

République Algérienne Démocratique et Populaire

Université Mouloud MAMMERRI de TIZI-OUZOU (UMMTO)

Faculté de Génie Electrique et d'Informatique

Département d'Electronique



**Mémoire de Magister en Electronique
Option : Télédétection**

Présenté par :

BOUDJELLAL Sofiane

Thème

**« Détection et identification de personne par méthode
biométrique »**

Devant les membres du jury :

Président :	Mr.	AMEUR	Soltane	Professeur	UMMTO
Rapporteur :	Mme.	AMEUR	Zohra	M.C.A	UMMTO
Examineurs :	Mr.	DI AF	Moussa	Professeur	UMMTO
	Mr.	HADDAB	Salah	M.C.A	UMMTO
	Mr.	LAHDIR	Mourad	M.C.A	UMMTO

Remerciements

Je tiens à remercier sincèrement Madame Aneur, qui, en tant que Directrice de ce mémoire, s'est toujours montrée à l'écoute et très disponible tout au long de la réalisation de ce mémoire de Magister, ainsi pour l'inspiration, l'aide et le temps qu'elle a bien voulu me consacrer. Je tiens à exprimer toute ma reconnaissance à Monsieur AMEUR Soltane, Professeur à l'UMMTO et directeur du laboratoire de recherche LAMPA, pour l'honneur qu'il me fait de présider mon jury de thèse, Aussi je remercie vivement Monsieur DIAF Moussa Professeur à l'UMMTO, Monsieur LAHDIR Mourad Maître de conférences à l'UMMTO et Monsieur HADDAB Salah Maître de conférences à l'UMMTO, membres du jury, pour leur éminente contribution à l'évaluation de ce projet.

Enfin, j'adresse mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenu et encouragé au cours de la réalisation de ce travail.

Merci à toutes et à tous.

Dédicaces

- ✚ A mon père, à ma mère
- ✚ A mon frère et sœurs
- ✚ A ma femme et sa famille
- ✚ A la mémoire de mes grands parents
- ✚ A tout mes amis (es)
- ✚ A tout mes maîtres et professeurs : du primaire au supérieur
- ✚ A tous ceux qui ont contribué au développement des sciences en général et de l'électronique en particulier

Table des matières

Introduction générale	1
CHAPITRE I : LA BIOMETRIE	
I.1. Préambule	3
I.2. La biométrie	3
I. 2.1. Définition	3
I.3. Le marché mondial de la biométrie.....	4
I.4. Les parts de marché par technologie	5
I.5. Les techniques biométriques	5
I.5.1. L'analyse morphologique (physiologique).....	5
I.5.2. L'analyse comportementale	5
I.6. Architecture d'un système biométrique	6
I.6.1. Module d'apprentissage	6
I.6.2. Module de reconnaissance	6
I.6.3 .Module d'adaptation.....	7
I.7. Présentation de quelques technologies biométriques	8
I.7.1. Les empreintes digitales	8
I.7.2. La voix	9
I.7.3. L'iris	9
I.7.4. La signature	10
I.7.5. La géométrie de la main	10
I.7.6. Le visage	11
I.8 .Reconnaissance du visage	12
I.8.1. Approches globales	13
I.8.2. Méthodes locales	15
I.8.3. Méthodes hybrides	16
I.9. Principales difficultés de la reconnaissance de visage	16
I.9.1. Changement d'illumination	16
I.9.2. Variation de pose	17

Table des matières

I.9.3. Expressions faciales	17
I.9.4. Présence ou absence des composants structurels	17
I.9.5. Occultations partielles	18
I.10. Evaluation de performance	18
I.10.1. Evaluation de l'identification	18
I.10.2. Evaluation de la vérification	18
I.11 Discussion	20
CHAPITRE II : TECHNIQUES DE DETECTION ET DE RECONNAISSANCE DE VISAGES	
II.1. Préambule	21
II.2. Détection de visages	21
II.3. Approches de la détection de visage	21
II.3.1. Approches basées sur l'apparence	21
II.3.2. Approches basées sur les connaissances acquises	22
II.3.3 Approches basées sur le « Template-matching »	22
II.3.4 Approches basées sur des caractéristiques invariantes	24
II.3.4.1 Méthodes basées sur les caractéristiques du visage	24
II.3.4.2. Méthodes basées sur l'analyse de la couleur de la peau	24
II.4. Prétraitement	25
II.4.1 Normalisation Photométrique	26
II.4.2. Normalisation géométrique	27
II.5. Les algorithmes de reconnaissance de visage	28
II.5.1. Dimensionnalité de l'espace visage	28
II.5.2. Analyse en Composantes Principales	28
II.5.2.1. Présentation	28
II.5.2.2. Le processus de reconnaissance (Karhunen-Loeve)	29
II.5.3. Analyse Discriminante Linéaire	30
II.5.3.1. Présentation	30

Table des matières

II.5.4. DF-LDA	31
II.5.5. Analyse en Composantes Indépendantes	33
II.5.6. Le réseau de neurones	35
II.5.6.1. Introduction	35
II.5.6.2. Le réseau de neurones LVO	36
II.5.7. Machine à vecteurs de support	37
II.5.7.1. Le principe de SVM	37
II.5.7.2. La séparation optimale	37
II.5.7.3. Le concept d'hyperplan	37
II.5.7.4. La classification linéaire	38
II.5.7.5. La classification non linéaire	39
II.5.8. Modèle de Markov caché	39
II.5.9. Méthodes basées sur l'apparence locale	40
II.5.9.1. ACP modulaire	40
II.5.10. Template matching	40
II.5.10.1. L'algorithme du template matching	41
 CHAPITRE III : RECONNAISSANCE DE VISAGE PAR EIGENFACE 	
III.1. Préambule	42
III.2. Présentation de la méthode Eigenface	42
III.2.1. Choix du nombre de vecteurs propres	44
III.2.2. Utilisation des Eigenfaces pour la classification des images visage	46
III.2.3. Résumé de la méthode	47
III.3. Mesures de Distance	48
III.3.1. Distances Euclidiennes	48
III.3.1.1 Distance City Block (L1).....	48
III.3.1.2 Distance Euclidienne (L2).....	48

Table des matières

III.3.2. Distances dans l'Espace de Mahalanobis	49
III.3.2.1. De l'espace des images à l'espace de Mahalanobis	49
III.3.2.2. Mahalanobis L1 (MahL1)	50
III.3.2.3. Mahalanobis L2 (MahL2)	50
III.3.2.4. Cosinus de Mahalanobis (MahCosine)	50
III.3.3. Choix de la distance de similarité	51
III.4. Organigramme détaillé de l'approche Eigenface	51
III.4.1. Organigramme du prétraitement	52
III.4.2. Organigramme de la phase d'apprentissage	53
III.4.3. Organigramme de la phase d'identification	54
III.5. Les bases de données de visage	55
III.5.1 AT&T	55
III.5.2 FEI	55
III.5.3. MUCT	56
III.5.4. CVSRP	56
CHAPITRE IV : TESTS ET RESULTATS	
IV.1. Preambule	58
IV.2 Environnement du travail	58
IV.2.1 Environnement matériel	58
IV.2.2. Outils de développement	58
IV.2.2.1. Matlab 7.12.0 (R2011a)	58
IV.3. Présentation de l'application	58
IV.3.1. Interface Présentation	58
IV.4. Démarche à suivre pour la reconnaissance faciale	61
IV.5. Fonctionnement du système	63

Table des matières

IV.5.1. Phase d'apprentissage	63
IV.5.2. Phase d'identification	64
IV.6. Détermination du seuil d'identification	65
IV.7. Evolution du taux d'identification suivant le changement de poses et d'éclairages	70
IV.7.1. Le changement de pose	70
IV.7.2. Le changement des conditions d'éclairages	72
IV.8. Temps d'exécution	73
IV.9. L'amélioration des performances de PCA	75
IV.9.1. Les conditions optimales de pose et d'éclairage	75
IV.10. Avantages et inconvénients du système	76
IV.11. Discussion	77
Conclusion générale	78

Introduction

De nos jours on parle de plus en plus de l'insécurité dans divers secteurs ainsi que des moyens informatiques à mettre en œuvre pour contrer cette tendance. La vérification et l'identification des individus est l'un des moyens permettant d'assurer cette sécurité. L'être humain se sert quotidiennement de son système visuel pour identifier les personnes de façon automatique, bien que le processus mis en jeu soit complexe.

L'homme a mis en place des moyens de vérification d'identité qui sont liés, soit à ce que possède une personne telle qu'une carte d'identité ou un passeport, soit à ce que sait cette personne, c'est le cas du mot de passe ou un code PIN. Néanmoins, ces éléments peuvent être oubliés, volés ou falsifiés. Pour contourner ces limitations, un autre moyen de sécurité a été développé qui permet d'utiliser, non pas l'information qu'un individu possède ou connaît, mais une information (propre) intrinsèque à cette personne. Cette nouvelle façon d'identification des individus est la biométrie.

Les caractéristiques biométriques sont une solution alternative aux anciens moyens de vérification d'identité. L'avantage de ces caractéristiques biométriques est d'être universelles, c'est-à-dire présentes chez toutes les personnes à identifier. D'autre part, elles sont mesurables et uniques : deux personnes ne peuvent pas posséder exactement la même caractéristique. Elles sont aussi permanentes ce qui signifie qu'elles ne varient pas ou peu au cours du temps.

Pour que des caractéristiques collectées puissent être qualifiées de modalités biométriques, elles doivent être :

- ✓ universelles (exister chez tous les individus),
- ✓ uniques (permettre de différencier un individu par rapport à un autre),
- ✓ permanentes (autoriser l'évolution dans le temps),
- ✓ enregistrables (collecter les caractéristiques d'un individu avec son accord),
- ✓ mesurables (autoriser une comparaison future).

L'intérêt principal de la biométrie est donc de reconnaître et d'identifier automatiquement les identités des individus en utilisant leurs caractéristiques physiologiques ou comportementales. Les caractéristiques physiologiques peuvent inclure le visage, l'iris, les empreintes, la géométrie de la main. Les caractéristiques comportementales incluent la voix, la signature, la démarche etc.

Dans ce mémoire, nous nous intéressons à la biométrie du visage qui possède beaucoup d'avantages tels que, la facilité d'utilisation, l'acceptation par l'utilisateur (car elle est non intrusive) et le faible coût. Ainsi, la reconnaissance du visage est déjà intégrée dans des systèmes de sécurité biométriques utilisant un certain nombre d'algorithmes classiques.

La difficulté de la reconnaissance de visage par ordinateur varie énormément suivant que les conditions d'acquisition. Dans un environnement contrôlé, des paramètres tels que l'arrière-plan, la direction et l'intensité des sources lumineuses, l'angle de la prise de vue, la distance de la caméra au sujet sont des paramètres maîtrisés par le système. Dans un

environnement non contrôlé, une série de prétraitements sont souvent nécessaires avant de faire la reconnaissance proprement dite.

Les travaux de ce mémoire s'inscrivent dans un contexte global de recherche et développement de d'algorithmes pour la reconnaissance de visages. Ainsi, au cours de ce mémoire, nous aurons à concevoir un système de reconnaissance facial avec la méthode Eigenface qui se base sur une analyse en composante principale. L'ACP est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard.

Nous avons choisi d'articuler notre étude autour de quatre chapitres principaux. Le premier chapitre est consacré à la présentation générale de la biométrie. Il décrit le principe de fonctionnement des systèmes biométriques puis définit les outils utilisés pour évaluer leurs performances. Ensuite, la place de la reconnaissance faciale parmi les autres techniques biométriques est analysée. A travers ce chapitre, nous voulons positionner le problème de la reconnaissance faciale et présenter ses enjeux et intérêts par rapport aux autres techniques. Nous étudions également les principales composantes d'un système de reconnaissance faciale, notamment la détection, l'extraction de caractéristiques et la reconnaissance. Enfin, nous mettons en lumière les difficultés rencontrées par les systèmes de reconnaissance de visage.

Dans le second chapitre une analyse détaillée des différentes techniques développées au cours de ces dernières années dans les domaines de la détection de visage et de la reconnaissance est présentée. Notre objectif est de classer ces méthodes de manière efficace afin de mettre en évidence les particularités ainsi que les avantages et les inconvénients de chacune d'entre elles. Ceci nous permettra, par la suite, de mettre en avant les critères de performances qui vont guider le choix des solutions retenues dans le cadre de notre problématique. Nous évoquerons aussi la normalisation géométrique des images des visages et les différentes normalisations photométriques.

Dans le troisième chapitre nous présenterons notre algorithme qui se base sur une analyse en composante principale, qui est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard. Nous verrons plusieurs approches pour améliorer les performances de PCA, puis nous présentons les bases données utilisées dans le cadre de ce mémoire, FEI, CVSRP, AT&T, IFW et MUCT.

Dans le quatrième chapitre nous présenterons notre interface graphique qui a été créé à l'aide du GUI (Graphical User Interface) sous Matlab, nous analyserons les résultats des tests effectués sur plusieurs bases de données sous différentes conditions.

En conclusion, nous récapitulerons les principales contributions de ce mémoire avant d'exposer les perspectives envisagées.

Chapitre I

La biométrie

I.1 Préambule

Les systèmes biométriques sont de plus en plus utilisés depuis quelques années. L'apparition de l'ordinateur et sa capacité à traiter et à stocker les données ont permis la création des systèmes biométriques informatisés. Il existe plusieurs caractéristiques physiques uniques pour un individu, ce qui explique la diversité des systèmes appliquant la biométrie, selon ce que l'on prend en compte :

- ✓ L'empreinte digitale
- ✓ La géométrie de la main
- ✓ L'iris
- ✓ La rétine ... etc

Nous allons voir dans ce chapitre les principales technologies biométriques, puis nous allons nous focaliser sur les systèmes de reconnaissances faciales, leurs avantages et les problèmes liés à leurs applications.

