitale de l'individu que le système a reconnu) en utilisant les cinq méthodes de normalisation (la méthode de normalisation Min-Max, Z-Score, Median Absolute Deviation, QLQ et la normalisation double sigmoïde).

✓ Similarity score : affiche les scores de similarité de la personne à identifier avec la base de données pour chaque méthode de normalisation citée ci-dessus.

✓ System evaluation panel : cette partie sert à évaluer la performance du système de la reconnaissance faciale, du système de reconnaissance d'empreintes digitales et du système multimodal résultant de la fusion des scores de ces deux modalités en utilisant les cinq méthodes de normalisation susmentionnées, cette partie comprend :

- FAR&FRR Curve : ce bouton sert à tracer les courbes FAR vs FRR du système de la reconnaissance faciale, du système de la reconnaissance d'empreintes digitales et du système multimodal.
- Scores distributions : permet d'afficher la distribution des scores des personnes légitimes et des imposteurs des trois systèmes.
- DET Curves : trace les courbes DET des trois systèmes.

Des barres situées dans la partie inférieure de chaque volet permettent d'afficher les informations concernant chaque système comme le temps d'exécution, les valeurs de EER, elle dirigent aussi l'utilisateur de l'interface en lui indiquant ce que doit faire.

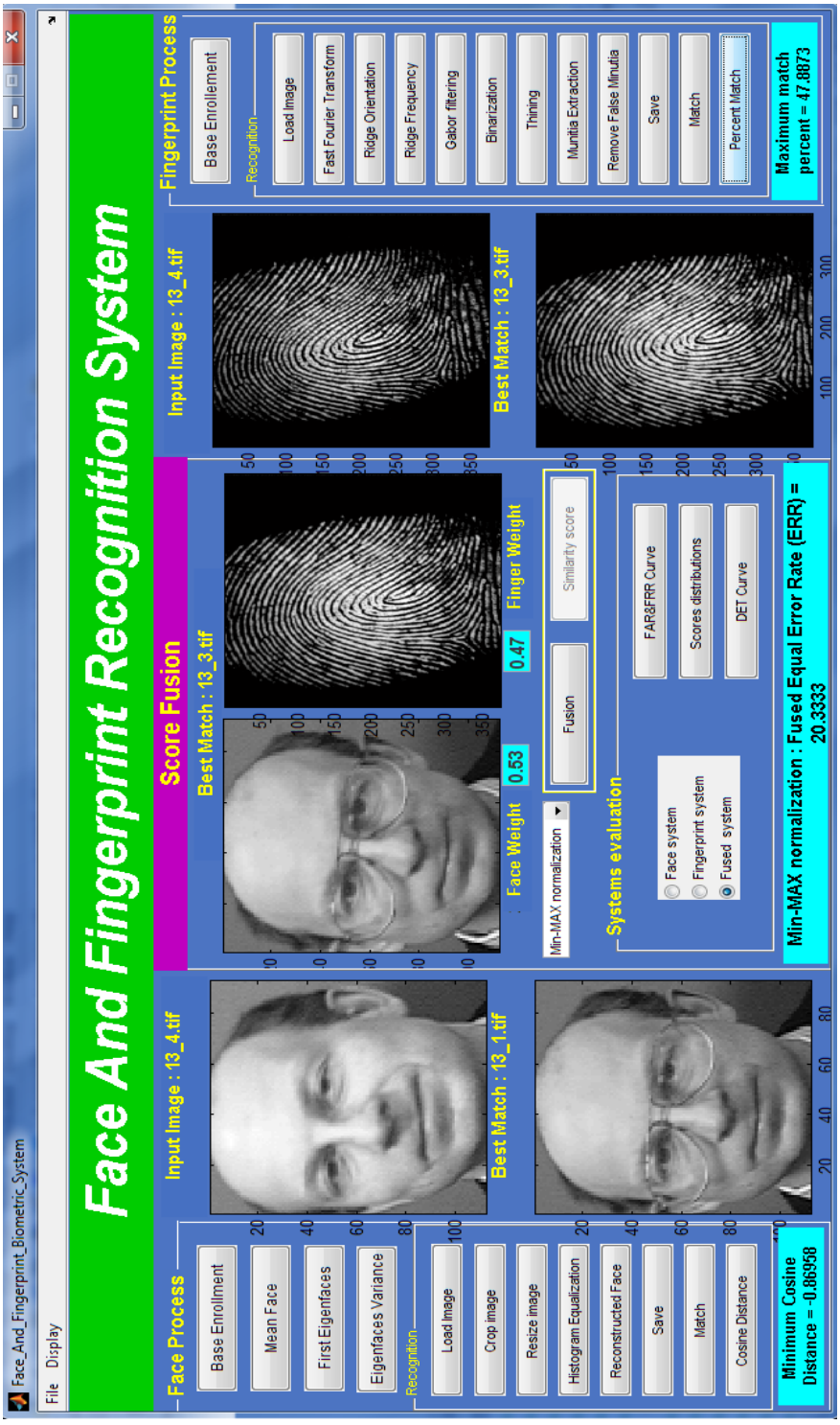


Figure 5.1. Interface graphique de la reconnaissance multimodale (visage et empreinte digitale).

5.4 Amélioration des images de test

La robustesse d'un système de reconnaissance est étroitement liée à la qualité des images utilisées. Dans la plupart des cas, ces images sont sujettes au bruit, elles sont floues et peu contrastées. Ceci est dû à plusieurs facteurs comme la qualité de capteurs utilisés. C'est ainsi qu'une étape de prétraitement, qui vise à minimiser voire supprimer les imperfections et les artéfacts des images afin de les rendre exploitables par la suite est nécessaire. En fait, plusieurs méthodes d'amélioration des images peuvent être distinguées, dans notre travail nous avons opté pour l'égalisation d'histogramme afin de normaliser les images de visages et la transformée de Fourier ainsi que le filtrage de Gabor pour améliorer les images des empreintes digitales.

5.4.1 Amélioration des images de visages

Comme nous l'avons indiqué ci-dessus, pour améliorer les images de visages nous utilisons l'égalisation d'histogramme, qui est une transformation ponctuelle d'intensité, consiste à rendre l'histogramme le plus plat possible, de façon à harmoniser la répartition des niveaux de gris afin d'augmenter les nuances dans l'image et donc son contraste et sa luminosité globale (figure 5.2).

Plus concrètement, soit la probabilité d'occurrence d'un pixel de niveau x_k dans l'image est :

$$P(x_k) = \frac{n_k}{n}, \qquad 0 \leq k < L \tag{5,1}$$

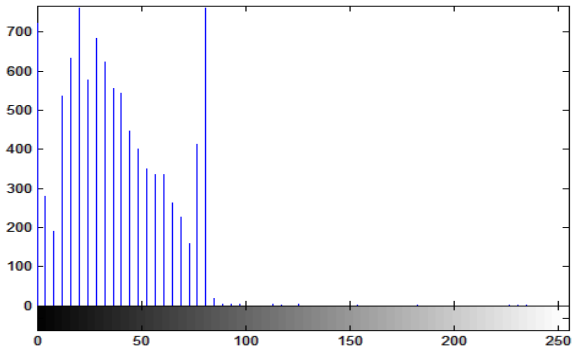
Avec n_k le nombre de pixels à niveau k de gris, L le nombre des niveaux de gris, n le nombre total de pixels de l'image et p définit alors l'histogramme normalisé sur [0,1]. Soit C la distribution cumulative de l'histogramme normalisé :

$$C(k) = \sum_{j=0}^{k} P(x_k) \tag{5,2}$$

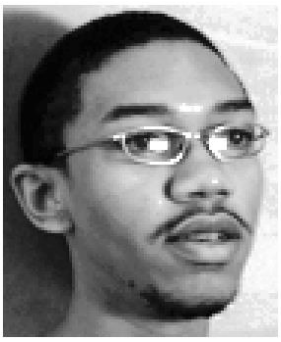
Nous cherchons une transformation Y = T(x) qui produit un niveau y pour chaque niveau x de l'image de telle façon que la distribution cumulative des différents niveaux de l'image transformée soit linéaire, cette transformation est donnée par :

$$Y_K = T(x_K) = L\,\frac{C(k)}{n} \tag{5.3}$$

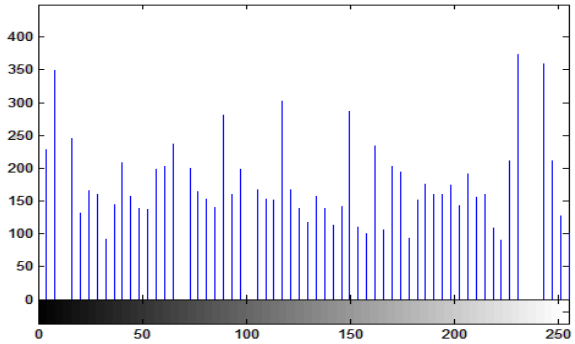
(a) Image en niveau gris



(b) Histogramme de l'image en niveaux gris



(c) Image améliorée



(b) Histogramme de l'image améliorée

Figure 5.2. Égalisation de l'histogramme.

5.4.2 Amélioration des images des empreintes digitales

Pour améliorer les images des empreintes digitales nous avons d'abord adopté l'algorithme proposé par Waston et al [133], qui utilisent l'égalisation d'histogramme puis ils appliquent la Transformée de Fourier Rapide (ou FFT pour Fast Fourier Transform) à chaque bloc 32 × 32 de l'image de l'empreinte comme suit :

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \times exp\left\{-j2\pi \times \left(\frac{ux}{M} + \frac{vy}{N}\right)\right\} \tag{5.4}$$

Avec u = 0, 1, 2, …, 31, et v = 0, 1, 2, …, 31.

On multiplie ensuite la FFT de chaque bloc par la puissance de son module, puis on calcule sa transformée inverse :

$$g(x,y) = F^{-1}\{F(u,v) \times |F(u,v)|^k\} \tag{5.5}$$

Où $F^{-1}(F(u, v))$ est définit par :

$$f(x,y) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v) \times \exp\left\{j2\pi \times \left(\frac{ux}{M} + \frac{vy}{N}\right)\right\} \quad (5.6)$$

Avec x = 0, 1, 2, …, 31, et y = 0, 1, 2, …, 31.

Le coefficient k dans la formule (5,5) est une constante déterminée expérimentalement. les hautes valeurs peuvent améliorer l'image en remplissant les petits trous des stries brisées, cependant les valeurs fortement élevées de cette constante peuvent engendrer des fausses bifurcations en joignant deux terminaisons. Nous avons choisis k = 0.10.

Ensuite nous avons appliqué l'algorithme proposé par Lin Hong et al [134] qui utilisent le filtrage de Gabor. Cet algorithme se compose cinq étapes (voir la figure 5.3).

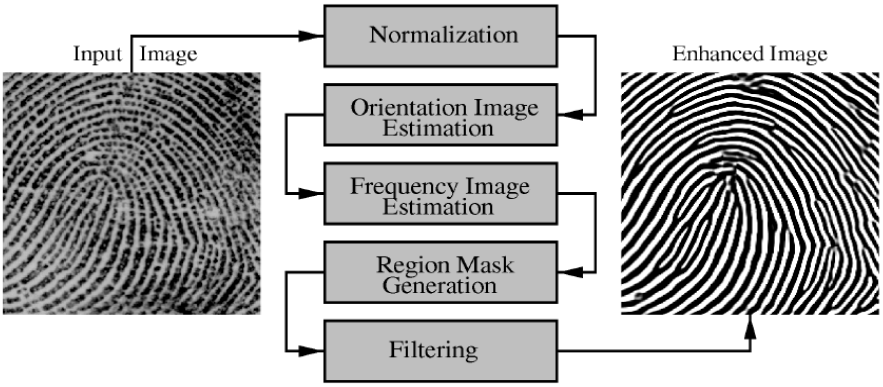


Figure 5.3. Processus d'amélioration de la qualité d'empreinte digitale [134].

- La normalisation : cette étape vise à réduire la variation de niveau de gris entre les crêtes et les vallées, elle ne change pas la clarté de la structure de l'empreinte digitale.

  Soit $I_{(N\times N)}$ l'image de l'empreinte digitale en niveau de gris. Sa moyenne est définit par :

$$M(I) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} I(i,j) \quad (5,7)$$

  Et sa variance est définie par :

$$VAR(I) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} (I(i,j) - M(I))^2 \quad (5,8)$$

  L'image normalisée Γ est définie par :

$$I(i,j) = \begin{cases} M_0 + \sqrt{\dfrac{VAR_0(I(i,j)-M)^2}{VAR}}, & si\ \ I(i,j) > M \\ M_0 - \sqrt{\dfrac{VAR_0(I(i,j)-M)^2}{VAR}}, & sinon \end{cases} \quad (5,9)$$

Où M_0 et VAR_0 sont respectivement le moyen et la variance désirés.

- L'orientation de l'image : nous calculons les gradients G_x et G_y pour chaque pixel en utilisant la convolution par le dérivé du filtre de Gauss (de taille 7×7). L'orientation dominante de chaque bloc w × w est ensuite estimée par l'équation suivante :

$$\theta_{ij} = \frac{\pi}{2} + \frac{1}{2}\tan 2(2G_{xy}, G_{xx} - G_{xy}) \quad (5,10)$$

Où
$$G_{xy} = \sum_w G_x G_y \quad (5,11)$$

$$G_{xx} = \sum_w {G_x}^2 \quad (5,12)$$

Et
$$G_{yy} = \sum_w {G_y}^2 \quad (5,13)$$

Cependant l'orientation obtenue peut contenir des erreurs à cause de bruit qui altère la structure des stries, cette orientation peut être adoucie par un filtre passe-bas. Nous avons choisis un noyau Gaussien pour lisser (smooth) chaque bloc :

$$\varphi_x(i,j) = \sum_{u=-\frac{w_\rho}{2}}^{\frac{w_\rho}{2}} \sum_{v=-\frac{w_\rho}{2}}^{\frac{w_\rho}{2}} F(u,v)\ \Phi_x\left(i - uw_\rho, j - vw_\rho\right) \quad (5,14)$$

$$\varphi_y(i,j) = \sum_{u=-\frac{w_\rho}{2}}^{\frac{w_\rho}{2}} \sum_{v=-\frac{w_\rho}{2}}^{\frac{w_\rho}{2}} F(u,v)\ \Phi_y\left(i - uw_\rho, j - vw_\rho\right) \quad (5,15)$$

Où
$$\Phi_x(i,j) = \cos(2\theta ij) \quad (5,16)$$

$$\Phi_y(i,j) = \sin(2\theta ij) \quad (5,17)$$

Et *F(u,v)* est le filtre de Gauss de taille $w_\rho \times w_\rho$. Ensuite nous calculons l'orientation par l'équation ci-dessous :

$$o(i,j) = \frac{1}{2}\tan^{-1}\left(\frac{\varphi_y(i,j)}{\varphi_x(i,j)}\right) \quad (5,18)$$

La figure suivante illustre l'orientation de l'image obtenue sans post-traitement et avec le lissage par le filtre de Gauss.

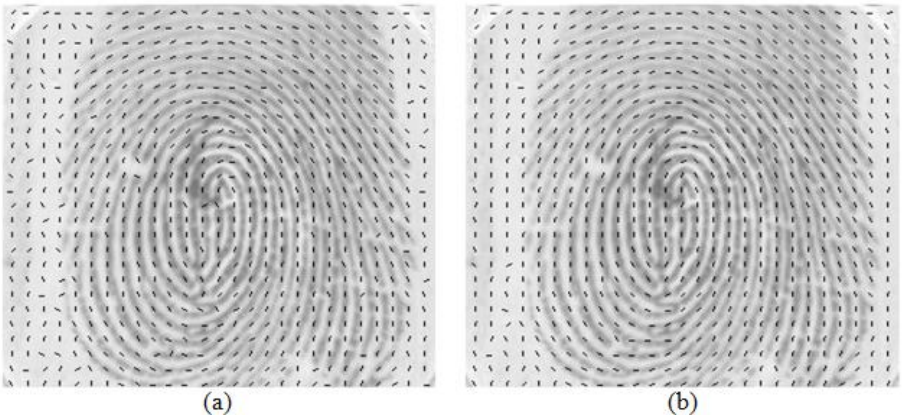


Figure 5.4. L'orientation de l'image, (a) sans lissage, (b) avec le lissage par un filtre de Gauss.

- L'image de fréquence : l'image de fréquence est une image $F_{(N\times N)}$ où F(i, j) représente la fréquence locale des stries en chaque pixel (i, j). Sur un voisinage local qui ne contient pas de points singuliers les crêtes et les vallées peuvent être modélisés par une onde sinusoïdale. Pour estimer la fréquence de l'image de l'empreinte digitale, on la divise en blocs. Nous calculons ensuite sa projection orthogonale à sa direction, on obtient un vecteur d'extrema, les maxima représentent les centres des stries et les minima correspondent aux centres des vallées. La période locale inter-strie D est la distance moyenne entre deux maxima consécutifs. S'il n'y a pas de minuties la fréquence F(i,j) est calculée par :

$$F(i,j) = \frac{1}{D} \tag{5,19}$$

- Filtre de Gabor : la fonction de Gabor est une oscillation harmonique composée d'une onde sinusoïdale avec une fréquence orientée modulée par une enveloppe Gaussienne. Le composant pair symétrique du filtre de Gabor est donné par :

$$H(x,y;\Omega,f) = \exp\left(-\frac{1}{2}\left[\frac{x_\Omega^2}{\delta_x^2}+\frac{y_\Omega^2}{\delta_y^2}\right]\right)\cos(2\pi f x_\Omega) \tag{5,20}$$

$$x_\Omega = x\cos\Omega + y\sin\Omega$$

$$y_\Omega = -x\sin\Omega + y\cos\Omega$$

Où Ω est l'orientation du filtre, f est la fréquence de la sinusoïde et δ_x et δ_y sont les constantes de l'enveloppe Gaussienne pour les axes x et y respectivement.
L'application du filtre de Gabor à l'image normalisée de l'empreinte digitale est obtenue par la convolution suivante :

$$E(i,j) = \sum_{u=-\frac{w_g}{2}}^{\frac{w_g}{2}} \sum_{v=-\frac{w_g}{2}}^{\frac{w_g}{2}} H\big(u, v, o\,(i,j), \mathrm{F(i,j)}\big)\, G\,(i-u, j-v) \qquad (5{,}21)$$

Avec w_g est la taille du filtre de Gabor.

La figure 5.5 illustre l'amélioration de l'image de l'empreinte digitale par la FFT et le filtre de Gabor.

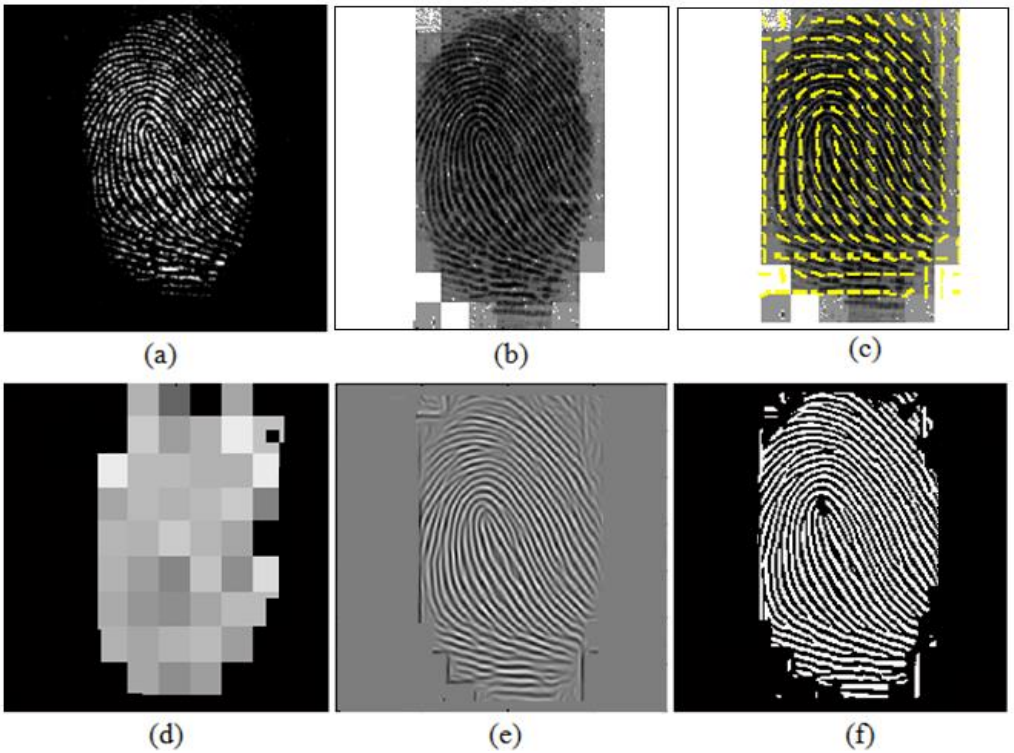


Figure 5.5. Amélioration de l'image de l'empreinte digitale, (a) image originale, (b) amélioration par FFT, (c) carte directionnelle, (d) carte fréquentielle, (e) amélioration par le filtre de Gabor, (f) image binarisée.

5.5 Evaluation du système biométrique

Pour évaluer la performance des systèmes biométriques nous allons tracer et étudier les courbes DET (décrit dans le chapitre 1) et FAR vs FRR :

La courbe FAR vs FRR :

Cette courbe représente les variations de taux de fausses acceptations et de taux de faux rejets par rapport au seuil de ressemblance (Threshold). Un exemple de courbe FAR vs FRR est illustrée à la figure 5.6.

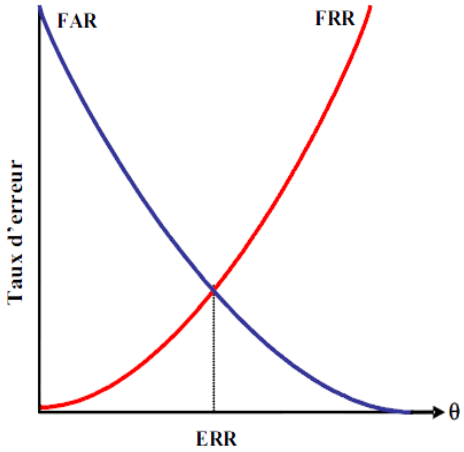


Figure 5.6. La courbe FAR vs FRR[1].