I.2. La biométrie

I.2.1. Définition

La biométrie peut être définie comme étant "la reconnaissance automatique d'une personne en utilisant des traits distinctifs". Une autre définition de la biométrie est "toutes caractéristiques physiques ou traits personnels automatiquement mesurables, robustes et distinctives qui peuvent être utilisées pour identifier un individu ou pour vérifier l'identité prétendue d'un individu" [1].

La biométrie consiste en l'analyse mathématique des caractéristiques biologiques d'une personne et a pour objectif de déterminer son identité de manière irréfutable. Contrairement à ce que l'on sait ou ce que l'on possède la biométrie est basée sur ce que l'on est et permet ainsi d'éviter la duplication, le vol, l'oubli ou la perte.

Un système biométrique peut avoir deux modes opératoires [4] :

- L'identification, elle permet d'établir l'identité d'une personne à partir d'une base de données, le système biométrique pose et essaye de répondre à la question, "qui est la personne X ? ", il s'agit d'une comparaison du type un contre plusieurs (1:N).
- La vérification ou l'authentification, le système biométrique demande à l'utilisateur son identité et essaye de répondre à la question, "est-ce la personne X? ". Dans une application de vérification l'utilisateur annonce son identité par l'intermédiaire d'un mot de passe, d'un numéro d'identification, d'un nom d'utilisateur, ou toute combinaison des trois. Le système sollicite également une information biométrique provenant de l'utilisateur, et compare la donnée caractéristique obtenue à partir de l'information entrée, avec la donnée enregistrée correspondante à l'identité prétendue, c'est une comparaison un à un (1:1). Le système trouvera ou ne trouvera pas d'appariement entre les deux. La vérification est communément employée dans des applications de contrôle d'accès et de paiement par authentification [2].

La biométrie offre beaucoup plus d'avantages que les méthodes existantes d'authentification personnelle telles que les clefs, les numéros d'identification (ID), les mots de passe et les cartes magnétiques. En effet elle fournit encore plus de sûreté et de convenance ce qui engendre d'énormes avantages économiques et elle comble les grandes failles de sécurité des mots de passe.

I.3. Le marché mondial de la biométrie

Régulièrement, un rapport sur le marché de la biométrie est édité par IBG (International Biometric Group). Cette étude est une analyse complète des chiffres d'affaires, des tendances de croissance, et des développements industriels pour le marché de la biométrie actuel et futur.

La lecture de ce rapport est essentielle pour des établissements déployant la technologie biométrique, les investisseurs dans les entreprises biométriques, ou les développeurs de solutions biométriques. Le chiffre d'affaires de l'industrie biométrique incluant les applications judiciaires et celles du secteur public, se développe rapidement. Une grande partie de la croissance sera attribuable au contrôle d'accès aux systèmes d'information (ordinateur / réseau) et au commerce électronique, bien que les applications du secteur public continuent à être une partie essentielle de l'industrie.

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

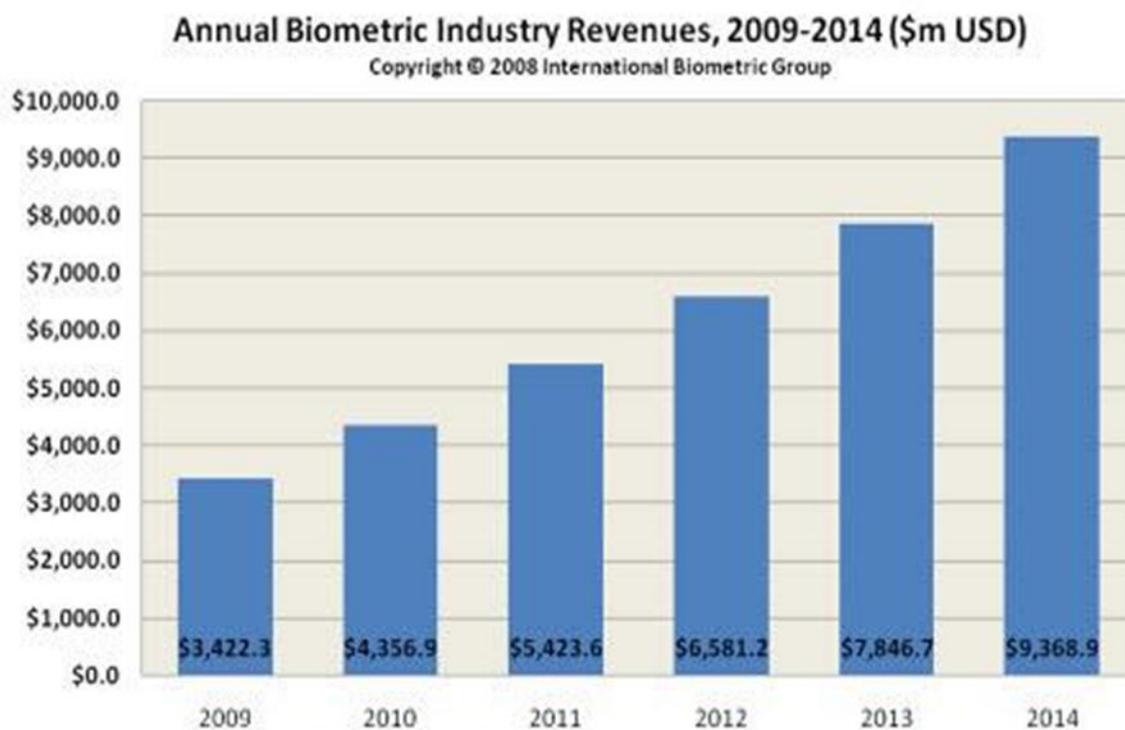


Figure1 : Evolution du marché international de la biométrie [3].

I.4. Les parts de marché par technologie

Les empreintes digitales continuent à être la principale technologie biométrique en termes de part de marché, près de 50% du chiffre d'affaires total (hors applications judiciaires). La reconnaissance du visage, avec 12% du marché (hors applications judiciaires), dépasse la reconnaissance de la main, qui avait avant la deuxième place en termes de source de revenus après les empreintes digitales.

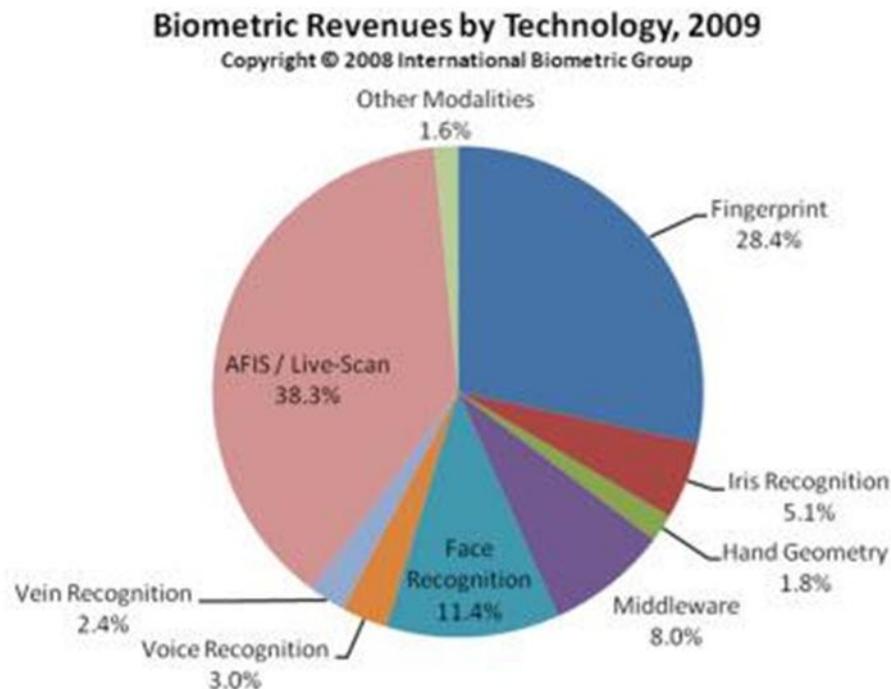


Figure 2 : Parts de marché des différentes méthodes biométriques [3]

I.5. Les techniques biométriques

Il existe plusieurs techniques biométriques utilisées dans plusieurs applications et secteurs, on peut en distinguer deux catégories :

I.5.1. L'analyse morphologique (physiologique)

Elle est basée sur l'identification de traits physiques particuliers qui, pour toute personne, sont uniques et permanents. Cette catégorie regroupe l'iris de l'œil, le réseau veineux de la rétine, la forme de la main, les empreintes digitales, les traits du visage, les veines de la main, etc.

I.5.2. L'analyse comportementale

Elle se base sur l'analyse de certains comportements d'une personne. Cette catégorie regroupe la reconnaissance vocale, la dynamique de frappe au clavier, la dynamique de la signature, l'analyse de la démarche, etc. Il existe, par ailleurs, une autre catégorie qui est l'étude des traces biologiques telles que : l'ADN, le sang, la salive, l'urine, l'odeur, etc.

I.6. Architecture d'un système biométrique

Il existe toujours au moins deux modules dans un système biométrique: le module d'apprentissage et celui de reconnaissance [4] [5]. Le troisième module est le module d'adaptation. Pendant l'apprentissage, le système va acquérir une ou plusieurs mesures biométriques qui serviront à construire un modèle de l'individu. Ce modèle de référence servira de point de comparaison lors de la reconnaissance. Le modèle pourra être réévalué après chaque utilisation grâce au module d'adaptation.

I.6.1. Module d'apprentissage

Au cours de l'apprentissage, la caractéristique biométrique est tout d'abord mesurée grâce à un capteur; on parle d'acquisition ou de capture. En général, cette capture n'est pas directement stockée et des transformations lui sont appliquées. En effet, le signal contient de l'information inutile à la reconnaissance et seuls les paramètres pertinents sont extraits. Le modèle est une représentation compacte du signal qui permet de faciliter la phase de reconnaissance, mais aussi de diminuer la quantité de données à stocker.

Il est à noter que la qualité du capteur peut grandement influencer les performances du système. Meilleure est la qualité du système d'acquisition, moins il y aura de prétraitements à effectuer pour extraire les paramètres du signal.

Cependant, les capteurs de qualité sont en général coûteux et leur utilisation est donc limitée à des applications de haute sécurité pour un public restreint. Le modèle peut être stocké dans une base de données.

I.6.2. Module de reconnaissance

Au cours de la reconnaissance, la caractéristique biométrique est mesurée et un ensemble de paramètres est extrait comme lors de l'apprentissage (figure 3). Le capteur utilisé doit avoir des propriétés aussi proches que possibles du capteur utilisé durant la phase d'apprentissage. Si les deux capteurs ont des propriétés trop différentes, il faudra en général appliquer une série de prétraitements supplémentaires pour limiter la dégradation des performances. La suite de la reconnaissance sera différente suivant le mode opératoire du système : identification ou vérification.

En mode identification, le système doit deviner l'identité de la personne. Il répond donc à une question de type : « Qui suis-je ? ». Dans ce mode, le système compare le signal mesuré avec les différents modèles contenus dans la base de données (problème de type 1 : N). En général, lorsque l'on parle d'identification, on suppose que le problème est fermé, c'est-à-dire que toute personne qui utilise le système possède un modèle dans la base de données.

En mode vérification, le système doit répondre à une question de type : « Suis-je bien la personne que je prétends être ? ».

L'utilisateur propose une identité au système et le système doit vérifier que l'identité de l'individu est bien celle proposée. Il suffit donc de comparer le signal avec un seul des modèles présents dans la base de données (problème de type 1 : 1). En mode vérification, on parle de problème ouvert puisque l'on suppose qu'un individu qui n'a pas de modèle dans la base de données (imposteur) peut chercher à être reconnu.

Identification et vérification sont donc deux problèmes différents. L'identification peut-être une tâche redoutable lorsque la base de données contient des milliers, voire des millions d'identités, tout particulièrement lorsqu'il existe des contraintes de type « temps réel » sur le

système. Ces difficultés sont analogues à celles que connaissent par exemple les systèmes d'indexation de documents multimédia.