- *EER (Equal Error Rate)* : L'efficacité du système biométrique est mesurée en fonction de taux d'erreurs égales EER, le point où le taux de fausse acceptation est égal au taux de faux rejet. En pratique la distribution des scores n'est pas continue et le point d'intersection des courbes FAR et FRR peut ne pas exister (figure 5.7 (b et c)). Dans ce cas EER est calculé comme suit :

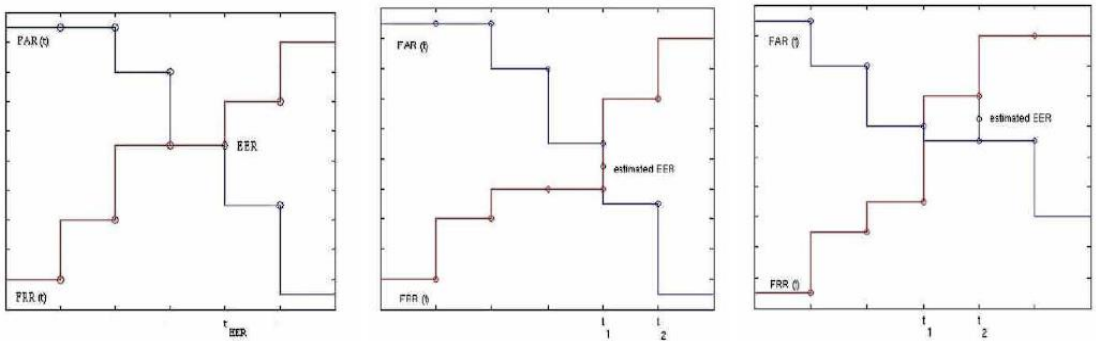
$$EER = \begin{cases} \dfrac{\mathrm{FAR}(t_1) + FRR(t_1)}{2} & ,si\ \ \mathrm{FAR}(t_1) - FRR(t_1) \leq\ \mathrm{FAR}(t_2) - FRR(t_2) \\ \dfrac{FAR(t_2) + FRR(t_2)}{2} & ,sinon \end{cases} \quad (5.22)$$

Où $t_1 = max_{t \in S} \{t / FRR(t) \leq FAR(t)\}$ (5.23)

Et $t_2 = min_{t \in S} \{t / FRR(t) \geq FAR(t)\}$ (5.24)

Et S est l'ensemble des seuils utilisés pour calculer la distribution des scores.

[1] *Mansfield, T. et al. (2001), "Biometric Product Testing Final Report".*



(a) Le point EER existe sur la courbe (b) EER est estimé à t1 (c) EER est estimé à t2

Figure 5.7. La courbe FAR vs FRR et le taux EER.

- *OP (Operating Point)* : pour assurer une haute sécurité, les systèmes biométriques opèrent avec un taux FAR très bas au lieu du EER. Dans ce cas, la performance du système est donnée par OP qui est défini en termes de pourcentage de FRR pour une valeur de FAR fixée. En pratique OP est calculé comme suit :

$$OP_{(FAR=\rho)} = FRR\left(t_{op}\right) \mid t_{op} = \max_{t \in S}\{t \mid \rho \leq FAR(t)\} \tag{5.25}$$

Où *S* est l'ensemble des seuils utilisés pour calculer la distribution des scores.

5.6 Résultats et discussion

Dans cette partie nous allons évaluer les performances des systèmes biométriques monomodaux (visage et empreinte digitale) et le système multimodal qui s'appuie sur la combinaison des scores de ces deux technologies biométriques. Nous allons tester la technique somme pondérée en affectant un poids $w_1 = 0.53$ à la modalité visage et un poids $w_2 = 1 - w_1 = 0.47$ à la modalité empreinte digitale. Nous rappelons qu'on a choisi :

- La méthode PCA pour la reconnaissance faciale et la distance cosinus pour comparer les visages à identifier avec ceux de la base de données.
- Et l'approche basée sur les minuties pour la reconnaissance des empreintes digitales. Deux méthodes sont utilisées pour améliorer la qualité des images des empreintes digitales à savoir la Transformée de Fourier Rapide et le filtrage de Gabor.

Les performances des systèmes de reconnaissance faciale, des empreintes digitales et du système multimodal résultant de la fusion des scores normalisées par les techniques Min-Max, Z-Score, MAD, QLQ et double sigmoïde sont visualisées par les courbes DET et FAR vs FRR suivantes :

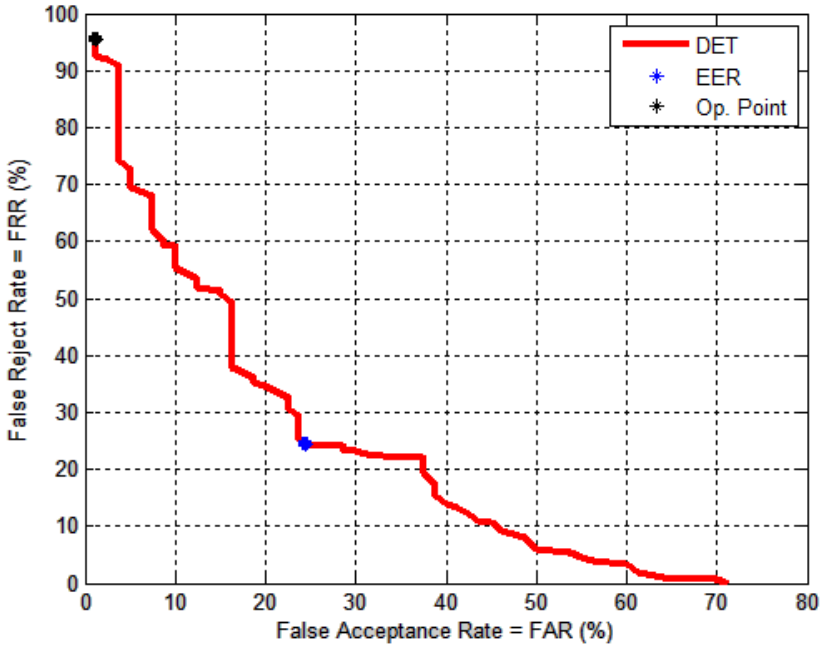


Figure 5.8. La courbe DET du système de la reconnaissance faciale.

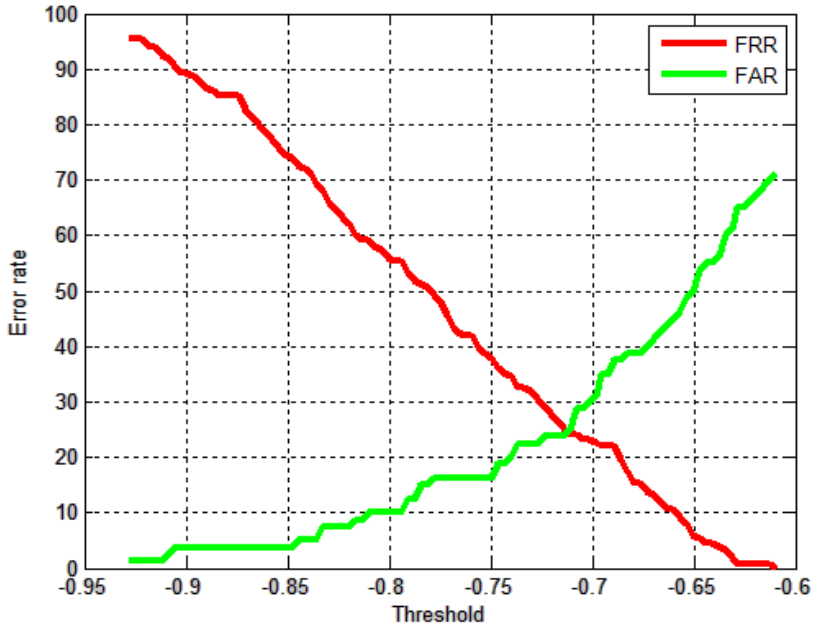


Figure 5.9. La courbe FAR vs FRR du système de la reconnaissance faciale.

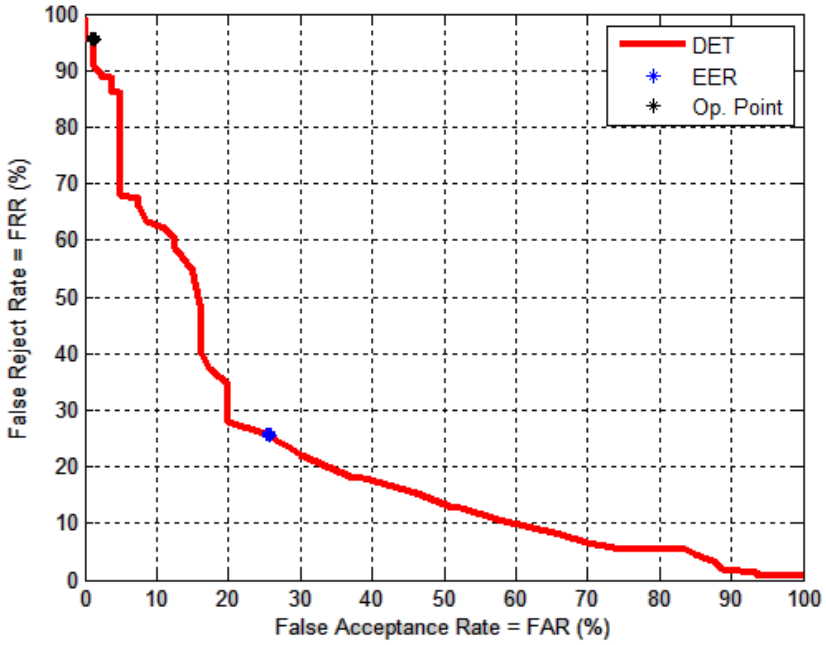


Figure 5.10. La courbe DET du système de la reconnaissance d'empreintes digitales.

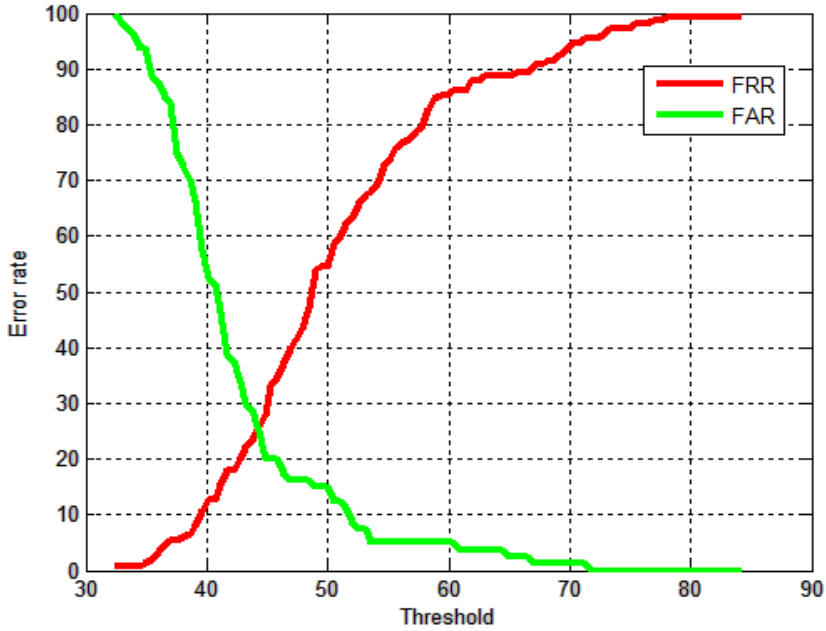


Figure 5.11. La courbe FAR vs FRR du système de la reconnaissance d'empreintes digitales.

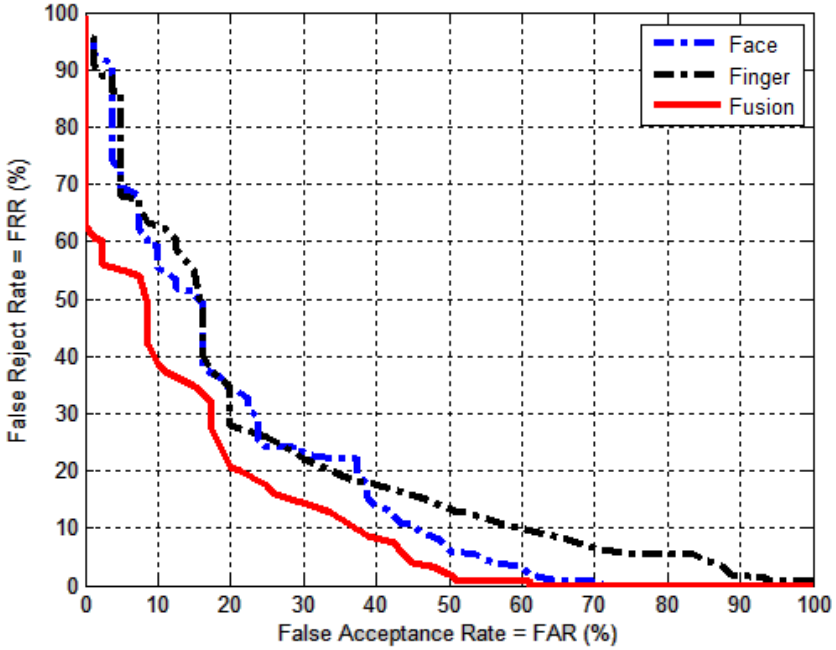


Figure 5.12. Courbes DET pour la fusion MinMax-somme pondérée.

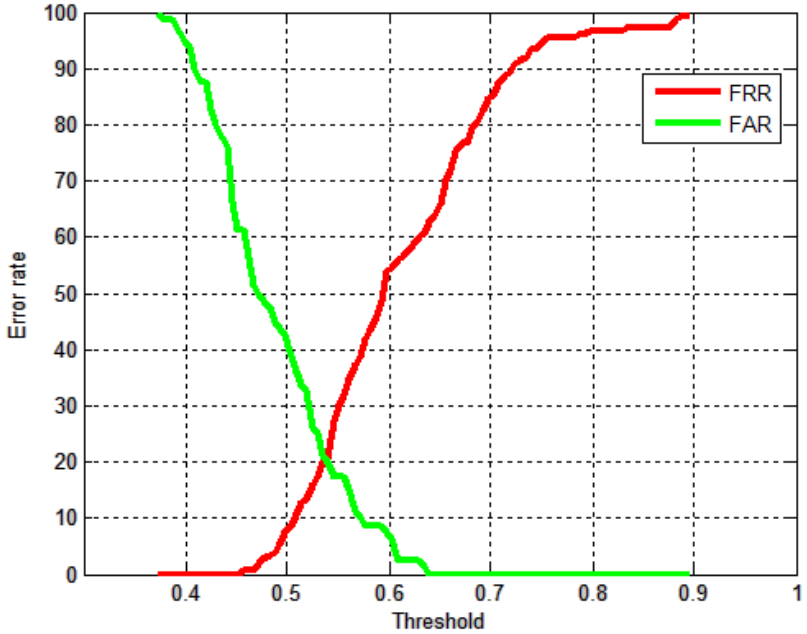


Figure 5.13. La courbe FAR vs FRR pour la fusion MinMax-somme pondérée.

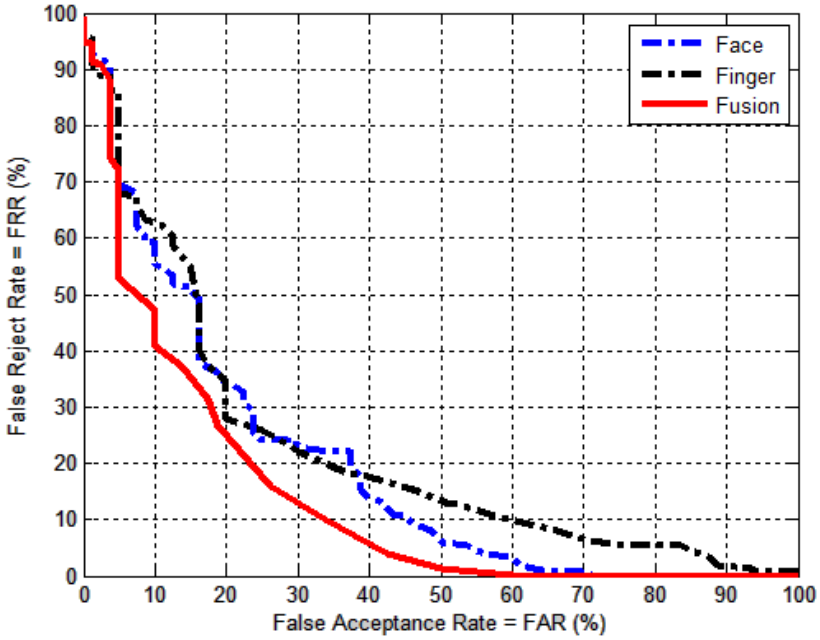


Figure 5.14. Courbes DET pour la fusion Z-Score-somme pondérée.

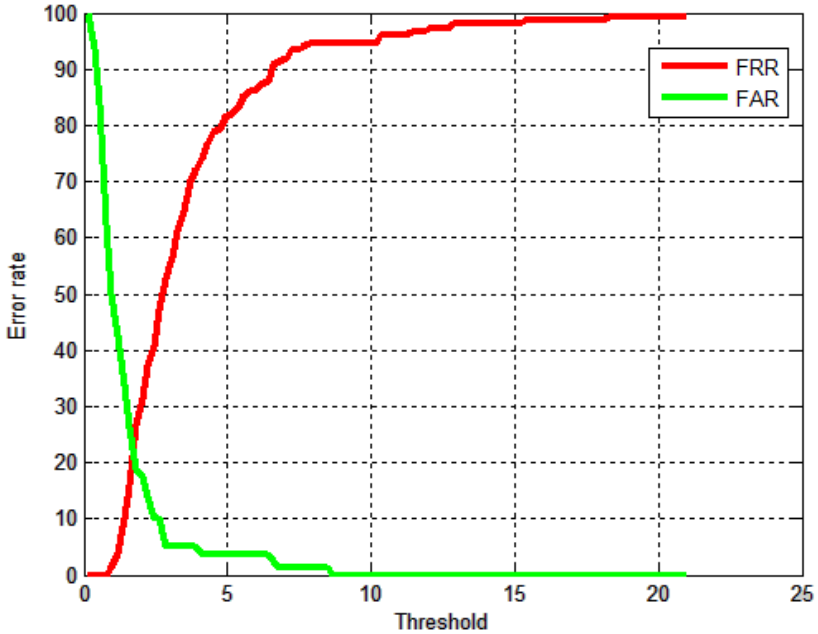


Figure 5.15. La courbe FAR vs FRR pour la fusion Z-Score-somme pondérée.

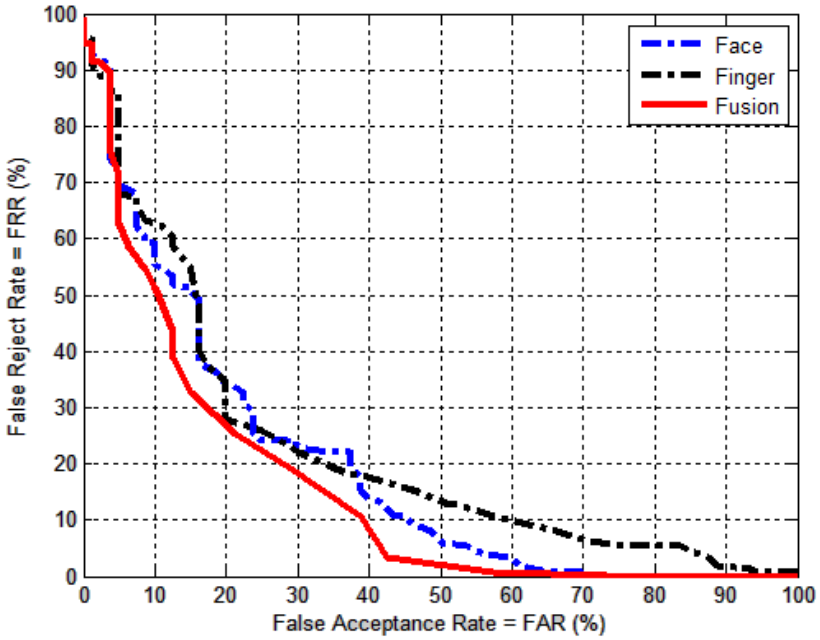


Figure 5.16. Courbes DET pour la fusion MAD-somme pondérée.

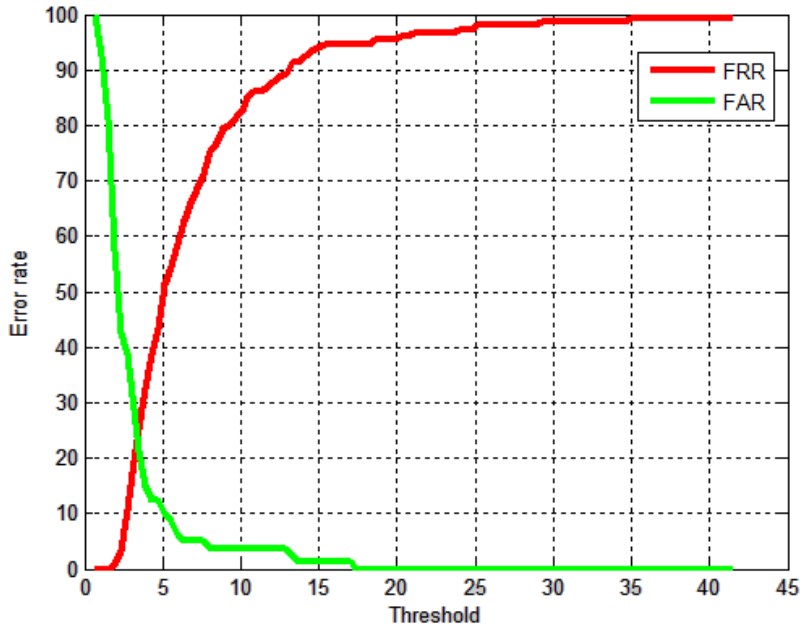


Figure 5.17. La courbe FAR vs FRR pour la fusion MAD-somme pondérée.

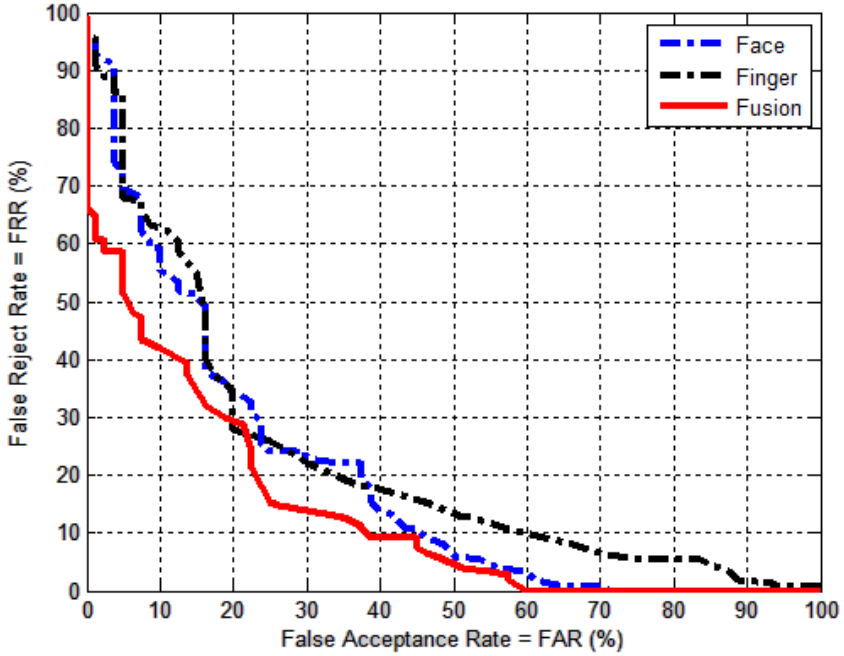


Figure 5.18. Courbes DET pour la fusion QLQ-somme pondérée.