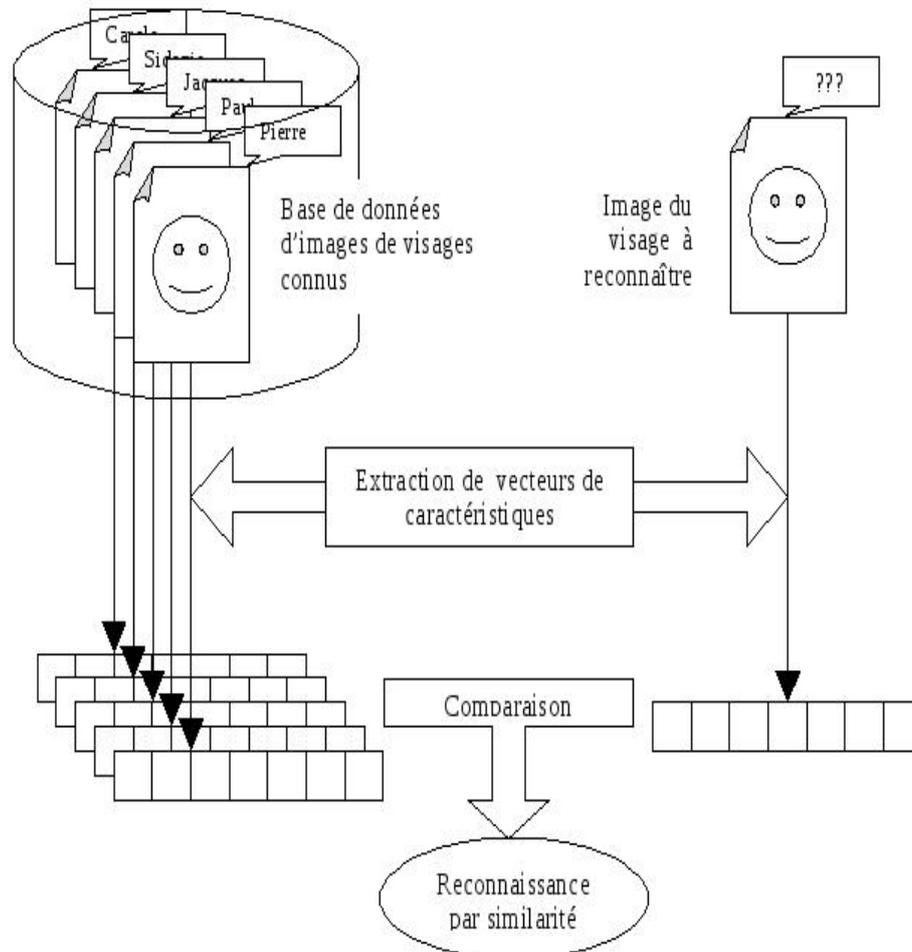


Figure 3 : Phase de reconnaissance d'un système de reconnaissance facial

I.6.3 .Module d'adaptation

Pendant la phase d'apprentissage, le système biométrique ne capture souvent que quelques instances d'un même attribut afin de limiter la gêne pour l'utilisateur. Il est donc difficile de construire un modèle assez général capable de décrire toutes les variations possibles de cet attribut. De plus, les caractéristiques de cette biométrie ainsi que ses conditions d'acquisition peuvent varier. L'adaptation est donc nécessaire pour maintenir voire améliorer la performance d'un système utilisation après utilisation. L'adaptation peut se faire en mode supervisé ou non-supervisé mais le second mode est de loin le plus utile en pratique. Si un utilisateur est identifié par le module de reconnaissance, les paramètres extraits du signal serviront alors à ré-estimer son modèle. En général, le taux d'adaptation dépend du degré de confiance du module de reconnaissance dans l'identité de l'utilisateur. Bien entendu, l'adaptation non-supervisée peut poser problème en cas d'erreurs du module de

reconnaissance. L'adaptation est quasi indispensable pour les caractéristiques non permanentes comme la voix [6].

I.7. Présentation de quelques technologies biométriques

Aucune biométrie unique ne pouvait répondre efficacement aux besoins de toutes les applications d'identification. Un certain nombre de techniques biométriques ont été proposées, analysées, et évaluées. Chaque biométrie a ses forces et ses limites, et en conséquence, chaque biométrie est utilisée dans une application particulière. Pour les caractéristiques physiques, nous décrivons la reconnaissance de visage, les empreintes digitales, la géométrie de la main et l'iris. Pour les caractéristiques comportementales, nous décrivons les biométries basées sur la voix et la signature.

Il existe d'autres méthodes biométriques basées sur les veines de la main, l'A.D.N, l'odeur corporelle, la forme de l'oreille, la forme des lèvres, le rythme de frappe sur un clavier, la démarche, qui ne seront pas développées dans ce chapitre.

I.7.1. Les empreintes digitales

A l'heure actuelle la reconnaissance des empreintes digitales est la méthode biométrique la plus utilisée. Les empreintes digitales sont composées de lignes localement parallèles présentant des points singuliers (minuties) et constituent un motif unique, universel et permanent. Pour obtenir une image de l'empreinte d'un doigt, les avancées technologiques ont permis d'automatiser la tâche au moyen de capteurs intégrés, remplaçant ainsi l'utilisation classique de l'encre et du papier. Ces capteurs fonctionnant selon différents mécanismes de mesure (pression, champ électrique, température) permettent de mesurer l'empreinte d'un doigt fixe positionné sur ce dernier (capteur matriciel) ou en mouvement (capteurs à balayage).

L'image d'empreinte d'un individu est capturée à l'aide d'un lecteur d'empreinte digitale puis les caractéristiques sont extraites de l'image puis un modèle est créé. Si des précautions appropriées sont suivies, le résultat est un moyen très précis d'authentification.

Les techniques d'appariement des empreintes digitales peuvent être classées en deux catégories : les techniques basées sur la détection locale des minuties et les techniques basées sur la corrélation. L'approche basée sur les minuties consiste à trouver d'abord les points de minuties puis trace leurs emplacements sur l'image du doigt (figure 4).

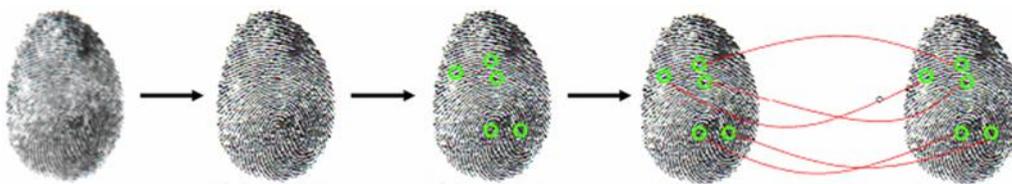


Figure 4. : Le processus de reconnaissance par empreinte digitale

Cependant, il y a quelques difficultés avec cette approche lorsque l'image d'empreinte digitale est d'une qualité médiocre, car l'extraction précise des points de minutie est difficile. Cette méthode ne tiens pas en compte la structure globale de crêtes et de sillons.

Les méthodes basées sur la corrélation sont capables de surmonter les problèmes de l'approche fondée sur les minuties. Ces méthodes utilisent la structure globale de l'empreinte, mais les résultats sont moins précis qu'avec les minuties. De plus, les techniques de

corrélation sont affectées par la translation et rotation de l'image de l'empreinte. C'est pour cela que les deux approches sont en général combinées pour augmenter les performances du système.

I.7.2.La voix

De tous les traits humains utilisés dans la biométrie, la voix est celle que les humains apprennent à reconnaître dès le plus jeune âge. Les systèmes de reconnaissance de locuteur peuvent être divisés en deux catégories : les systèmes dépendant du texte prononcé et les systèmes indépendants du texte. Dans le premier cas, l'utilisateur est tenu d'utiliser un texte (un mot ou une phrase) fixe prédéterminé au cours des séances d'apprentissage et de reconnaissance. Alors que, pour un système indépendant du texte le locuteur parle librement sans texte prédéfini.

Cette dernière catégorie est plus difficile, mais elle est utile dans le cas où l'on a besoin de reconnaître un locuteur sans sa coopération. La recherche sur la reconnaissance de locuteur est en pleine croissance, car elle ne nécessite pas de matériel cher, puisque la plupart des ordinateurs personnels de nos jours sont équipés d'un microphone. Toutefois, la mauvaise qualité et le bruit ambiant peuvent influencer la vérification et par suite réduire son utilisation dans les systèmes biométriques. Dans un système de reconnaissance de locuteur le signal est premièrement mesuré puis décomposé en plusieurs canaux de fréquences passe-bande. Ensuite, les caractéristiques importantes du signal vocal sont extraites de chaque bande.

Parmi les caractéristiques les plus communément utilisées sont les coefficients Cepstraux. Ils sont obtenus par le logarithme de la transformée de Fourier du signal vocal dans chaque bande. Finalement, la mise en correspondance des coefficients Cepstraux permet de reconnaître la voix. Dans cette étape, généralement on fait appel à des approches fondées sur les modèles de Markov cachés, la quantification vectorielle, ou la déformation temps dynamique.



Figure 5: Spectre d'un signal voix

I.7.3.L'iris

L'utilisation de l'iris comme caractéristique biométrique unique de l'homme a donné lieu à une technologie d'identification fiable et extrêmement précise. L'iris est la région, sous forme d'anneau, située entre la pupille et le blanc de l'œil, il est unique. L'iris a une structure extraordinaire et offre de nombreuses caractéristiques de texture qui sont uniques pour chaque individu. Les algorithmes utilisés dans la reconnaissance de l'iris sont si précis que la planète toute entière pourrait être inscrite dans une base de données de l'iris avec peu d'erreurs d'identification.

L'image de l'iris est généralement capturée à l'aide d'une caméra standard. Cependant, cette étape de capture implique une coopération de l'individu. De plus, il existe plusieurs contraintes liées à l'utilisation de cette technologie. Par exemple, il faut s'assurer que l'iris de

l'individu est à une distance fixe et proche du dispositif de capture, ce qui limite l'utilisation de cette technologie.

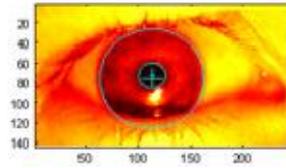


Figure 6 : Photo d'iris

I.7.4. La signature

La vérification de la signature analyse la façon dont un utilisateur signe son nom. Les caractéristiques dynamiques de la signature comme la vitesse et la pression, sont aussi importantes que la forme géométrique de la signature. Il existe deux approches pour vérifier la signature statique et dynamique. Dans la vérification de signature statique, seules les formes géométriques de la signature sont utilisées pour authentifier une personne. Dans cette approche, en règle générale, la signature est normalisée à une taille connue ensuite décomposée en éléments simples.

La forme et les relations de ses éléments sont utilisées comme caractéristiques d'identification. Quant à la deuxième approche de la vérification de signature elle utilise, en plus de la forme géométrique, les caractéristiques dynamiques telles que l'accélération, la vitesse et les profils de trajectoire de la signature. Il est à noter que la signature est une biométrie comportementale, elle évolue dans le temps et est influencée par les conditions physiques et émotionnelles de la personne.



Figure 7 : Signature

I.7.5. La géométrie de la main

La géométrie de la main est une technologie biométrique récente. Comme son nom l'indique, elle consiste à analyser et à mesurer la forme de la main, c'est-à-dire mesurer la longueur, la largeur et la hauteur de la main d'un utilisateur et de créer une image 3-D. Des LEDs infrarouges et un appareil-photo numérique sont utilisés pour acquérir les données de la main.

Cette technologie offre un niveau raisonnable de précision et est relativement facile à utiliser. Cependant elle peut être facilement trompée par des jumeaux ou par des personnes ayant des formes de la main proches. Les utilisations les plus populaires de la géométrie de la main comprennent l'enregistrement de présence et le contrôle d'accès. Par contre, les systèmes de

capture de la géométrie de la main sont relativement grands et lourds, ce qui limite leur utilisation dans d'autres applications comme l'authentification dans les systèmes embarqués : téléphones portables, voitures, ordinateurs portables, etc.



Figure 8 : Dispositif de reconnaissance par géométrie de la main

I.7.6. Le visage

Nos visages sont des objets complexes avec des traits qui peuvent varier dans le temps. Cependant, les humains ont une capacité naturelle à reconnaître les visages et d'identifier les personnes dans un coup d'œil. Bien sûr, notre capacité de reconnaissance naturelle s'étend au-delà de la reconnaissance du visage, où nous sommes également en mesure de repérer rapidement des objets, des sons ou des odeurs. Malheureusement, cette aptitude naturelle n'existe pas dans les ordinateurs. C'est ainsi qu'est né le besoin de simuler artificiellement la reconnaissance afin de créer des systèmes intelligents autonomes. Simuler notre capacité naturelle de la reconnaissance des visages dans les machines est une tâche difficile mais pas impossible. Tout au long de notre vie, de nombreux visages sont vus et conservés naturellement dans nos mémoires formant une sorte de base de données.

La reconnaissance des visages par ordinateur nécessite également une base de données qui est habituellement construite en utilisant des images du visage ou parfois des images différentes d'une même personne pour tenir compte des variations dans les traits du visage. Les systèmes actuels de reconnaissance du visage sont composés d'un module d'acquisition d'images avec une caméra. Il procède d'abord à une détection du visage dans l'image acquise. Ensuite, l'image du visage détectée est normalisée pour être transmise au module de reconnaissance qui va la traiter en utilisant des algorithmes afin d'extraire une signature du visage.

Finalement, cette signature est comparée, à l'aide d'un classificateur, avec les signatures déjà existantes dans une base de données locale, afin d'identifier l'individu en question. Les différentes étapes de la reconnaissance des visages sont illustrées dans la **figure 9**. Durant la dernière décennie de recherche, la performance des systèmes de reconnaissance du visage s'est grandement améliorée, mais les résultats sont encore loin d'être parfaits. Ces systèmes sont très sensibles aux variations d'illumination et de pose.

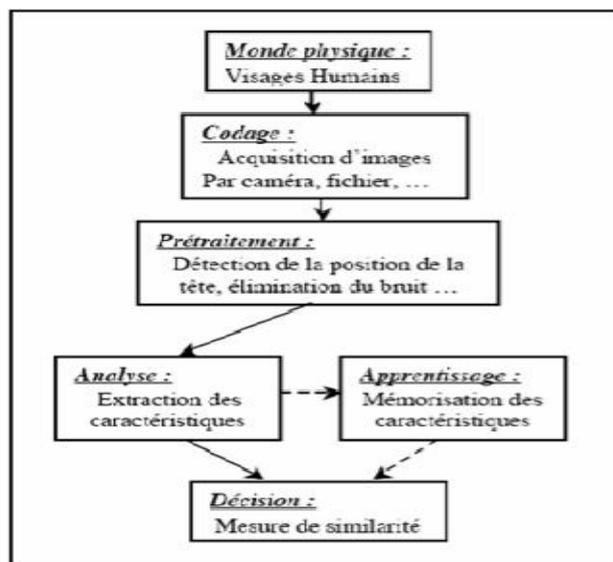


Figure 9 : Schéma synoptique d'un système de reconnaissance faciale

Après cet aperçu rapide de quelques systèmes biométriques les plus connus, nous allons détailler, dans la section suivante, le fonctionnement des systèmes de reconnaissance de visages avec les méthodes mathématiques mises en œuvre pour reconnaître un visage depuis une image fixe.