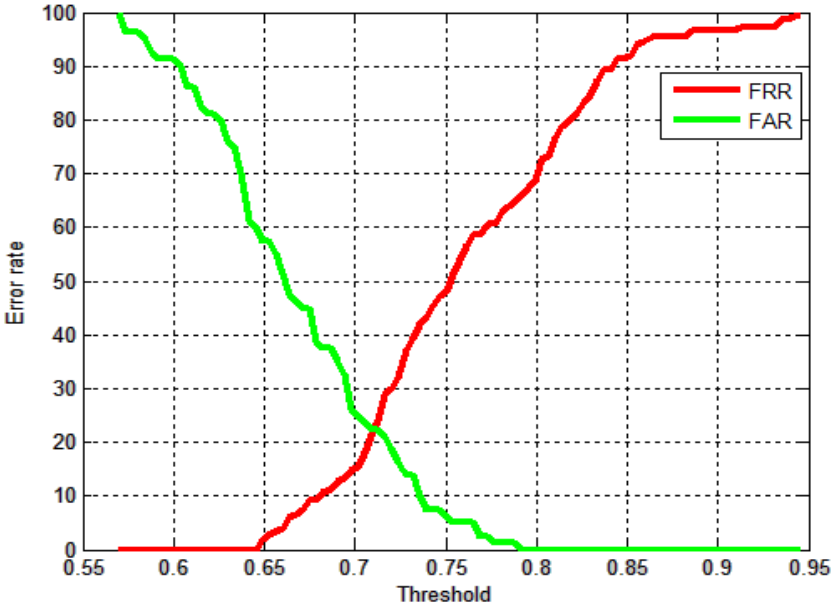


Figure 5.19. La courbe FAR vs FRR pour la fusion QLQ-somme pondérée.

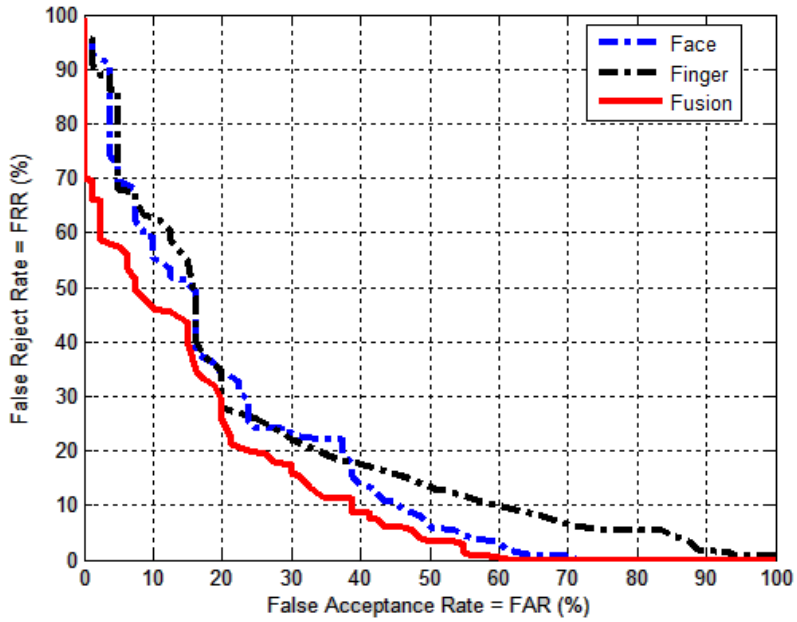


Figure 5.20. Courbes DET pour la fusion Double Sigmoïde-somme pondérée.

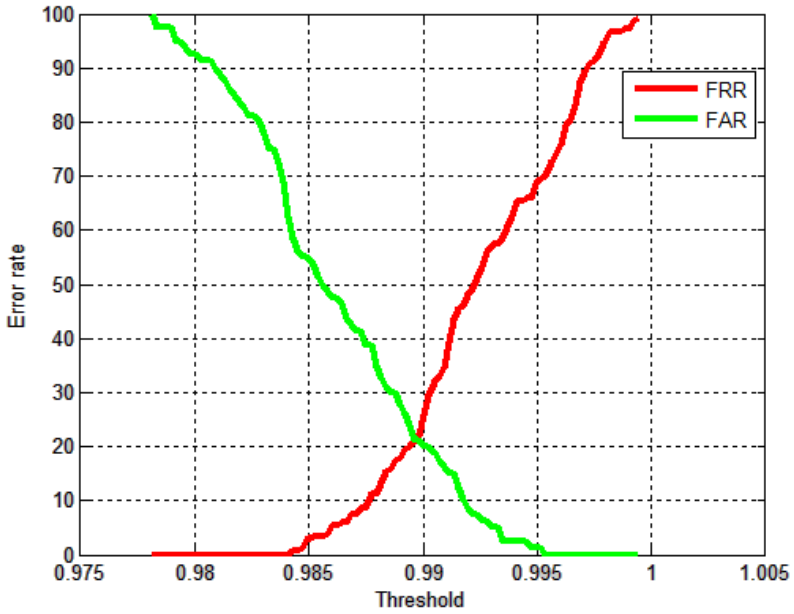


Figure 5.21. La courbe FAR vs FRR pour la fusion Double Sigmoïde-somme pondérée.

Les courbes obtenues régissent les remarques suivantes :

- La fusion des scores de deux modalités apportent une amélioration des performances quel que soit la technique de normalisation utilisée.
- La méthode Min-Max donne le meilleur résultat avec un EER = 20,33%.
- A partir des distributions de scores, nous avons déduit les paramètres de normalisation de :
  - la technique QLQ, qui sont C_v = 0.015 et W_v = 0.65 pour les scores du visage et C_e=0.01 et W_e = 0.029 pour les scores de l'empreinte digitale, et de
  - la technique double sigmoïde (DS), qui sont r_1 = 0.4, t = 1 et r_2 = 1.03 pour les scores du visage et r_1 = 13.2, t = 13.57 et r_2 = 14.3 pour les scores de l'empreinte digitale.

  Ces deux techniques donnent respectivement un taux d'erreurs égales EER_{QLQ} = 21,91% et EER_{DS} = 21,29%. Ces résultats peuvent être améliorés si nous avons choisi convenablement les paramètres de normalisation.
- La méthode MAD donne de mauvais résultats (EER = 23,29%) en comparaison avec les autres méthodes.
- Les taux EER des différentes techniques de normalisation sont résumés dans la table 5.1 et les courbes ROC de chaque système d'authentification sont illustrées sur la figure 5.22.

Système monomodal		Système multimodal (fusion des scores des modalités visage et empreinte digitale par la méthode somme pondérée)				
Visage	Empreinte digitale	Techniques de normalisation				
		Min-Max	Z-Score	MAD	QLQ	double sigmoïde
EER = 24,5%	EER = 25,5%	EER = 20,33%	EER = 22,70%	EER = 23,29%	EER = 21,91%	EER = 21,29%

Table 5.1. Les taux EER des différents systèmes d'authentification.

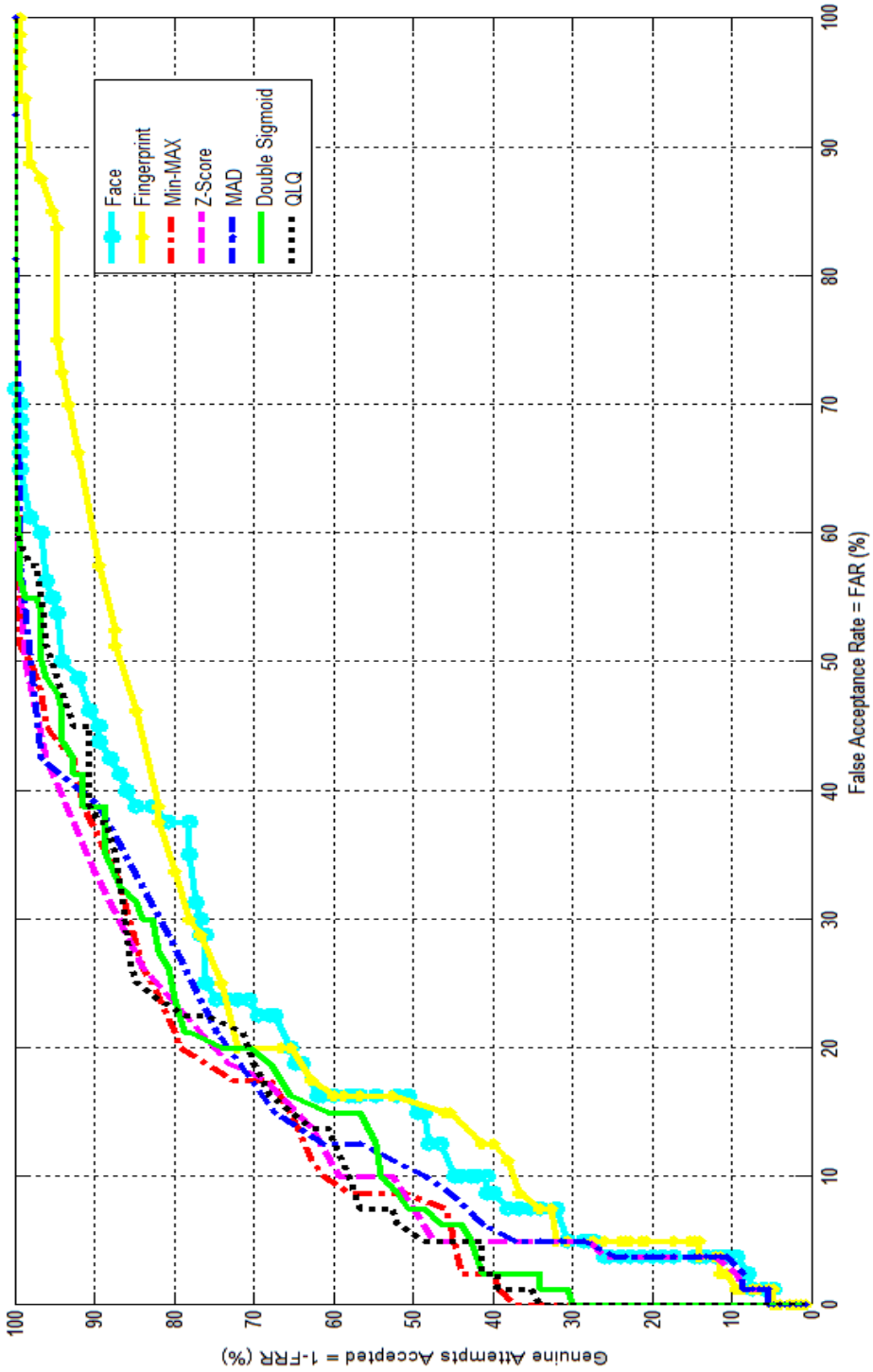


Figure 5.22. Comparaison des courbes ROC pour les différentes méthodes de normalisation.