I.8 .Reconnaissance du visage

La reconnaissance faciale est une tâche que les humains effectuent naturellement et sans effort dans leurs vies quotidiennes. La grande disponibilité d'ordinateurs puissants et peu onéreux ainsi que des systèmes informatiques embarqués ont suscité un énorme intérêt dans le traitement automatique des images et des vidéos numériques au sein de nombreuses applications, incluant l'identification biométrique, la surveillance, l'interaction homme-machine et la gestion de données multimédia.

La reconnaissance faciale, en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de la recherche, ceci étant dû aux avancées rapides dans des technologies telles que les appareils photo numériques, Internet et les dispositifs mobiles, le tout associé à des besoins en sécurité sans cesse en augmentation.

La reconnaissance faciale possède plusieurs avantages sur les autres technologies biométriques : elle est naturelle, non intrusive et facile à utiliser.

Idéalement, un système de reconnaissance faciale doit pouvoir identifier des visages présents dans une image ou une vidéo de manière automatique. Le système peut opérer dans les deux modes suivants : authentification ou identification ; on peut également noter qu'il existe un autre type de scénario de reconnaissance faciale mettant en jeu une vérification sur une liste de surveillance ("watch-list"), où un individu est comparé à une liste restreinte de suspects.

Le principe de fonctionnement de base d'un système de reconnaissance faciale (**Figure 9**) peut être résumé en quatre étapes : les deux premières s'effectuent en amont du système (détection [7] et normalisation du visage [8]) et les deux dernières représentent la reconnaissance (extraction et comparaison des caractéristiques).

La performance des systèmes de reconnaissance faciale s'est significativement améliorée depuis les premiers travaux qui ont été menés dans les années 1960-70 [9] [10] et de nombreux algorithmes de reconnaissance du visage ont été proposés depuis.

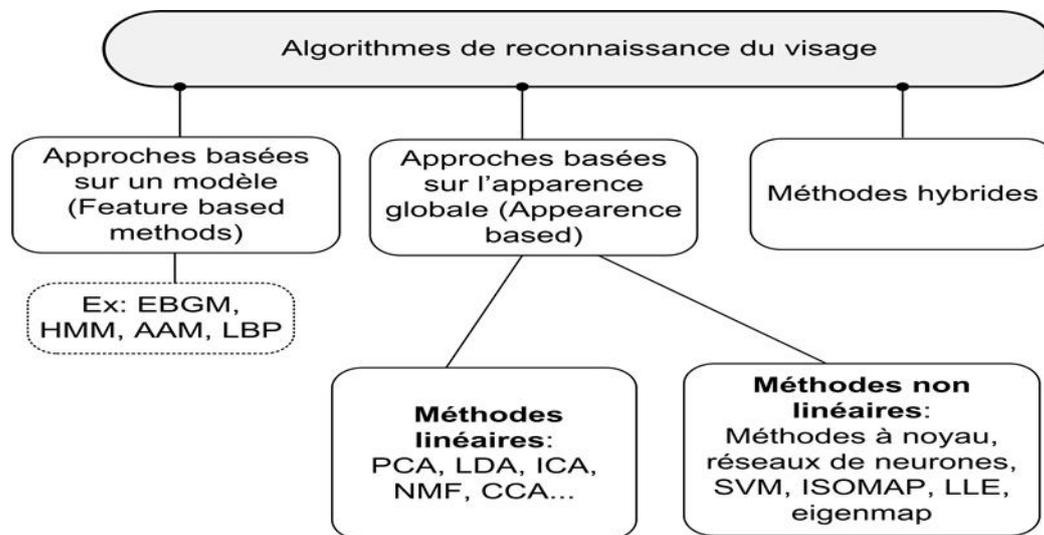


Figure 10 : Les différentes approches d'un système de reconnaissance biométrique.

Les méthodes de reconnaissance de visages peuvent être classées en trois grandes approches (Figure10). Une approche globale dans laquelle on analyse le visage (l'image pixellisée du visage) dans son entier, une approche locale basée sur un modèle, dans laquelle le système essaie de détecter, regrouper et reconnaître les différents éléments constitutifs du visage tel que le nez, les yeux et la bouche. Enfin, il existe des méthodes hybrides qui combinent les deux approches précédentes.

I.8.1. Approches globales

La particularité des algorithmes basés sur l'apparence c'est l'utilisation directe des valeurs d'intensité des pixels de l'image entière du visage comme caractéristiques sur lesquelles la décision de reconnaissance sera fondée. L'inconvénient de cette approche c'est la taille importante des données à traiter. En effet, dans ces méthodes, une image de taille $n=p \times q$ pixels, est représentée par un vecteur de la même taille, dans un espace de grande dimension.

Dans le but de réduire la dimension des données de départ, plusieurs méthodes ont été proposées en vue de transformer le vecteur des données originales dans un autre espace, de faible dimension, sans pour autant éliminer les informations discriminatoires qui seront utilisées durant l'étape de classification. La transformation des vecteurs de données peut être réalisée de manière linéaire ou non-linéaire.

Par ailleurs, les méthodes globales peuvent être à leur tour classifiées en deux grandes catégories à savoir les méthodes linéaires et les méthodes non linéaires. Ces méthodes appelées aussi méthodes de projections sont basées sur la décomposition de l'image sur un sous espace réduit et sur la recherche d'un vecteur de caractéristiques optimal décrivant l'image du visage à reconnaître. Ainsi, les vecteurs images originaux de grande taille seront projetés dans un espace de plus faible dimension. Les nombreuses méthodes de projection

existantes privilégient la bonne représentation des données suivant un point de vue. Par exemple, l'analyse en composantes principales ou le classical multidimensional scaling (MDS) [12] permettent de maximiser la variance des données dans la représentation.

La représentation exprime alors la forme générale du jeu de données. Les méthodes supervisées comme l'analyse discriminante projettent les données de façon à valoriser la séparation des classes. Le Locally Linear Embedding (LLE) [40], le Laplacian Eigenmaps [32] et le Hessian-based Locally Linear Embedding (HLLE) [40] supposent que les données sont situées sur une surface localement linéaire.

Ces méthodes permettent alors le dépliement de l'espace à l'aide de projections linéaires locales. Nous allons présenter ces approches non supervisées linéaires et non-linéaires basées sur les noyaux ou sur l'apprentissage de variétés. Les méthodes supervisées seront aussi présentées comme l'analyse discriminante linéaire (LDA).

Les techniques linéaires projettent linéairement les données d'un espace de grande dimension (par exemple, l'espace de l'image originale) sur un sous-espace de dimension inférieure. Malheureusement, ces techniques sont incapables de préserver les variations non convexes des variétés (géométriques donc au sens mathématique du terme) de visages afin de différencier des individus. Dans un sous-espace linéaire, les distances euclidiennes et plus généralement les distances de Mahalanobis, qui sont normalement utilisées pour faire comparer des vecteurs de données, ne permettent pas une bonne classification entre les classes de formes "visage" et "non-visage" et entre les individus eux-mêmes. Ce facteur crucial limite le pouvoir des techniques linéaires pour obtenir une détection et une reconnaissance du visage très précises.

La technique linéaire la plus connue et sans aucun doute l'analyse en composantes principales (PCA), également appelée transformée de Karhunen-Loeve. Le PCA fut d'abord utilisé afin de représenter efficacement des images de visages humains [11]. En 1991, cette technique a été reprise dans le cadre plus spécifique de la reconnaissance faciale sous le nom de méthode des Eigenfaces [12].

Cependant, le PCA classique nécessite que les images de visage soient mises sous formes de vecteurs, ce qui a pour effet de détruire la structure géométrique de l'image.

Pour ne pas perdre les informations de voisinage lors du passage de l'image en vecteur, une méthode PCA bidimensionnelle (2-D PCA) [13] a été étudiée. Cette méthode prend en entrée des images et non plus des vecteurs.

Il existe d'autres techniques également construites à partir de décompositions linéaires comme l'analyse discriminante linéaire (LDA) [14] ou encore l'analyse en composantes indépendantes (ICA) [15].

Tandis que le PCA construit un sous-espace pour représenter de manière "optimale" (mathématiquement parlant) seulement "l'objet" visage, le LDA construit un sous-espace discriminant pour distinguer de façon "optimale" les visages de différentes personnes. Elle permet donc d'effectuer une véritable séparation de classes. Des études comparatives (par exemple, [14]) montrent que les méthodes basées sur le LDA donne généralement de meilleurs résultats que les méthodes basées sur le PCA.

L'algorithme ICA, quant à lui, est une généralisation de l'algorithme PCA avec lequel il coïncide dans le cas de données gaussiennes. L'algorithme ICA est basé sur le concept intuitif de contraste et permet d'éliminer la redondance statistique des données de départ. Ce dernier a été rendu célèbre notamment avec l'expérience de la "cocktail party" [16] qui met en avant la résolution d'un problème de séparation (décorrélation) de sources audio. Bartlett et al. ont

fourni deux architectures différentes pour l'ICA : une première architecture (ICA I) qui construit une base d'images statistiquement indépendantes et une deuxième architecture (ICA II) qui fournit une représentation en code factoriel des données.

Bien que ces méthodes globales linéaires basées sur l'apparence évitent l'instabilité des toutes premières méthodes géométriques qui ont été mises au point, elles ne sont pas assez précises pour décrire les subtilités des variétés (géométriques) présentes dans l'espace de l'image originale. Ceci est dû à leurs limitations à gérer la non-linéarité en reconnaissance faciale : les déformations de variétés non linéaires peuvent être lissées et les concavités peuvent être remplies, causant des conséquences défavorables.

Afin de pouvoir traiter ce problème de non-linéarité en reconnaissance faciale, de telles méthodes linéaires ont été étendues à des techniques non linéaires basées sur la notion mathématique de noyau ("kernel") comme le Kernel PCA [17] et le Kernel LDA [18]. Ici, une projection non linéaire (réduction de dimension) de l'espace de l'image sur l'espace de caractéristiques ("feature space") est effectuée ; les variétés présentes dans l'espace de caractéristiques résultant deviennent simples, de même que les subtilités des variétés qui sont préservées. Bien que les méthodes basées sur le noyau peuvent atteindre une bonne performance sur les données d'entraînement, il ne peut pas en être de même pour de nouvelles données en raison de leur plus grande flexibilité ; contrairement aux méthodes linéaires.

I.8.2. Méthodes locales

Les méthodes locales, basées sur des modèles, utilisent des connaissances a priori que l'on possède sur la morphologie du visage et s'appuient en général sur des points caractéristiques de celui-ci. Kanade présenta un des premiers algorithmes de ce type [10] en détectant certains points ou traits caractéristiques d'un visage puis en les comparant avec des paramètres extraits d'autres visages. Ces méthodes constituent une autre approche pour prendre en compte la non-linéarité en construisant un espace de caractéristiques local et en utilisant des filtres d'images appropriés, de manière à ce que les distributions des visages soient moins affectées par divers changements.

Les approches Bayésiennes (comme la méthode BIC [19]), les machines à vecteurs de support (SVM) [20], la méthode des modèles actifs d'apparence (AAM) ou encore la méthode "local binary pattern" (LBP) [21] ont été utilisées dans ce but.

Toutes ces méthodes ont l'avantage de pouvoir modéliser plus facilement les variations de pose, d'éclairage et d'expression par rapport aux méthodes globales. Toutefois, elles sont plus lourdes à utiliser puisqu'il faut souvent placer manuellement un assez grand nombre de points sur le visage alors que les méthodes globales ne nécessitent de connaître que la position des yeux afin de normaliser les images, ce qui peut être fait automatiquement et de manière assez fiable par un algorithme de détection [22].

I.8.3.Méthodes hybrides

Les méthodes hybrides permettent d'associer les avantages des méthodes globales et locales en combinant la détection de caractéristiques géométriques (ou structurales) avec l'extraction de caractéristiques d'apparence locales. Elles permettent d'augmenter la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales.

L'analyse de caractéristiques locales (LFA) [23] et les caractéristiques extraites par ondelettes de Gabor (comme l'Elastic Bunch Graph Matching, EBGM), sont des algorithmes hybrides typiques.

Plus récemment, l'algorithme Log Gabor PCA (LG-PCA) [24] effectue une convolution avec des ondelettes de Gabor orientées autour de certains points caractéristiques du visage afin de créer des vecteurs contenant la localisation et la valeur d'amplitudes énergétiques locales ; ces vecteurs sont ensuite envoyés dans un algorithme PCA afin de réduire la dimension des données.

I.9. Principales difficultés de la reconnaissance de visage

Pour le cerveau humain, le processus de la reconnaissance de visages est une tâche visuelle de haut niveau. Bien que les êtres humains puissent détecter et identifier des visages dans une scène sans beaucoup de peine, construire un système automatique qui accomplit de telles tâches représente un sérieux défi. Ce défi est d'autant plus grand lorsque les conditions d'acquisition des images sont très variables. Il existe deux types de variations associées aux images de visages : inter et intra sujet. La variation inter-sujet est limitée à cause de la ressemblance physique entre les individus. Par contre la variation intra-sujet est plus vaste. Elle peut être attribuée à plusieurs facteurs que nous analysons ci-dessous.

1.9.1. Changement d'illumination

L'apparence d'un visage dans une image varie énormément en fonction de l'illumination de la scène lors de la prise de vue (voir **figure 11**). Les variations d'éclairage rendent la tâche de reconnaissance de visage très difficile. En effet, le changement d'apparence d'un visage dû à l'illumination, se révèle parfois plus critique que la différence physique entre les individus, et peut entraîner une mauvaise classification des images d'entrée. Ceci a été expérimentalement observé dans Adini et al [25] où les auteurs ont utilisé une base de données de 25 individus. L'identification de visage dans un environnement non contrôlé reste donc un domaine de recherche ouvert. Les évaluations FRVT [26] ont révélé que le problème de variation d'illumination constitue un défi majeur pour la reconnaissance faciale.



Figure 11 : Exemple de variation d'éclairage

I.9.2. Variation de pose

Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images. Cette difficulté a été démontrée par des tests d'évaluation élaborés sur les bases FERET et FRVT [26]. La variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Quand le visage est de profil dans le plan image (orientation $< 30^\circ$), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux).

Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible (voir figure 12).



Figure 12 : Exemples de variation de poses.

I.9.3. Expressions faciales

Un autre facteur qui affecte l'apparence du visage est l'expression faciale (voir fig13). La déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. L'information faciale se situant dans la partie supérieure du visage reste quasi invariable. Elle est généralement suffisante pour effectuer une identification. Toutefois, étant donné que l'expression faciale modifie l'aspect du visage, elle entraîne forcément une diminution du taux de reconnaissance. L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui reste non résolu. L'information temporelle fournit une connaissance additionnelle significative qui peut être utilisée pour résoudre ce problème [27].



Figure 13 : Exemples de variation d'expressions.

I.9.4. Présence ou absence des composants structurels

La présence des composants structurels telle que la barbe, la moustache, ou bien les lunettes peut modifier énormément les caractéristiques faciales telles que la forme, la couleur, ou la taille du visage. De plus, ces composants peuvent cacher les caractéristiques faciales de

base causant ainsi une défaillance du système de reconnaissance. Par exemple, des lunettes opaques ne permettent pas de bien distinguer la forme et la couleur des yeux, et une moustache ou une barbe modifie la forme du visage.

I.9.5. Occultations partielles

Le visage peut être partiellement masqué par des objets dans la scène, ou par le port d'accessoire tels que lunettes, écharpe...etc. Dans le contexte de la biométrie, les systèmes proposés doivent être non intrusifs c'est-à-dire qu'on ne doit pas compter sur une coopération active du sujet. Par conséquent, il est important de savoir reconnaître des visages partiellement occultés. Gross et al [27] ont étudié l'impact du port de lunettes de soleil, et du cache-nez occultant la partie inférieure du visage sur la reconnaissance faciale. Ils ont utilisé la base de données AR [28]. Leurs résultats expérimentaux semblent indiquer que, dans ces conditions, les performances des algorithmes de reconnaissance restent faibles.

I.10. Evaluation de performance

La performance d'un système d'identification peut se mesurer principalement à l'aide de trois critères : sa précision, son efficacité (vitesse d'exécution) et le volume de données qui doit être stocké pour chaque locuteur. Nous nous concentrerons dans cette section sur le premier aspect. Comme nous l'avons vu précédemment, l'identification et la vérification sont des modes opératoires différents. Elles nécessitent donc des mesures de précision différentes que nous étudierons dans les deux sous-sections suivantes.

I.10.1. Evaluation de l'identification

Le taux d'identification est la mesure la plus couramment utilisée mais il n'est pas toujours suffisant. En effet, en cas d'erreur, il peut être utile de savoir si le bon choix se trouve dans les N premiers. On trace alors le score cumulé (cumulative match score) qui représente la probabilité que le bon choix se trouve parmi les N premiers [29].

Dans le cas où il existe plusieurs modèles pour chaque individu dans la base de données, les mesures classiques des systèmes de recherche dans une base de données (database retrieval system) peuvent être utilisées. La précision (precision) est le rapport entre le nombre de modèles correctement retrouvés par le système dans la base de données et le nombre total de modèles retrouvés. Le rappel (recall) est le rapport entre le nombre de modèles correctement retrouvés dans la base de données et le nombre total de modèles qui auraient dû être retrouvés.

I.10.2. Evaluation de la vérification

Lorsqu'un système fonctionne en mode vérification, celui-ci peut faire deux types d'erreurs. Il peut rejeter un utilisateur légitime et dans ce premier cas on parle de faux rejet (false rejection). Il peut aussi accepter un imposteur et on parle dans ce second cas de fausse acceptation (false acceptance). La performance d'un système se mesure donc à son taux de faux rejet (False Rejection Rate ou FRR) et à son taux de fausse acceptation (False Acceptance Rate ou FAR).

La vérification est un problème de décision similaire à la détection d'un signal dans le bruit en théorie de l'information. Il peut être formulé de la manière suivante. Soient H_0 l'hypothèse : "la capture C provient d'un imposteur" et H_1 l'hypothèse : "la capture C provient de

l'utilisateur légitime". Il faut donc choisir l'hypothèse la plus probable. On considère que la capture C provient d'un utilisateur légitime si :

$$P(H_1/C) > P(H_0/C) \quad (I.1)$$

En appliquant la loi de Bayes on obtient :

$$\frac{P(H_1/C)P(H_1)}{P(C)} > \frac{P(H_0/C)P(H_0)}{P(C)} \quad (I.2)$$

et donc :

$$\frac{P(C/H_1)}{P(C/H_0)} > \frac{P(H_0)}{P(H_1)} \quad (I.3)$$

Le taux de vraisemblance $\frac{P(C/H_1)}{P(C/H_0)}$ Est comparé à un seuil θ appelé seuil de décision. Les valeurs $P(H_0)$ et $P(H_1)$ qui représentent respectivement la probabilité pour qu'un imposteur ou un utilisateur légitime essayent d'accéder au système sont des valeurs difficile à estimer.

Nous avons représenté sur la (figure.14), la distribution hypothétique des taux de vraisemblance qu'obtiendraient les utilisateurs légitimes et les imposteurs d'un système de vérification donné. Les FAR et FRR sont représentés en hachuré. Idéalement, le système devrait avoir des FAR et FRR égaux à zéro. Comme ce n'est jamais le cas en pratique, il faut choisir un compromis entre FAR et FRR. Plus le seuil de décision θ est bas, plus le système acceptera d'utilisateurs légitimes mais plus il acceptera aussi d'imposteurs. Inversement, plus le seuil de décision θ est élevé, plus le système rejettera d'imposteurs mais plus il rejettera aussi d'utilisateurs légitimes.

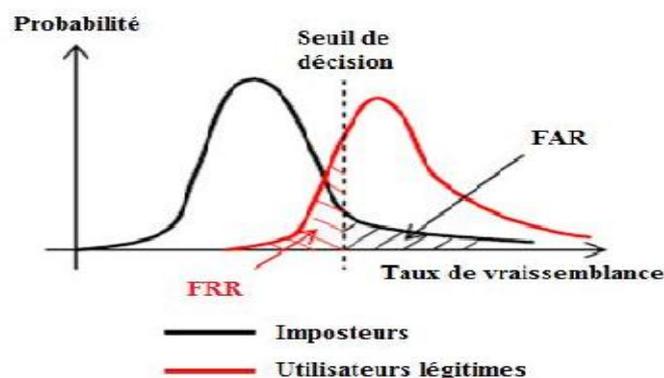


Figure 14 : Distributions des taux de vraisemblance des utilisateurs légitimes et des imposteurs d'un système biométrique

Il est donc impossible en faisant varier le seuil de décision de faire diminuer les deux types d'erreurs en même temps.

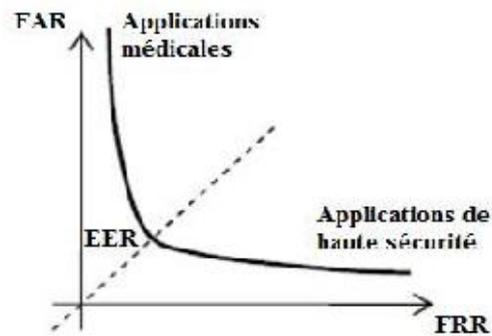


Figure 15 : Courbe ROC

La courbe dite ROC (Receiver Operating Characteristic), représentée à la figure 15, permet de représenter graphiquement la performance d'un système de vérification pour les différentes valeurs de θ . Le taux d'erreur égal (Equal Error Rate ou EER) correspond au point $FAR=FRR$, c'est-à-dire graphiquement à l'intersection de la courbe ROC avec la première bissectrice. Il est fréquemment utilisé pour donner un aperçu de la performance d'un système. Cependant, il est important de souligner que l'EER ne résume en aucun cas toutes les caractéristiques d'un système biométrique.

Le seuil θ doit donc être ajusté en fonction de l'application ciblée : haute sécurité, basse sécurité ou compromis entre les deux.

I.11 Discussion :

Dans ce chapitre, nous avons passé en revue les principales technologies biométriques. Ensuite, nous avons présenté dans un premier temps l'architecture de base d'un système biométrique pour se focaliser, dans un deuxième temps, sur la reconnaissance de visages.

Ces dernières années, de nombreuses approches pour la reconnaissance automatique de visages ont été proposées. Elles peuvent être regroupées en deux types d'approches principales. Les approches globales pour lesquelles les caractéristiques sont extraites directement de l'image du visage en entier, et les approches locales basées sur l'extraction de traits locaux du visage, tel que les yeux, le nez ou la bouche. Nous avons décrit également leurs avantages et leurs inconvénients puis dégager les problèmes liés à leur application dans un environnement réel.

Chapitre II

Techniques de détection et de
reconnaissance de visage

II.1. Préambule

Dans ce chapitre, nous présenterons un état de l'art sur les techniques de détection de visage et les différentes méthodes les plus connues de reconnaissance faciale. Enfin, nous terminerons le chapitre par une synthèse des méthodes et techniques étudiées.

II.2. Détection de visages

La détection de visages dans l'image est un traitement indispensable et crucial avant la phase de reconnaissance. En effet, le processus de reconnaissance de visages ne pourra jamais devenir intégralement automatique s'il n'a pas été précédé par une étape de détection efficace. Le traitement consiste à rechercher dans une image la position des visages et de les extraire sous la forme d'un ensemble d'images dans le but de faciliter leur traitement ultérieur.

Selon la définition proposée dans [30] un visage est considéré correctement détecté si la taille d'image extraite ne dépasse pas 20% de la taille réelle de la région faciale, et qu'elle contient essentiellement les yeux, le nez et la bouche. Elle sera ensuite affinée par un prétraitement.

II.3. Approches de la détection de visage

Il existe plusieurs méthodes pour la détection de visages. Elles peuvent être subdivisées en quatre différents types [30].

II.3.1. Approches basées sur l'apparence

Ces approches appliquent généralement des techniques d'apprentissage automatique. Ainsi, les modèles sont appris à partir d'un ensemble d'images représentatives de la variabilité de l'aspect facial. Ces modèles sont alors employés pour la détection. L'idée principale de ces méthodes est de considérer que le problème de la détection de visage est un problème de classification (visage, non-visage). Une des approches les plus connues de détection de visage est l'Eigenface [12]. Elle consiste à projeter l'image dans un espace et à calculer la distance euclidienne entre l'image et sa projection. En effet, en codant l'image dans un espace, on dégrade l'information contenue dans l'image, puis on calcule la perte d'information entre l'image et sa projection. Si cette perte d'information est grande (évaluée à partir de la distance, que l'on compare à un seuil fixé a priori), l'image n'est pas correctement représentée dans l'espace : elle ne contient pas de visage.

Cette méthode donne des résultats assez encourageants, mais le temps de calcul est très important. Dans Rowley et al. [31], les auteurs proposent un système de détection de visage basé sur la classification par des réseaux de neurones. Leur technique est divisée en deux étapes: la localisation des visages en utilisant un réseau de neurones et la vérification des résultats obtenus. Les auteurs ont construit un réseau de neurones qui, à partir d'une image prétraitée de 20x20 pixels, indique s'il s'agit d'un visage ou non. Le prétraitement consiste à égaliser l'histogramme. L'image est balayée avec des fenêtres de 20x20. Pour détecter les visages de différentes tailles, une analyse multi-résolutions est effectuée. L'extension a aussi été proposée afin de déterminer un réseau de neurones indiquant le degré de rotation d'un visage. Ainsi, le système est capable de détecter des visages ayant subi des rotations dans le plan et de différentes échelles. L'un des avantages des réseaux de neurones est leur robustesse au bruit. Malheureusement, les réseaux de neurones, sont souvent difficiles à construire. Leur structure (nombre de couches cachées pour les perceptrons par exemple) influe beaucoup sur les résultats et il n'existe pas de méthode pour déterminer automatiquement cette structure.

La phase d'apprentissage est difficile à mener puisque les exemples doivent être correctement choisis (en nombre et en configuration).

II.3.2. Approches basées sur les connaissances acquises

Ces méthodes sont basées sur la définition de règles strictes à partir des rapports entre les caractéristiques faciales. Elles s'intéressent aux parties caractéristiques du visage comme le nez, la bouche et les yeux. Ces méthodes sont conçues principalement pour la localisation de visage. Dans [32], Kotropoulous et Pitas utilisent une méthode à base de règles. Les caractéristiques du visage sont localisées à l'aide de la méthode de projection proposée par Kanade [33] pour détecter les contours d'un visage. Soit $I(x,y)$ l'intensité de la luminance du pixel (x,y) de l'image $m*n$, les projections horizontale et verticale de cette image sont définies par l'équation suivante :

$$HI(x) = \sum_{y=1}^n I(x,y) \quad \text{et} \quad VI(x) = \sum_{x=1}^m I(x,y) \quad (2.1)$$

Le profil horizontal de l'image originale est calculé en premier. Les deux minima locaux sont déterminés, ils correspondent aux bords gauche et droit du visage. Ensuite, le profil vertical est à son tour calculé. Les minima locaux de ce profil vertical correspondent aux positions de la bouche, du nez et des yeux. L'inconvénient de cette méthode est qu'elle n'arrive pas à détecter le visage lorsque ce dernier se trouve sur un arrière-plan complexe. Yang and Huang [34] quant à eux, ont étudié les évolutions des caractéristiques du visage en fonction de la résolution. Quand la résolution de l'image d'un visage est réduite progressivement, par sous-échantillonnage ou par moyenne, les traits macroscopiques du visage disparaissent. Ainsi, pour une résolution faible, la région du visage devient uniforme. Yang et Huang se sont basés sur cette observation pour proposer une méthode hiérarchique de détection de visages.