Les temps d'exécution sur un code non optimisé par rapport aux différentes étapes du traitement sont résumés dans la table 5.2:

Enrôlement de la base des visages	Visage moyen	Amélioration de l'image de visage (l'égalisation d'histogramme)	Visage reconstruit	Appariement des visages	Enrôlement de la base des empreintes	Amélioration de l'image de l'empreinte (FFT)
0.53s	**0.01s**	**0.005s**	**0.038s**	**0.05s**	**3.22mn**	**0.10s**
Carte directionnelle de l'empreinte digitale	**Carte fréquentielle de l'empreinte digitale**	**Amélioration de l'image de l'empreinte (filtre de Gabor)**	**Binarisation de l'empreinte**	**Extraction des minuties**	**Élimination des fausses minuties**	**Appariement des empreintes**
0.59s	**0.47s**	**1.30s**	**0.008s**	**0.10s**	**0.03s**	**16.92s**

Table 5.2. Les temps d'exécution des différentes étapes du traitement.

5.7 Conclusion

Ce chapitre est consacré à l'évaluation des performances des systèmes biométriques. Nous avons implémenté un système de reconnaissance faciale basé sur l'approche PCA et un système de reconnaissance des empreintes digitales basé sur l'extraction des minuties. Les tests de ces deux algorithmes sont effectués sur les bases ORL et FVC, nous avons obtenu un taux d'erreurs égales EER = 24,5% pour la première modalité et EER = 25,5% pour la deuxième. Nous combinons ensuite les scores issus de ces deux techniques en utilisant la méthode de la somme pondérée pour construire un système multimodal. Nous affectons un poids $w_1 = 0.53$ au visage et un poids $w_2 = 0.47$ à l'empreinte digitale.

Nous avons eu une amélioration des performances des systèmes biométriques monomodaux pour toutes les techniques de normalisation des scores testées à savoir Min-Max, Z-Score, MAD, QLQ et double sigmoïde. Les meilleurs résultats sont ceux obtenus par la méthode Min-Max qui donne un taux d'erreurs égales EER = 20,33%. Ce dernier peut encore être diminué par l'affectation de poids spécifiques pour chaque individu.

Conclusion générale

Dans ce mémoire nous nous sommes intéressés au système biométrique multimodal dans le but d'améliorer la reconnaissance des individus. Nous avons présenté la biométrie, ses avantages et ses inconvénients qui mènent à l'introduction de la biométrie multimodale. Ensuite nous avons présenté l'état de l'art de la reconnaissance faciale, un domaine qui a reçu une attention accrue de la part des chercheurs. Plusieurs algorithmes sont développés cependant quelques problèmes notamment ceux dus à variation d'acquisition des images tels que la variation de pose, de changement d'illumination et les expressions faciales restent un plus grand challenge pour cette technologie. Par la suite, nous avons introduit la reconnaissance des empreintes digitales, technologie qui domine le marché mondial de la biométrie. Nous avons présenté l'état de l'art de cette modalité avant de décrire les différents niveaux de fusion des systèmes biométriques. Ensuite nous avons présenté un système multimodal basé sur la fusion au niveau des scores de modalités visage et empreinte digitale en utilisant deux algorithmes qui sont PCA qui représente les images de visages et l'algorithme basé sur l'extraction des minuties pour l'identification des empreintes. Pour combiner les scores nous avons opté pour la méthode somme pondérée. A cause de la non disponibilité des bases multimodales réelles nous effectuons nos tests sur une base virtuelle conçue en associant chaque visage de la base ORL à une et une seule empreinte digitale de la base FVC. Nous avons obtenu une amélioration des performances des systèmes biométriques monomodaux pour toutes les techniques de normalisation des scores testées ; ce qui montre l'apport de l'approche multimodalité par rapport à la monomodalité.

Par ailleurs, nous voyons les perspectives suivantes :

- Nous envisageons d'utiliser une base bimodale réelle et assez grande pour confirmer l'approche de la multimodalité.
- Mettre au point la fusion au niveau caractéristiques qui donne une riche information par rapport à celle fournie par les scores. Cependant cela augmente la complexité ainsi que le temps de calcul ce qui impose l'utilisation de techniques d'optimisation.
- Bien que la multimodalité entraîne une amélioration des performances des systèmes biométriques monomodaux, la robustesse d'un système multimodal dépend de celle des systèmes unimodaux qui le construit. Il a donc fallu de développer des algorithmes de reconnaissance unimodale. Il semble intéressant de les combiner si nous souhaitons améliorer davantage les performances de la reconnaissance.

Annexe A

Analyse Discriminante Linéaire (LDA)

Si les données d'apprentissage appartiennent à des classes prédéfinies, il est intéressant d'utiliser l'Analyse Discriminante Linéaire de *Fisher* (LDA). C'est une méthode globale basée sur la maximisation de la distance "inter-classe" tout en minimisant la distance "intra-classe".

Comme dans le PCA, nous rassemblons les images de la base d'apprentissage dans une grande matrice d'images Γ où chaque colonne représente une image Γ_i. Puis nous calculons l'image moyenne Ψ.

✓ Ensuite, nous calculons pour chaque classe C_i, l'image moyenne Ψ_{Ci} :

$$\Psi_{C_i} = \frac{1}{q_i}\sum_{k=1}^{q_i}\Gamma_k \qquad (A.1)$$

Où q_i est le nombre d'images dans la classe C_i.

✓ Nous soustrayons de chaque image Γ_i de chaque classe C_i l'image moyenne Φ_i:

$$\Phi_i = \Gamma_i - \Psi_{C_i} \qquad (A.2)$$

✓ Nous calculons ensuite les différentes matrices de dispersion à savoir :

- La Matrice de Dispersion Intra-Classe (S_w)

$$S_w = \sum_{i=1}^{c}\sum_{\Gamma_k \in C_i}(\Gamma_K - \Psi_{C_i})(\Gamma_K - \Psi_{C_i})^T \qquad (A.3)$$

- La Matrice de Dispersion Inter-Classe (S_b)

$$S_b = \sum_{i=1}^{c} q_i(\Psi_{C_i} - \Psi)(\Psi_{C_i} - \Psi)^T \qquad (A.4)$$

- La Matrice de Dispersion Totale (S_T)

$$S_T = \sum_{i=1}^{M} (\Gamma_i - \Psi)(\Gamma_i - \Psi)^T \qquad (A.5)$$

Avec c le nombre total de classes (i.e. le nombre d'individus), q_i le nombre d'images dans la classe C_i et M le nombre total d'images.

✓ Le calcul de la projection est obtenu en maximisant critère Fisher J(W) :

$$W_{opt} = arg\ max_{\ w} \frac{|W^T S_b W|}{|W^T S_w W|} \qquad (A.6)$$

✓ Ce problème est ramené à un problème généralisé aux valeurs propres [135] :

$$S_b W = \lambda_w S_w W \qquad (A.7)$$

Ou nous calculons directement les valeurs propres de $S_w^{-1}\ S_b$.

✓ Ainsi, la projection vectorielle d'une image apprise réajustée par rapport à la moyenne Φ_i est définie par :

$$g(\Phi_i) = W^T \Phi_i \qquad (A.8)$$

✓ l'image test Φ_i se projette sur W^T :

$$g(\Phi_t) = W^T \Phi_t \qquad (A.9)$$

✓ Enfin, nous effectuons la mesure de distance (Par exemple, la distance Euclidienne) entre l'image test et l'image projetée sur l'espace vectoriel engendré par W^T :

$$d_{ti} = \sqrt{\sum_{k=1}^{c} (g(\Phi_t) - g(\Phi_i))^2} \qquad (A.10)$$

✓ L'image test est dans la classe dont la distance est minimale par rapport à toutes les autres distances de classe.

Annexe B

La figure B.1 représente des échantillons de la base ORL utilisée pour les tests, cette base est disponible en libre téléchargement à l'adresse suivante :
http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

Figure B.1. Échantillons de la base ORL.

La figure B.2 représente des échantillons de la base FVC utilisée pour les tests, chaque ligne représente les images du même doigt. Cette base est disponible en libre téléchargement à l'adresse suivante :
http://www.mif.vu.lt/atpazinimas/finger/FVC/2002/DB1_A/

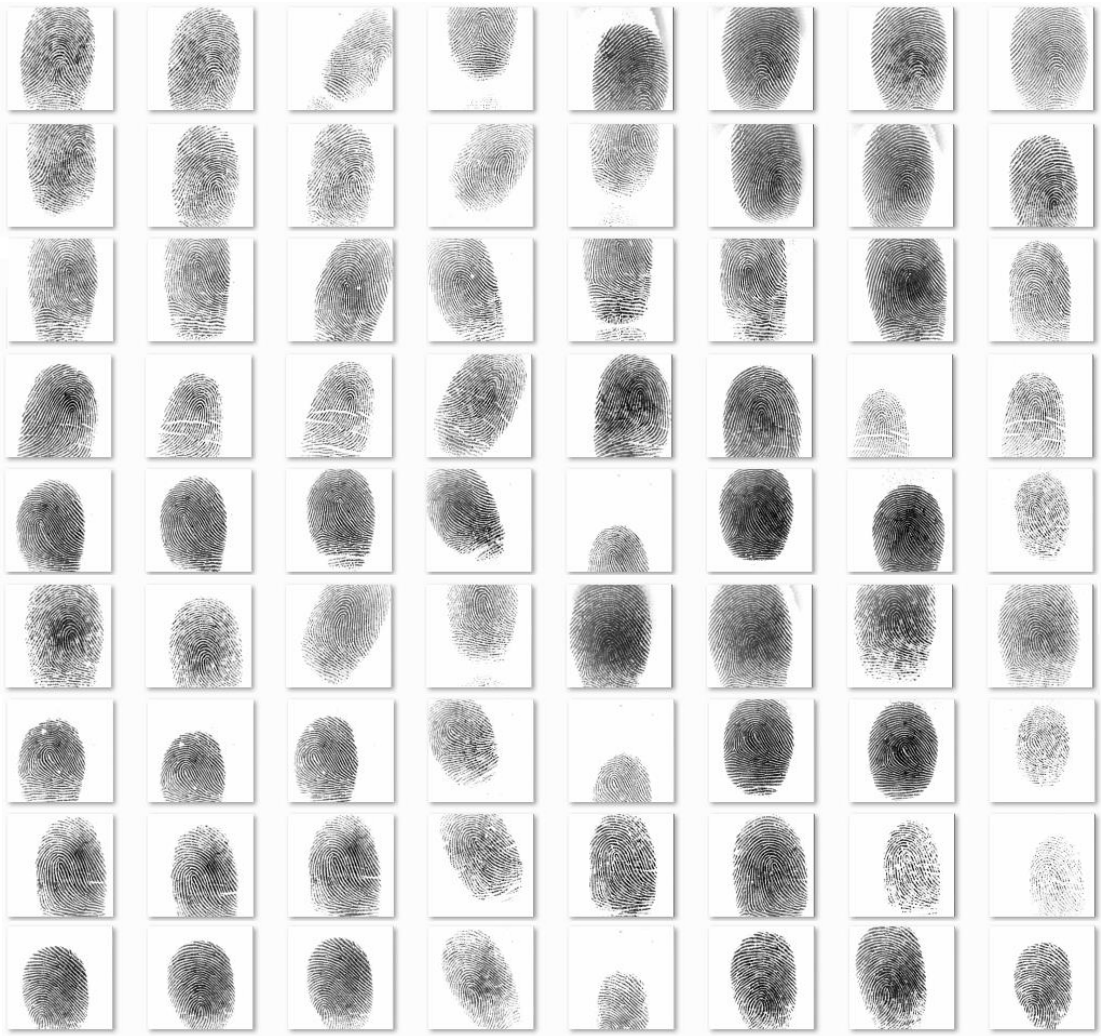


Figure B.2. Échantillons de la base FVC.