En commençant par les images à faible résolution, un ensemble de candidats de visage est déterminé à l'aide d'un ensemble de règles permettant de rechercher les régions uniformes dans une image. Les candidats de visage sont ensuite vérifiés en cherchant l'existence de traits faciaux proéminents grâce au calcul des minima locaux à des résolutions supérieures. Une caractéristique intéressante de cette technique « descendante » de recherche de zone d'intérêt (informations globales vers des informations plus détaillées) est de réduire le temps de calcul nécessaire par l'utilisation d'images sous-échantillonnées. Malheureusement, cette technique occasionne de nombreuses fausses détections et un taux faible de détection.

II.3.3 Approches basées sur le « Template-matching »

Les templates peuvent être définis soit "manuellement", soit paramétrés à l'aide de fonctions. L'idée est de calculer la corrélation entre l'image candidate et le template. Ces méthodes rencontrent encore quelques problèmes de robustesse liés aux variations de lumière, d'échelle, etc. Sinha [35] utilise un ensemble d'invariants décrivant le modèle du visage. Afin de déterminer les invariants aux changements de luminosité permettant de caractériser les différentes parties du visage (telles que les yeux, les joues, et le front); cet algorithme calcule ainsi les rapports de luminance entre les régions du visage et retient les directions de ces rapports (par exemple, la région 1 est-elle plus claire ou plus sombre que la région 2). La figure 16 montre un modèle prédéfini correspondant à 23 relations. Ces relations prédéfinies sont classifiées en 11 relations essentielles (flèches) et 12 relations confirmations (gris). Chaque flèche représente une relation entre deux régions (**figure 16**). Une relation est vérifiée si le rapport entre les deux régions qui lui correspond dépasse un seuil. Le visage est localisé si le nombre de relations essentielles et de confirmation dépasse lui aussi un seuil.

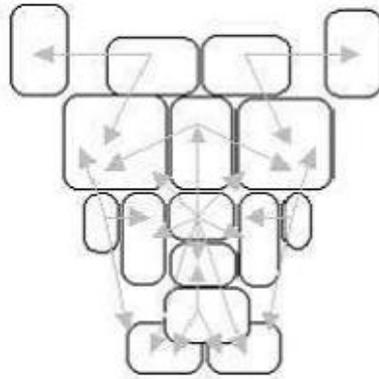


Figure16 : Modèle de visage composé de 16 régions (les rectangles) associées à 23 relations (flèches). [Sin94]

Yuille et al. [36] ont utilisé un template déformable pour modéliser les caractéristiques faciales. Ce template adapte un modèle élastique, connu a priori, aux caractéristiques faciales (ex. yeux). Dans cette approche les caractéristiques faciales sont décrites par des templates paramétriques. Une fonction d'énergie est définie pour lier les contours, les sommets et les vallées dans l'image d'entrée aux paramètres correspondants dans le template. Le meilleur ajustement du modèle élastique est trouvé en minimisant une fonction d'énergie des paramètres. Bien que leurs résultats expérimentaux démontrent une bonne performance pour le suivi de caractéristiques non rigides, un inconvénient de cette approche est que le template déformable doit être initialisé dans la proximité de l'objet d'intérêt. Pour détecter les caractéristiques faciales pour la reconnaissance de visage, Brunelli et Poggio [37] ont utilisé, pour chaque région extraite, un détecteur approprié. Ils se sont aussi inspirés de la méthode de Kanade.



Figure 17: Différentes régions utilisées pour la phase de template matching

Pour les régions yeux, nez et bouche (figure17), ils utilisent la direction du gradient vertical et horizontal. La bouche et le nez sont localisés en utilisant des stratégies similaires. La position verticale est déterminée grâce aux standards anthropométriques. D'abord, une estimation fine de leur position réelle est obtenue en cherchant les pics de la projection horizontale du gradient vertical pour le nez et les vallées de la projection horizontale de l'intensité pour la bouche.

La position des sourcils et leur épaisseur peuvent être trouvées par une analyse similaire. La recherche est une fois encore limitée à la fenêtre d'intérêt, juste au-dessus des yeux, et les sourcils sont trouvés en utilisant la carte du gradient vertical. Le détecteur du sourcil cherche les paires de pics du gradient ayant des directions opposées.

II.3.4 Approches basées sur des caractéristiques invariantes

Ces approches sont utilisées principalement pour la localisation de visage. Les algorithmes développés visent à trouver les caractéristiques structurales existantes même si la pose, le point de vue ou la condition d'éclairage changent. Puis ils emploient ces caractéristiques invariables pour localiser les visages. Nous pouvons citer deux familles de méthodes appartenant à cette approche : Les méthodes basées sur la couleur de la peau et les méthodes basées sur les caractéristiques de visage qu'on détaillera ci-dessous.

II.3.4.1 Méthodes basées sur les caractéristiques du visage

En général ces algorithmes utilisent tout d'abord une hypothèse sur la position du haut du visage ensuite l'algorithme de recherche parcourt le visage de haut en bas afin de trouver l'axe des yeux caractérisé par une augmentation soudaine de la densité de contours (mesurée par le rapport noir/blanc le long des plans horizontaux). La longueur entre le haut du visage et le plan de l'œil est alors utilisée comme une longueur de référence pour construire un « template » facial flexible. Ce « template » couvrant des caractéristiques telles que les yeux et la bouche est initialisé à partir de l'image d'entrée. La forme initiale du « template » est obtenue en utilisant la longueur anthropométrique en respectant la longueur de référence.

Le template flexible est alors ajusté par rapport aux positions finales des caractéristiques en utilisant un algorithme de réglage fin qui emploie une fonction de coût basée contour.

Bien que ces algorithmes réussissent à détecter les caractéristiques d'ethnies différentes puisqu'ils ne se basent pas sur les informations de niveaux de gris et de couleur, ils n'arrivent pas cependant à détecter correctement ces caractéristiques si l'image du visage contient des lunettes ou bien si les cheveux couvrent le front.

II.3.4.2. Méthodes basées sur l'analyse de la couleur de la peau

Les méthodes de détection basées sur l'analyse de la couleur de la peau sont des méthodes efficaces et rapides. Elles réduisent l'espace de recherche de la région visage dans l'image. De plus, la couleur de la peau est une information robuste face aux rotations, aux changements d'échelle et aux occultations partielles. Plusieurs espaces couleur peuvent être utilisés pour détecter, dans l'image, les pixels qui ont la couleur de la peau. L'efficacité de la détection dépend essentiellement de l'espace couleur choisi. Les espaces couleur les plus utilisés sont :

- L'espace RVB, mis au point en 1931 par la Commission Internationale de l'Eclairage (CIE). Il consiste à représenter l'espace des couleurs à partir de trois rayonnements monochromatiques de couleurs : Rouge-Vert-Bleu. Cet espace correspond à la façon dont les couleurs sont généralement codées informatiquement, ou plus exactement à la manière dont les écrans à tubes cathodiques (ordinateurs, TV) représentent ces couleurs.
- L'espace HSL (Hue, Saturation, Luminance), appelé aussi TSL (Teinte, Saturation, Luminance) en Français, s'appuie sur les travaux du peintre Albert H. Munsell. C'est un modèle de représentation dit "naturel", car il est proche de la perception physiologique de la couleur par l'œil humain. En effet, le modèle RGB aussi adapté soit-il pour la

représentation informatique de la couleur ou bien l'affichage sur les périphériques de sortie, ne permet pas de sélectionner facilement une couleur.

Le modèle HSL consiste à décomposer la couleur selon des critères physiologiques :

- ✓ la teinte (en Anglais Hue), correspondant à la perception de la couleur,
- ✓ la saturation, décrivant la pureté de la couleur, c'est-à-dire son caractère vif ou terne,
- ✓ la luminance, indiquant la quantité de lumière de la couleur, c'est-à-dire son aspect clair ou sombre.

Il existe d'autres modèles naturels de représentation proches du modèle HSL :

- ✓ HSB : Hue, Saturation, Brightness soit en Français Teinte, Saturation, Brilliance. La brillance décrit la perception de la lumière émise par une surface.
- ✓ HSV : Hue, Saturation, Value soit en Français Teinte, Saturation, Valeur.
- ✓ HSI : Hue, Saturation, Intensity soit en Français Teinte, Saturation, Intensité.
- ✓ HCI : Hue, Chrominance, Intensity soit Teinte, Chrominance, Intensité.
- ✓ Le modèle YCrCb est un signal non-linéaire codé à partir du signal RVB. Le paramètre Y représente la luminance (c'est-à-dire l'information en noir et blanc), tandis que Cr et Cb permettent de représenter la chrominance, c'est-à-dire l'information couleur.
- ✓ Le modèle colorimétrique $L^*a^*b^*$ (aussi connu sous le nom de CIELab), dans lequel une couleur est repérée par trois valeurs :
 - ✓ L^* , la luminance, exprimée en pourcentage (0 pour le noir à 100 pour le blanc).
 - ✓ a^* et b^* deux gammes de couleur allant, respectivement, du vert au rouge et du bleu au jaune avec des valeurs allant de -120 à +120.

Le modèle $L^*a^*b^*$ couvre ainsi l'intégralité du spectre visible par l'œil humain et le représente de manière uniforme. Il permet donc de décrire l'ensemble des couleurs visibles indépendamment de toute technologie graphique.

Après avoir vu quelques algorithmes de détection de visage, nous passerons à la phase suivante qui est la reconnaissance. Avant d'arriver à cette dernière, nous devons passer par une étape transitoire de prétraitement qui est la normalisation de l'image de visage détecté dans le but de minimiser les effets du bruit engendré durant la chaîne d'acquisition de l'image.

II.4. Prétraitement

La phase de prétraitement vient après la phase de détection. Elle permet de préparer l'image du visage de telle sorte qu'elle soit exploitable dans la phase d'enrôlement. On l'appelle aussi phase de normalisation puisqu'elle ramène à un format prédéfini toutes les images extraites de l'image brute. Elle consiste généralement en un centrage du visage dans l'image et une élimination des zones non informatives.

Pour garantir la bonne performance du système de reconnaissance de visages, il est important que toutes les images soient de taille identique, à la même échelle et au même format concernant les couleurs (par exemple, les images couleur sont parfois converties en niveaux de gris). Ceci améliore incontestablement le fonctionnement de l'étape d'extraction de signatures et par conséquent la qualité de cette dernière.

La normalisation est constituée de deux processus : géométrique et photométrique. La normalisation géométrique est nécessaire parce que la taille du visage à l'intérieur de l'image acquise peut varier en fonction de la distance entre le module d'acquisition et la personne.

Le visage doit donc être extrait à partir de l'image et une transformation géométrique, pour obtenir une taille fixe, est appliquée. L'approche standard consiste à définir l'emplacement des centres des yeux dans une position constante au sein de l'image de sortie.

L'étape de normalisation photométrique tente d'éliminer ou de réduire les effets de l'illumination de l'image.

II.4.1 Normalisation Photométrique :

Dans le domaine de la reconnaissance par le visage, un certain nombre de méthodes de normalisation de l'illumination ont été présentées. Ces méthodes peuvent être classées en deux grandes catégories : les méthodes basées sur la modification de la dynamique et des méthodes basées sur l'estimation de la réflectance de l'image. Les quatre normalisations présentées ci-dessous sont les plus utilisées dans la reconnaissance du visage et font partie de ces deux catégories.

Il s'agit de :

- ✓ L'égalisation d'histogramme : Le but de cette normalisation est de répartir uniformément les niveaux de gris de l'image afin d'améliorer son contraste.
- ✓ La correction Gamma : cette normalisation permet de réduire l'écart entre les zones sombres et claires du visage par une transformation non linéaire des niveaux de gris.
- ✓ Le MultiScale-Retinex [94] : L'objectif de cette méthode est d'estimer la partie luminance de l'image pour en déduire la réflectance.
- ✓ Le lissage anisotropique [26] : cette approche estime aussi la luminance de l'image mais en se basant sur l'anisotropie de la réflectance.

Nous avons opté pour la méthode de l'égalisation d'histogramme qui a amélioré le taux d'identification de notre système de reconnaissance.

✓ Égalisation d'histogramme

Cette normalisation appartient à la catégorie des méthodes basées sur l'amélioration de la dynamique de l'image. Le but de l'égalisation est d'harmoniser la répartition des niveaux de gris de l'image. Elle permet de tendre vers un même nombre de pixels pour chacun des niveaux de l'histogramme. Cette opération vise à augmenter les nuances dans l'image et donc son contraste (Figure18).

Plus concrètement, si n_i est le nombre de pixels à un niveau i de gris, la probabilité qu'un pixel x de l'image ait un niveau i est :

$$P(x_i) = \frac{n_i}{n}, \quad i \in 0, \dots, L \quad (2.2)$$

avec n le nombre total des pixels de l'image et L le nombre des niveaux de gris. La probabilité P représente dans ce cas l'histogramme de l'image normalisée à $[0; 1]$. Soit c la distribution cumulative de l'histogramme normalisé P , elle est donnée par

$$c(i) = \sum_{j=0}^i P(x_j) \quad (2.3)$$

L'idée est de trouver une transformation $y = T(x)$ qui, pour chaque niveau x de l'image, produira un niveau y de telle façon que la distribution cumulative des différents niveaux de l'image transformée soit linéaire. Cette fonction est définie par :

$$y_i = T(x_i) = L \frac{c(i)}{n} \quad (2.4)$$

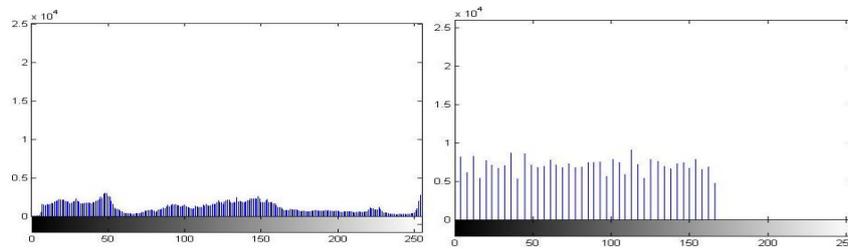
avec L le niveau maximal de gris.