Bibliographie

[1] Peter Gregory and Michael A. Simon, "Biometrics For Dummies", Cisa, Cissp, 2008.

[2] John D. Woodward, Jr., Christopher Horn, Julius Gatune, Aryn Thomas, "biometrics, A Look at Facial Recognition", documented briefing by RAND Public Safety and Justice for the Virginia State Crime Commission, 2003.

[3] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar, "Handbook of Fingerprint Recognition", Springer, p. 340, 2005.

[4] F. Perronnin, J.-L. Dugelay, "Introduction à la Biométrie - Authentification des Individus par Traitement Audio-Vidéo", Revue Traitement du Signal, vol. 19, no. 4, 2002.

[5] Anil K. Jain, P. Flynn, A. Ross, "Handbook of biometrics", Springer, 2007.

[6] P. Meenen, R. Adhami, "Fingerprinting for Security", IEEE potentials, vol. 20, n°3, p. 33-38, 2001.

[7] R. Chellappa, C. Wilson, S. Sirohey, "Human and Machine Recognition of Faces: A Survey", Proceedings of IEEE, vol. 83, p.705-740, 1995.

[8] W. Zhao, R. Chellappa, A. Rosenfeld, P.Phillips, "Face Recognition: A Literature Survey", UMD CAR-TR, 2000.

[9] Bruce E. Koenig "Spectographic voice identification: A forensic survey", FBI, Engeineering section, Technical Service Division, 1986.

[10] P. Jonathon Phillips, Alvin Martin, C. l. Wilson, Mark Przybocki, "An introduction to evaluating biometric systems". Computer, vol. 33, n° 2, p. 56-63, 2000.

[11] A.K. Jain, R. Arun, P. Salil, "An Introduction to Biometric Recognition", IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics, vol. 14, p. 1, 2004.

[12] International Biometric Group, http://www.biomrtricgroup.com.

[13] A. K. Jain and A. Ross, "Multibiometric systems", Communications of the ACM, special issue on multimodal interfaces, vol. 47, no. 1, p. 34-40, 2004.

[14] Y. Chen, S. Dass, A. Jain, "Fingerprint Quality Indices for Predicting Authentication Performance", Proceedings of Fifth International Conference on Audio- and Video-ased Biometric Person Authentication (AVBPA), p. 160-170, 2005.

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

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

[17] J. Mahier, M. Pasquet, C. Rosenberger, F. Cuozzo, "Biometric authentication. Encyclopedia of Information Science and Technology", p. 346-354, 2008.

[18] A. Ross and A. Jain, "Information fusion in biometrics", Pattern Recognition Letters, vol. 24, n°. 13, p. 2115-2125, 2003.

[19] A. Ross, K. Nandakumar, A. Jain, "Handbook of Multibiometrics". Springer-Verlag New York, Inc., 2006.

[20] K. Nandakumar, "Integration of Multiple Cues in Biometric Systems". Master's thesis, Michigan State University, 2005.

[21] K. Chang, K. Bowyer, P. Flynn, "An Evaluation of Multimodal 2D+3D Face Biometrics". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, n°. 4, p. 619-624, 2005.

[22] Y. Adini, Y. Moses, S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, p. 721-732, 1997.

[23] D. Blackburn, J. Bone, P. Phillips, "Facial Recognition Vendor Test 2000: Evaluation Report", Technical Report A269514, National Institute of Standards and Technology, p.70, 2001.

[24] B. Fasel, J. Luettin, "Automatic Facial Expression Analysis: A Survey", Pattern Recognition, vol. 36, p. 259-275, 2003.

[25] M.S. Bartlett, G. Littlewort, I. Fasel, J.R. Movellan, "Real Time Recognition of Facial Expressions: Development and Applications to Human Computer Interaction", International Conference on Computer Vision and Pattern Recognition (CVPR), vol.5, p. 53-58, 2003.

[26] R. Gross, J. Shi, J.F. Cohen, "Quo vadis Face Recognition?", Proceedings of Third Workshop on Empirical Evaluation Methods in Computer Vision, 2001.

[27] M.Yang, J. Kriegman, N. Ahuja. "Detecting faces in images: A survey. Dans IEEE Transactions on Pattern Analysis and Machine ntelligence", vol. 24, n°. 1, p. 34-58, 2002.

[28] A.Schwaninger, S. Ryf, F.Hofer, "Configural information is processed differently in perception and recognition of faces", Vision Research, vol. 43 p.1501-15015, 2003.

[29] C. Kotropoulos, I. Pitas, "Rule based face detection in frontal views", Acoustics, Speech, and Signal processing, vol. 4, p. 2537-2540, 1997

[30] A. Yuille, P. Hallinan, D. Cohen, "Feature Extraction from Faces Using Deformable Templates",” Int’l J. Computer Vision, vol. 8, n°. 2, p. 99-111, 1992.

[31] A.Mohamed, Y.Weng, J.Jiang, S. Ipson, "Face detection based neural networks using robust skin color segmentation", 5th International Multi-Conference on systems, Signal and Devices, Amman, Jordan, 2008.

[32] H. Rowley. "Neural network-Based Face Detection". Thesis, School of Computer Science-Carnegie, Mellon Unversity, 1999.

[33] Y. Sabrina, M. T. Laskri, "Un modèle basé templates matching /réseau de neurones pour la reconnaissance des visages humains", JED’2007, Annaba, Algérie, 2007.

[34] k. Sobottka, I. Pitas, "Face localiszation and facial exraction based on shape and color information", Proccedings of the 3rd International Conference on Image Processing, ICP’96,vol. 3, p. 483-486, Lausanne, Switzerland, 1996.

[35] A.J. O’Toole, H. Abdi, "Low-dimensional representation of faces in higher dimensions of the face space", Opt. Soc. Am. vol. 10, n°. 23, p.405-411, 1993.

[36] A.K. Jain, B. Chandrasekaran, "Dimensionality and sample size considerations in pattern recognition practice", P.R. Krishnaiah, L.N.Kanal (Eds.), Handbook of Statistics, vol. 2, pp. 835-855, 1982.

[37] A.K. Jain, B. Chandrasekaran, "Dimensionality and sample size considerations in pattern recognition practice", P.R. Krishnaiah, L.N.Kanal (Eds.), Handbook of Statistics, vol. 2, North-Holland, Amsterdam, p. 835-855, 1987.

[38] S.J. Raudys, A.K. Jain, "Small sample size effects in statistical pattern recognition: recommendations for practitioners", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, n°. 3, p. 252-264, 1991.

[39] R.P.W. Duin. "Small sample size generalization", G. Borgefors (Eds.), SCIA'95, Proceeding of the Ninth Scandinavian Conference on Image Analysis, vol. 2, Uppsala, Sweden, 6-9, p. 957-964, 1995.

[40] K. Delac, M. Grgic, S. Grgic, "Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set", Technical Report, University of Zagreb, FER, 2004.

[41] M. A. Turk and A. P. Pentland, "Face recognition using Eigenfaces", IEEE Conference on Computer Vision and Pattern Recognition, p. 586-590, Hawai, 1992.

[42] M. Kirby, L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, n°. 1, p. 103-108, 1990

[43] A. M. Martínez, A. C. Kak, "PCA versus LDA". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, n°. 2, p. 228-233, 2001.

[44] D. J. Kriegman, J. P. Hespanha, P. N. Belhumeur, "Eigenfaces vs. fisherfaces : Recognition using class-specific linear projection", European Conference on Computer Vision, p. 43-58, 1996.

[45] G. Bahtiyar, "Holistic Face Recognition by Dimension Reduction", Master's thesis, Department of Electrical and Electronics Engineering, Graduate School of Natural and Applied Sciences of the Middle East Technical University, 2003.

[46] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection". IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997.

[47] D.A. Socolinsky, A. Selinger, "Thermal face recognition in an operational scenario", IEEE Conference on Computer Vision and Pattern Recognition, p. 1012-1019, 2004.

[48] W. Shi-Qian, W. Li-Zhen, F. Zhi-Jun, L. Run-Wu, Y. Xiao-Qin."Infrared face recognition based on blood perfusion and sub–block duct in wavelet domain", International Conference on Wavelet Analysis and Pattern Recognition, 2007.

[49] P.N. Belhumeur, D.J. Kriegman, "What is the set of images of an object under all possible lighting conditions ? ", Computer Vision and Pattern Recognition, 1996, Proceedings CVPR '96, 1996 IEEE Computer Society Conference, p. 270-277, 1996.

[50] M. S. Bartlett, J. R. Movellan, T. J. Sejnowski, "Face recognition by independent component analysis", Transactions on Neural Networks, 2002.

[51] I. Buciu, I. Pitas, "Application of non-negative and local non negative matrix factorization to facial expression recognition", International Conference on Pattern Recognition, p. 288-291, 2004.

[52] Y. Wang, Y. Jia, C. Hu, M. Turk, "Non-negative matrix factorization framework for face recognition", Internation Journal of Pattern Recognition and Artificial Intelligence, vol. 19, n°. 4, p. 495-511, 2005.

[53] M. Visani, C. Garcia, J. M. Jolion, "Normalized radial basis function networks and bilinear discriminant analysis for face recognition", IEEE Conference on Advanced Video and Signal Based Surveillance, p. 342-347, 2005.

[54] H. Cevikalp, M. Neamtu, M. Wilkes, A. Barkana, "Discriminative common vectors for face recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, n°. 1, p. 4-13, 2005.

[55] B. Schölkopf, A. Smola, K.-R. Müller,"Nonlinear component analysis as a kernel eigenvalue problem", Neural Computation, vol. 10, n°. 5, p. 1299-1319, 1998.

[56] S. Mika, G. Ratsch, J.Weston, B. Schölkopf, K.-R. Müller."Fisher Discriminant Analysis With Kernels", Neural Networks for Signal Processing IX, p. 41-48, 1999.

[57] H. Kim, H. Park, H. Zha, "Distance preserving dimension reduction for manifold learning", International Conference on Data Mining, 2007.

[58] S. Biswas, K. W. Bowyer, Patrick J. Flynn,"Multidimensional scaling for matching low–resolution facial images", Biometrics : Theory, Applications and Systems, 2010.

[59] M. H. Yang, "Face recognition using extended isomap", International Conference on Image Processing, p. 117-120, 2002.

[60] G. Hagen, T. Smith, A. Banasuk, R.R. Coifman, I. Mezic, "Validation of low-dimensional models using diffusion maps and harmonic averaging", IEEE Conference on Decision and Control, 2007.

[61] M. N. Teli, "Dimensionality reduction using neural networks", Technical report, 02 2008.

[62] M. K. Fleming, G. W. Cottrell, "Categorization of faces using unsupervised feature extraction", IEEE International Joint Conference on Neural Networks, vol. 2, p. 65-70, San Diego, 1990.

[63] R. Brunelli, T. Poggio, "Face recognition : Features versus templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, n°. 10, p. 1042-1052, 1993.

[64] H. A. Rowley, S. Baluja, T. Kanade, "Neural network-based face detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, n°. 1, p. 23-38, 1998.

[65] A. Lanitis, C. J. Taylor, T. F. Cootes, "Automatic face identification system using flexible appearance models", Image and Vision Computing, vol. 13, n°. 5, p. 393-401, 1995.

[66] B.S. Manjunath, R. Chellappa, C.V.D. Malsburg, "A feature based approach to face recognition, in: Proceedings", IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, p. 373-378, 1992.