Figure 18 : (a) Image RGB

(b) Image niveau de gris

(c) égalisation d'histogramme



Histogramme niveaux de gris

Histogramme égalisée

Figure 19: Histogramme de l'image avant et après égalisation

II.4.2. Normalisation géométrique

En utilisant un algorithme de reconnaissance se basant sur la réduction de l'espace, nous ne pouvons pas négliger un point très important qui est la normalisation géométrique des images de visage. Cette normalisation géométrique consiste à extraire la zone du visage de l'image originale, ensuite une rotation du visage est effectuée afin d'aligner l'axe des yeux avec l'axe horizontal. Enfin, une réduction proportionnelle à la distance entre les centres des deux yeux est appliquée. On obtient alors une image de visage dont la distance entre les centres des yeux est fixe.

Les dimensions de l'image du visage sont calculées à partir de la distance à obtenir entre les centres des deux yeux.

II.5. Les algorithmes de reconnaissance de visage

Dans cette partie de ce chapitre, on décrit quelques algorithmes classiques de reconnaissance de visages, notamment : PCA, LDA, LVQ, SVM, HMM, Template matching, EGM...

II.5.1. Dimensionnalité de l'espace visage

Une image du visage est un signal à deux dimensions, acquis par un capteur digital (caméra numérique, scanner...). Ce capteur codera la couleur ou l'intensité des différents points de l'image dans une matrice de pixels à deux dimensions. Après une normalisation de l'image et une mise à l'échelle vers une taille fixe (ex : (m, n)), l'image du visage peut être considérée comme un vecteur dans un espace multidimensionnel \mathbb{R}^m . Ce vecteur est obtenu en mettant, simplement, dans un ordre fixe les éléments de la matrice des pixels. Afin de définir une image (ou un point) dans l'espace des images, nous devons spécifier une valeur pour chaque pixel de cette image. Le nombre de points constituant cet espace devient rapidement très grand, même pour les images de petite dimension. Cette dimensionnalité pose un certain nombre de problèmes pour les algorithmes de reconnaissance, qui se basent sur cette représentation de l'image, à savoir :

- ✓ dans un contexte de la reconnaissance, travailler dans un grand espace pose un problème de complexité de calcul.
- ✓ pour les méthodes paramétriques, le nombre de paramètres à estimer peut rapidement dépasser le nombre d'échantillons d'apprentissage, ce qui pénalise l'estimation.
- ✓ pour les méthodes non paramétriques, le nombre d'exemples nécessaires afin de représenter efficacement la distribution des données peut être insuffisant.

Le cas du visage est assez particulier. Le visage est formé par des surfaces lisses et une texture régulière. Il possède une forte symétrie (dans le cas des images frontales) et il est formé à partir de mêmes objets (yeux, nez et bouche), qui ont les mêmes positions pour tous les visages. Ces spécificités donnent lieu à plusieurs constatations :

- ✓ les différents pixels de l'image du visage sont fortement corrélés.
- ✓ les images des différents visages sont aussi corrélées.
- ✓ un important nombre de points dans l'espace des images ne représentent pas des visages.

D'après ces constatations, nous pouvons déduire que les visages forment une classe particulière dans l'espace des images. L'ensemble des images de visage forment alors un sous-espace de l'espace des images. D'une façon plus formelle, si on considère Γ l'espace des images 2D d'une taille fixe, $l = n * m$. On a $\Gamma \subset \mathbb{R}^l$. Notons F l'espace des images des visages normalisés à une taille $n * m$. Nous avons $F \subset \Gamma$, $\Gamma \subset \mathbb{R}^l$.

II.5.2. Analyse en Composantes Principales

II.5.2.1. Présentation

L'analyse en composantes principales (ACP) est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension.

Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard.

Ces poids sont obtenus en projetant l'image dans un espace de visage engendré par les visages propres (Eigenfaces). Ainsi au lieu de stocker un vecteur image de taille Résolution R , on stocke un vecteur de taille $N \leq \text{nombre d'images d'apprentissage } M \ll R$. Dans l'approche ACP, la normalisation d'éclairage est toujours indispensable [38]. Les visages propres sont des images de la même taille que les images d'apprentissage et qui montrent des visages ayant un aspect fantomatique. Mathématiquement, elles sont les composantes principales de la distribution des visages, ou les vecteurs propres (Eigenvectors) de la matrice de covariance de l'ensemble des images de visage. Chaque image de visage de l'ensemble d'apprentissage peut être exactement représentée en termes de combinaison linéaire des Eigenfaces et du visage moyen [39].

Le nombre possible d'Eigenfaces est égal au nombre d'images de visage dans l'ensemble d'apprentissage. Néanmoins les visages peuvent être approximés, en utilisant seulement les meilleurs Eigenfaces (ayant les plus larges valeurs propres qui représentent en fait, la plupart de *variance* dans l'ensemble d'images de visage) qui définissent le "Low Dimensional Space", permettant ainsi de diminuer les calculs. Dans le cas où on ne prend que les N ($N < M$) grandes valeurs propres (Eigenvalues) et les vecteurs propres qui leurs sont associés, on réduit encore plus la mémoire de stockage et le temps de calcul, mais on réduit aussi les performances dû à la négligence d'une partie de l'information. C'est pour cela qu'on a pris dans notre application $N=M$, bénéficiant à la fois d'une réduction de l'espace ($R \leq M$) acceptable, et aussi en ayant la totalité de l'information, garantissant ainsi les performances optimales. Une variante de l'ACP, plus utilisée dans la communauté du traitement du signal est connue sous le nom de transformée de Karhunen-Loève [40]. L'ACP est une technique rapide, simple et populaire dans l'identification de modèle, c'est l'une des meilleures techniques. Les projections de l'ACP sont optimales pour la reconstruction d'une base de dimension réduite. Cependant, l'ACP n'est pas optimisée pour la séparabilité (discrimination) de classe. Une alternative qui est l'analyse discriminante linéaire LDA tient compte de ceci.

II.5.2.2. Le processus de reconnaissance (Karhunen-Loeve)

Au début de la phase d'apprentissage, on fait l'acquisition, la lecture et la normalisation des images d'apprentissage (de taille R). Puis on calcule le visage moyen de ces images. On soustrait ensuite le visage moyen de chaque image d'apprentissage normalisée (ces nouvelles images représentent les colonnes d'une matrice S). On fait le calcul de la matrice de covariance qui est définie par la formule $L=S^t*S$. On calcule ensuite les vecteurs propres V et les valeurs propres D de la matrice L . Puis on passe à l'étape de calcul des visages propres selon la formule $U = S * V * (\text{abs}(D))^{-0.5}$. Et finalement on calcule les poids des visages de la base (de taille M) en les projetant dans le sous-espace engendré par les visages propres "Face Space".

Pendant la phase de vérification, on fait l'acquisition, la lecture et la normalisation de l'image de vérification (de taille R). Puis on soustrait le visage moyen (calculé auparavant lors de la phase d'apprentissage) de l'image de vérification normalisée. Ensuite on calcule le poids de l'image (de taille M) en utilisant les visages propres comme une base de projection. Et finalement on utilise la distance euclidienne comme une mesure de similarité [41].

II.5.3. Analyse Discriminante Linéaire

II.5.3.1. Présentation

L'analyse discriminante linéaire (*LDA*) et les discriminants linéaires de Fisher (*FLD*) relatifs, sont utilisés pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement.

Les combinaisons résultantes peuvent être employées comme classificateur linéaire, ou généralement dans la réduction de caractéristiques avant la classification postérieure.

LDA est étroitement lié à l'ACP (*PCA en anglais*), du fait que tous les deux recherchent des combinaisons linéaires des variables qui représentent au mieux les données. *LDA* essaye explicitement de modéliser la différence entre les classes des données. *PCA* quant à elle, ne tient pas compte des différences entre les classes.

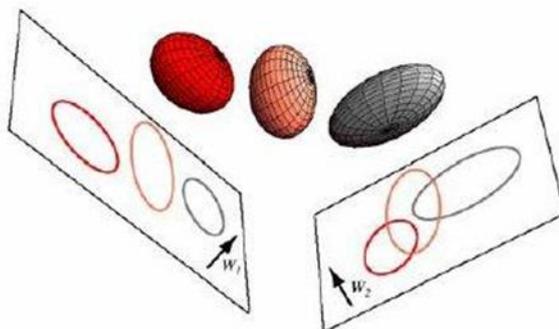


Figure 20 : Illustration du principe de séparation optimale des classes par le LDA. Trois distributions 3D sont projetées sur deux sous-espaces 2D décrits par les vecteurs W_1 et W_2 . Puisque le LDA essaye de trouver la plus grande séparation parmi les classes, on voit bien que W_1 est ici le vecteur optimal

Chaque visage, qui se compose d'un grand nombre de Pixel, est réduit à un plus petit ensemble de combinaisons linéaires avant la classification

Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixel, qui forment un *template*. Les combinaisons linéaires obtenues en utilisant *FLD* s'appellent les *Fisherfaces*, en analogie avec les *Eigenfaces* [42]. *LDA* est une technique qui cherche les directions qui sont efficaces pour la discrimination entre les données.

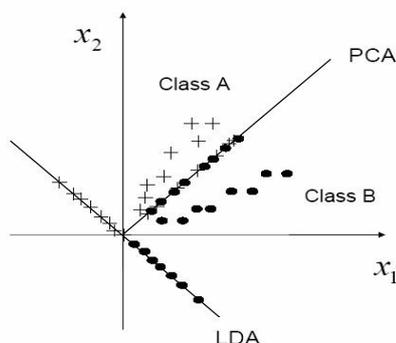


Figure 21 : Les projections PCA et LDA d'un ensemble de données

LDA est plutôt connue pour sa maximisation de l'éparpillement d'inter-classe “*the between-class scatter*” et sa réduction au minimum de l'éparpillement d'intra-classe “*the within-class scatter*”, qui se manifeste par le groupement des vecteurs de poids de la même classe (faible distance entre ces vecteurs), et par la séparation des vecteurs de poids de classes différentes (grande distance entre ces vecteurs) [43].

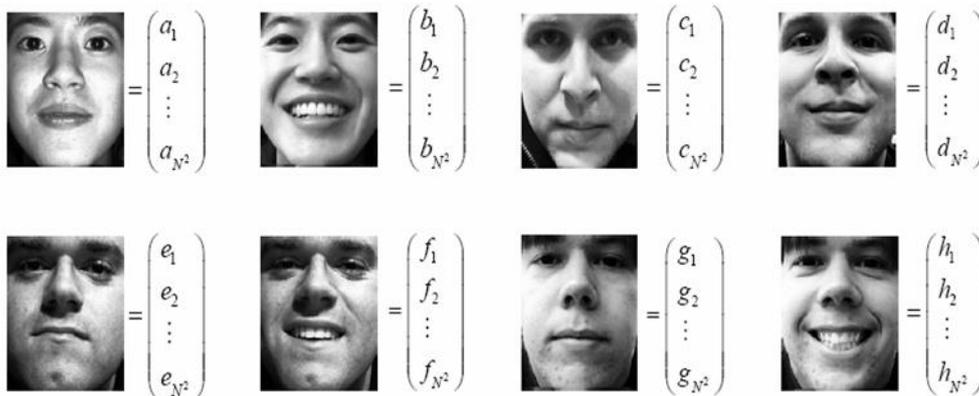
II.5.4. DF-LDA

La plupart des méthodes traditionnelles basées sur *LDA*, souffrent de l'inconvénient que leurs critères d'optimalité ne sont pas directement liés aux capacités de classification de la représentation obtenue des caractéristiques.

D'ailleurs, leur exactitude de classification est affectée par le problème de la petite dimension de l'échantillon (« SSS », small sample size) qui est souvent rencontré en reconnaissance de visages. L'algorithme de DF-LDA “DIRECT FRACTIONAL-STEP LDA” traite ces deux imperfections d'une façon efficace et rentable, offrant des performances supérieures à celles des techniques de reconnaissance de visages classiques.