[67] T.S. Lee, "Image representation using 2-d Gabor wavelets", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, n°. 10, p. 959-971, 1996.

[68] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, Malsburg, R.Wurtz, "Distortion invariant object recognition in the dynamic link architecture", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, n°. 3, 42 (3), p. 300-311, 1993.

[69] L. Wiskott, J. M. Fellous, N. Kuiger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, n°. 7, p. 775-779, 1997.

[70] B. Kepenekci, F.B. Tek, G. Bozdagi Akar, "Occluded face recognition based on Gabor wavelets", ICIP 2002, Rochester, NY, p. 3-10, 2002.

[71] B.S. Manjunath, R. Chellappa, C.V.D. Malsburg, "A feature based approach to face recognition", Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, p. 373-378, 1992.

[72] H.S. Le, H. Li. "Recognizing frontal face images using hidden Markov models with one training image per person", Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), vol. 1, p. 318-321, 2004.

[73] T. Ahonen, A. Hadid, M. Pietikainen, "Face recognition with local binary patterns" ECCV, p. 469-481, 2004.

[74] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, n°. 2, p. 91-110, 2004.

[75] S.C. Chen, J. Liu, Z.-H. Zhou, Makin, G. FLDA, "applicable to face recognition with one sample per person", Pattern Recognition, vol. 37, n°. 7, p. 1553-1555, 2004.

[76] X. Tan, S.C. Chen, Z.-H. Zhou, F. Zhang, "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft kNN ensemble", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, n°. 4, p. 875-886,2005.

[77] P. Viola, M.J. Jones, "Robust real-time face detection", International Journal of Computer Vision, vol. 57, p. 137-154, 2004.

[78] R. Snelick, U. Uludag, A. Mink, M. Indovina, A. Jain, "Large-Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, n°. 3, pp. 450-455, 2005.

[79] A.M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, n°. 6, p. 748-763, 2002.

[80] X. Tana, C. Songcan, "Face recognition from a single image per person: A survey", Pattern Recognition, 2006.

[81] P. W. Hallinan, G. G. Gorden, A. L. Yuille, P. Giblin, D. Mumford, "Two-and Three-dimensional Patterns of the Face", A K Peters, Ltd., Natick, MA, 1999.

[82] B. Heisele, T. Serre, M. Pontil, T. Poggio, "Component-based face detection", Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, p. 657-662, 2001.

[83] N.P. Costen, T.F. Cootes, C.J. Taylor, "Compensating for ensemble specific effects when building facial models", Image Vision Computing. 20, p. 673-682, 2002.

[84] T. F. Cootes, C. J. Taylor, "Constrained active appearance models", International Conference on Computer Vision, p. 748-754, 2001.

[85] M. Loève, "Fonctions aléatoires du second ordre", Processus stochastiques et mouvements browniens, 1948.

[86] M. Turk and A. Pentland, "Eigenfaces for Recognition", J. Cognitive Neuroscience, vol. 3, n°. 1, p. 71-86, 1991.

[87] V. Perlibakas, "Distance measures for pca-based face recognition", Pattern Recognition Letters, vol. 25, n°. 6, p. 711-724, 2004.

[88] H. Moon , P.J. Phillips, "Analysis of pca-based face recognition algorithms", Empirical Evaluation Techniques in Computer Vision, 1998.

[89] P. J.Phillips, H. Wechsler, J. Huang, P. Rauss, "The feret database and evaluation procedure for face-recognition algorithms", Image and Vision Computing, 1998.

[90] X. Chen, P. J. Flynn, K. W. Bowyer, "IR and visible light face recognition", Computer Vision and Image Understanding, vol. 99, n°. 3, p.332-358, 2005.

[91] T. K. Kim, H. W. Kim, W. J. Hwang, J. V. Kittler, "Independent component analysis in a local facial residue space for face recognition", Pattern Recognition, vol. 37, n°. 9, p. 1873-1885, 2004.

[92] Francis Galton, Fingerprint, McMillan, London, 1892.

[93] International Biometric Group, The Henry Classification, www.biometricgroup.com

[94] Maltoni Davide, Dario Maio, Anil K. Jain, Salil Prabhakar, Handbook of fingerprint recognition, Springer, New York, 2003.

[95] M. Sezgin et B.Sankur, "Survey over image thresholding technique and quantitative performance evaluation", Journal of Electronic Imaging, vol. 13, p. 146-165, 2004.

[96] L.Lam, S.W. Lee et C.Y. Suen, "Thinning Methodologies-A Comprehensive Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, vol. 9, p. 869-885, 1992.

[97] Arcelli C., and Baja G.S.D., "A Width Independent Fast Thinning Algorithm", IEEE Transaction on pattern Analysis and Machine Intelligence, vol. 4, n°. 7, pp. 463-474, 1984.

[98] N.Galy , "Etude d'un système complet de reconnaissance d'empreintes digitales pour un capteur microsystème à balayage", p. 80, 2005.

[99] S.Kim, D.Lee, J. Kim, "Algorithm for Detection and Elimination of False Minutiae in Fingerprint Image", Lecture Notes in Computer Science, Spring Verlag, vol. 2091, p. 235-240, 2001.

[100] Z.Bian, D.Zhang, W.Shu, "Knowledge-Based Fingerprint Post-Processing", International Journal of Pattern Recognition and artificial Intelligence, vol. 16, n°.1,p. 53-67, 2002.

[101] A. Farina, Z.M. Kovacs-Vajna, A. Leone, "Fingerprint minutiae extraction from skeletonized binary images", Pattern Recognition, vol. 32, p. 877-889, 1999.

[102] D. Maio, D. Maltoni, "Direct Gray-Scale Minutiae Detection In Fingerprints", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 19, n°. 1, p. 27-39, 1997.

[103] Francesco Turroni, "Fingerprint Recognition: Enhancement, Feature Extraction and Automatic Evaluation of Algorithms", dottorato di ricerca in Informatica Ciclo XXIV Universita di Bologna , 2012

[104] A. Ravishankar Rao, "A Taxonomy For Texture Description and Identification", Springer Verlag, New York, 1990.

[105] X. Jiang, W. Yau, W. Ser, "Detecting the fingerprint minutiae by adaptative tracing the gray-level ridge", Pattern Recognition, vol. 34, p. 999-1013, 2001.

[106] N. Ratha, R. Bolle, "Automatic Fingerprint Recognition Systems", Springer, New York, 2004.

[107] A.M Bazen, S.H. Gerez, "An intrinsec Coordinate System for Fingerprint Matching", Proc. Int. Conf. on Audio- and Video-Based Biometric Person Authentification, p.198-204, 2001.

[108] University of Bologna http: //biolab.csr.unibo.it, de 9/02/2009.

[109] E. Kaymaz, S. Mitra, "Analysis and Matching of Degraded and Noisy Fingerprints" Proc. of SPIE (Application of Digital Image Processing), vol. 1771, p. 498-508, 1992.

[110] J.D. Stosz, L.A. Alyea, "Automated System for Fingerprint Authentification Using Pores and Ridge Structure", Proc. of SPIE (Automatic Systems for the Identification and Inspection of Humans), vol. 2277, p. 210-223, 1994.

[111] L . M.Munoz-Serrano, "Sistema automtico de reconocimiento de huella dactilar basado en informacin de textura", master's thesis, ETSIT, université Politecnia de Madrid, 2004.

[112] A.K. Jain, L. Hong, S. Pankanti, R. Bolle, "An identity authenfication system using fingerprints", Proceedings of the IEEE, vol. 85, n° 9, 1997.

[113] C. Sanderson, K. Paliwal, "Information fusion and person verification using speech and face information", Tech. Rep. IDIAP-RR p. 02-33, IDAIP, 2002.

[114] S. Lyengar, L. Prasad, H. Min, "Advances in Distributed Sensor Technology", 1995.

[115] A. Ross, A. Jain, "Fingerprint Mosaicking", Proceedings of the IEEE International Conference on Acoustic Speech and Signal Processing, 2002.

[116] Y. Moon, H. Yeung, K. Chan, S. Chan, "Template synthesis and image mosaicking for fingerprint registration : An experimental study", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, p. 409-412, 2004.

[117] G. N. Bebis, A. Gyaourova, S. Singh, I. T. Pavlidis, "Face recognition by fusing thermal infrared and visible imagery", Image and Vision Computing, vol. 24, n° 7, p. 727-742, 2006.

[118] Xiao-Yuan Jing, Yong-Fang Yao, David Zhang, Jing-Yu Yang, MiaoLi, "Face and palmprint pixel level fusion and kernel DCV-RBF classifer for small sample biometric recognition", Pattern Recognition, vol. 44, n° 11, p. 3209-3224, 2007.

[119] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classification", John Wiley & Sons, 2001.

[120] L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition : An analysis of its behavior and performance", IEEE Transactions on Systems, Man, and Cybernetics, vol. 27, 1997.

[121] Hao, Anderson, Daugman, "Combining crypto with biometrics efficiently", IEEE Transactions on Computers, vol. 55, 2006.

[122] L. Lam, C. Y. Suen, "Optimal combinations of pattern classifiers", Pattern Recognition Letters, vol. 16, n° 9, p. 945-954, 1995.

[123] L. Xu, A. Krzyzak, C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition", IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, n° 3, p. 418-435, 1992.

[124] A. Jain, K. Nandakumar, A. Ross, "Score normalization in multimodal biometric systems", Pattern Recognition, vol. 38, n°. 12, p. 2270-2285, 2005.

[125] P. Huber, "Robust Statistics", JohnWiley & Sons, 1981.

[126] R. Cappelli, D. Maio, D. Maltoni, "Combining Fingerprint Classifiers", Proceedings of the First International Workshop on Multiple Classifier Systems, p. 351-361, Springer-Verlag, London, UK, 2000.

[127] F. Hampel, P. Rousseeuw, E. Ronchetti, W. Stahel, "Robust Statistics : The Approach Based on Influence Functions", John Wiley & Sons, 1986.

[128] J. Kittler, M. Hatef, R. Duin, J. Matas, "On Combining Classifiers", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, n°. 3, p. 226-239, 1998.

[129] A. Jain, A. Ross, "Learning User-specific Parameters in a Multibiometric System". Proceedings of International Conference on Image Processing (ICIP), p. 57-60, New York, NY, USA, 2002.

[130] Y.Wang, T. Tan, A. Jain, "Combining face and iris biometrics for identity verification", Proceedings of Fourth International Conference on Audio- and Video-Based Authentication (AVBPA), p. 805-813, Guildford, U.K., 2003.

[131] M. Fuents, D. Mostefa, J. Kharroubi, S. Gracia-salicetti, B. Dorizzi, G. Chollet, "Vérification de l'Identité par Fusion de données Biométrique : Signature En-Ligne et Parole", soumis à la Conférence Internationale Francophone sur l'écrit et Document, CIFED'02, 2002.

[132] L. Hong, A. Jain, "Integrating Faces and Fingerprints for Personal Identification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, n°. 12, p. 1295-1307, 1998.

[133] C.I. Waston, G.T. Candela, P.J. Grother, "Comparison of fft Fingerprint Filtering Methods for Neural Network Classification", NISTIR, 5493, 1994.

[134] L. Hong, Y. Wan, A. Jain, "Fingerprint image enhancement Algorithm and performance evaluation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.20, n°. 8, p. 777-789, 1998.

[135] G. Golub, "The Generalized Eigenvalue Problem", Lectures on Matrix Computation, Ph.D. program of the Dipartimento di Matematica “Istituto Guido Castelnuovo”, Roma, 2004.