L'algorithme de *DF-LDA* est représenté comme suit [44]:

- ✓ L'acquisition et la lecture des images d'apprentissage



- ✓ Normalisation des images
- ✓ Calcul du visage moyen m
- ✓ Calcul du visage moyen de chaque classe

$$\begin{aligned}
 \bar{x} &= \frac{1}{2} \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_{N^2} + b_{N^2} \end{pmatrix}, & \bar{y} &= \frac{1}{2} \begin{pmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_{N^2} + d_{N^2} \end{pmatrix}, \\
 \bar{z} &= \frac{1}{2} \begin{pmatrix} e_1 + f_1 \\ e_2 + f_2 \\ \vdots \\ e_{N^2} + f_{N^2} \end{pmatrix}, & \bar{w} &= \frac{1}{2} \begin{pmatrix} g_1 + h_1 \\ g_2 + h_2 \\ \vdots \\ g_{N^2} + h_{N^2} \end{pmatrix}
 \end{aligned} \tag{2.5}$$

- ✓ Soustraction des visages d'apprentissage

$$\begin{aligned} \vec{a}_m &= \begin{pmatrix} a_1 - x_1 \\ a_2 - x_2 \\ \vdots \\ a_{N^2} - x_{N^2} \end{pmatrix}, \quad \vec{b}_m = \begin{pmatrix} b_1 - x_1 \\ b_2 - x_2 \\ \vdots \\ b_{N^2} - x_{N^2} \end{pmatrix}, \quad \vec{c}_m = \begin{pmatrix} c_1 - y_1 \\ c_2 - y_2 \\ \vdots \\ c_{N^2} - y_{N^2} \end{pmatrix}, \quad \vec{d}_m = \begin{pmatrix} d_1 - y_1 \\ d_2 - y_2 \\ \vdots \\ d_{N^2} - y_{N^2} \end{pmatrix}, \\ \vec{e}_m &= \begin{pmatrix} e_1 - z_1 \\ e_2 - z_2 \\ \vdots \\ e_{N^2} - z_{N^2} \end{pmatrix}, \quad \vec{f}_m = \begin{pmatrix} f_1 - z_1 \\ f_2 - z_2 \\ \vdots \\ f_{N^2} - z_{N^2} \end{pmatrix}, \quad \vec{g}_m = \begin{pmatrix} g_1 - w_1 \\ g_2 - w_2 \\ \vdots \\ g_{N^2} - w_{N^2} \end{pmatrix}, \quad \vec{h}_m = \begin{pmatrix} h_1 - w_1 \\ h_2 - w_2 \\ \vdots \\ h_{N^2} - w_{N^2} \end{pmatrix} \end{aligned} \quad (2.6)$$

- ✓ La construction des scatter matrices (une matrice par classe) :

$$\begin{aligned} S_1 &= (\vec{a}_m \vec{a}_m^T + \vec{b}_m \vec{b}_m^T), \quad S_2 = (\vec{c}_m \vec{c}_m^T + \vec{d}_m \vec{d}_m^T), \\ S_3 &= (\vec{e}_m \vec{e}_m^T + \vec{f}_m \vec{f}_m^T), \quad S_4 = (\vec{g}_m \vec{g}_m^T + \vec{h}_m \vec{h}_m^T) \end{aligned} \quad (2.7)$$

- ✓ Le calcul de la *within-class scatter matrix* SW:

$$S_W = S_1 + S_2 + S_3 + S_4 \quad (2.8)$$

- ✓ Le calcul de la *weighted between-class scatter matrix*

$$\hat{S}_{BTW} = \sum_{i=1}^C \Phi_i \Phi_i^T \quad (2.9)$$

avec,

$$\Phi_i = \left(\frac{L_i}{L} \right)^{\frac{1}{2}} \sum_{j=1}^C (\omega(d_{ij}))^{\frac{1}{2}} (\bar{Z}_i - \bar{Z}_j) \quad (2.10)$$

est la moyenne de la classe Z_i ,

$L_i = \dim(Z_i)$, L est le nombre d'image d'apprentissage, C est le nombre de classe, d_{ij} est la distance euclidienne entre les moyennes des classes i et j , La fonction pesante $\omega(d_{ij})$ est une fonction monotoniquement décroissante de la distance d_{ij}

(il est recommandé d'utiliser : $\omega(d_{ij}) = (d_{ij})^{-2p}$ avec $p=2,3,\dots$)

$$\text{Le calcul de la matrice } S_{TOT} = S_{WTH} + \hat{S}_{BTW} \quad (2.11)$$

la métrique utilisée dans le *DF-LDA* et :

$$\Psi = \arg \max_{\Psi} \frac{|(\Psi^T \hat{S}_{BTW} \Psi)|}{|(\Psi^T S_{TOT} \Psi)|} \quad (2.12)$$

II.5.5. Analyse en Composantes Indépendantes

PCA est une technique optimale de recherche de représentation réduite qui minimise l'erreur de reconstruction, cependant les vecteurs de base tenant compte de l'erreur de reconstruction peuvent ne pas être optimaux pour coder l'information appropriée de l'image à la classification. L'*analyse en composantes indépendantes (ICA)* est une généralisation de *PCA* qui utilise en plus des statistiques d'ordre 2, les statistiques d'ordre plus supérieur, ce qui peut produire une représentation de données plus puissante.

Le but d'*ICA* est de trouver des vecteurs de base (images de base) localisés dans l'espace et qui sont statistiquement indépendants, en réduisant au minimum la dépendance statistique.

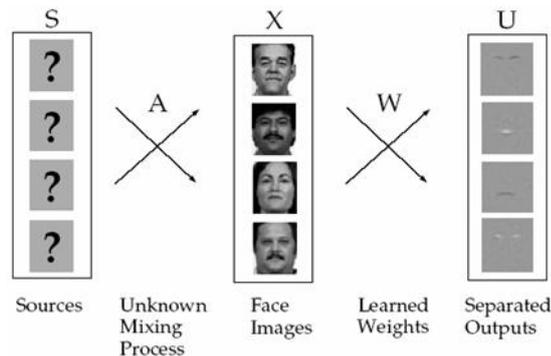


Figure22 : Le modèle de synthèse d'image de l'architecture 1 d'ICA

Avant l'apprentissage, on soustrait l'image moyenne des images d'apprentissage X et on les filtre par un filtre de blanchiment (whitening filter)

$$W_z = 2(XX^T)^{-\frac{1}{2}} \tag{2.13}$$

Pour supprimer les statistiques d'ordre 1 et 2 (la moyenne et les covariances sont mises à zéro et les variances sont égalisées). Pour trouver un ensemble d'images composantes indépendantes, les images d'apprentissage X sont considérées comme étant une combinaison linéaire d'images de base statistiquement indépendantes S, où A est une matrice de mélange inconnue.

Les images de base sont retrouvées par une matrice de filtre d'apprentissage $W_i = W \times W_z$, qui produit des sorties U statistiquement indépendantes. Ces images de base sont considérées comme un ensemble de caractéristiques faciales statistiquement indépendantes, où les valeurs des pixels de chaque image caractéristique sont statistiquement indépendantes des valeurs des pixels des autres images caractéristiques [45].

The diagram shows a grayscale face image on the left, followed by an equals sign and a series of terms: $b_1 * u_1 + b_2 * u_2 + \dots + b_n * u_n$. Each u_i is represented by a small grayscale image showing a specific facial feature (like eyes, nose, or mouth) on a dark background.

Figure 23 : La représentation d'ICA = (b1, b2,...,bn)

La représentation d'ICA est constituée des coefficients b de la combinaison linéaire des images de base U qui forment chaque image de visage. Le nombre de composantes indépendantes extraites par l'algorithme d'ICA est égal à la dimension de l'entrée [45] (le nombre d'images d'apprentissage), et si on veut contrôler ce nombre, au lieu d'exécuter ICA sur les n images originales, on l'exécute sur un ensemble de m combinaisons linéaires de ces images, où $m < n$.

Communément, on choisit pour ces combinaisons linéaires un sous ensemble formé des premières composantes principales de la matrice de donnée, ainsi l'exécution d'ICA sur l'ensemble de ces m composantes principales produit m images sources U_i statistiquement indépendantes.

Dans cette première architecture d'ICA, les images de base U_i sont statistiquement indépendantes, mais les coefficients b_i ne le sont pas. Au lieu de séparer les images de visages en ensembles d'images indépendantes, il existe une deuxième architecture d'ICA, dans laquelle les coefficients sont statistiquement indépendants, en séparant les pixels en un ensemble de variables indépendantes codant les images.

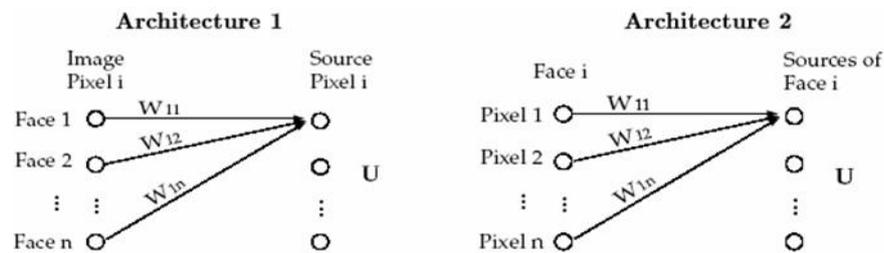


Figure 24 : Les deux architectures d'ICA : l'architecture 1 pour trouver des images de base statistiquement indépendantes (la séparation des images de visages produit des images composantes indépendantes), et l'architecture 2 pour trouver un code factoriel (la séparation des pixels produit un code factoriel)

$$\begin{matrix}
 & & \mathbf{a}_1 & & \mathbf{a}_2 & & \dots & & \mathbf{a}_n \\
 \text{Image} & = & u_1 * & + & u_2 * & + & \dots & + & u_n * \\
 \text{Face} & & \text{Image} & & \text{Image} & & & & \text{Image} \\
 \text{1} & & & & & & & & \\
 \end{matrix}$$

Figure 25 : Le modèle de synthèse d'image de l'architecture 2 d'ICA

Dans l'architecture 2 d'ICA, chaque image de l'ensemble de données est considérée comme étant créée à partir d'un ensemble d'images de base en A , et d'un vecteur de causes d'images "image causes" fondamentales statistiquement indépendantes en S . L'algorithme d'ICA essaye d'inverser les images de base en trouvant un ensemble de filtres W_i qui produisent des sorties statistiquement indépendantes.

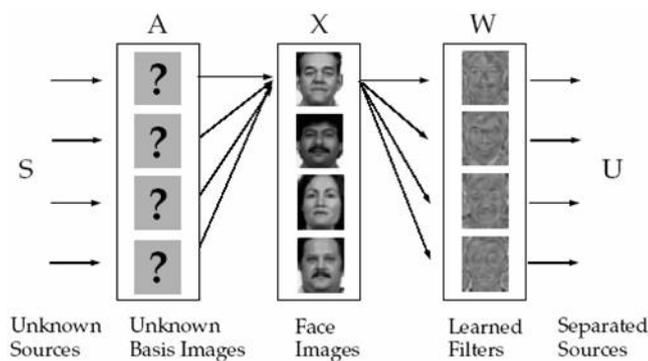


Figure 26 : La représentation factorielle d'ICA = (u_1, u_2, \dots, u_n)

II.5.6. Le réseau de neurones

II.5.6.1. Introduction

L'application des réseaux de neurones artificiels dans la reconnaissance de visages a visé plusieurs problèmes : la classification Homme Femme, la reconnaissance de visages et la classification des expressions faciales [46]. Un réseau de neurones est un système de traitement de l'information qui a été développé comme généralisations des modèles mathématiques assortissant la connaissance humaine. Ils se composent d'un grand nombre d'unités de traitement hautement reliées appelées neurones, travaillant ensemble pour exécuter une tâche de classification donnée.

Il est un processeur parallèle distribué, ayant une prospérité naturelle pour stocker une connaissance expérimentale. Il ressemble au cerveau humain en trois aspects : la connaissance est acquise par le réseau par un processus d'apprentissage, des forces de connexion reliées ensemble, connues sous le nom de poids synaptiques, sont employées pour stocker la connaissance, et chaque neurone a un état interne appelé seuil ou fonction d'activation (ou fonction de transfert) utilisée pour classifier les vecteurs.

Malheureusement, ils sont souvent difficiles à construire puisque leur structure (nombre de couches cachées et nombre de neurones par couche pour les perceptrons) influe beaucoup sur les résultats et il n'existe pas de méthode pour déterminer automatiquement cette structure. Néanmoins, nous donnons des exemples d'études ayant utilisées les réseaux de neurones pour la reconnaissance de visages.

Cottrell et al. [51] ont cherché des méthodes d'analyse automatique du visage les plus proches possible de la réalité biologique. Ainsi, un réseau de neurones dit auto-supervisé, c'est-à-dire dont la couche d'entrée et la couche de sortie sont identiques et égales à l'image d'un visage, effectue une analyse en composantes principales.

Le nombre de composantes principales est donné par le nombre de neurones de la couche cachée. Chaque neurone de la couche cachée correspond aux valeurs propres de la décomposition. Lawrence et al. [52] choisissent de classer les vecteurs de caractéristiques, extraites par cartes de Kohonen, à l'aide d'un réseau de neurones convolutionnel. Ce type de réseau de neurones est partiellement invariant à des transformations globales telles que la translation, la rotation et les changements d'échelle.

Les résultats expérimentaux ont montré la supériorité des réseaux de neurones convolutionnels sur les réseaux auto-supervisés de Cottrell [51].

Lin et al. [53] ont utilisé un réseau de neurones à décision probabiliste alliant ainsi les avantages des approches statistiques et des réseaux de neurones. Il a été montré que les performances de cette solution sont comparables à la méthode de Lawrence et al. , tout en étant beaucoup moins coûteux en termes de temps de calcul.

On trouve plusieurs types de réseaux de neurones : l'apprentissage compétitif (non-supervisé, parfois appelé *VQ*), le *LVQ* (apprentissage compétitif supervisé), les cartes auto organisatrices, le perceptron multicouches avec apprentissage par rétro propagation, les réseaux probabilistes, les réseaux à fonctions radiales de base (*FRB*), et les réseaux récurrents [47].

Une classification par réseau de neurones comporte les étapes suivantes :

Tout d'abord une phase de prétraitement des images d'apprentissage et l'association à chaque image d'apprentissage (entrée du réseau) un vecteur de sortie, puis vient l'étape d'initialisation (création des couches du réseau). On fait l'apprentissage (supervisé) du réseau, jusqu'à atteindre une certaine erreur minimale (le réseau apprend à bien classifier les images d'apprentissage). On présente ensuite au réseau une nouvelle image à identifier (phase de reconnaissance ou de simulation ou d'activation du réseau) qui sera finalement affectée à une classe donnée [48].

II.5.6.2. Le réseau de neurones LVQ

Parmi tous les types de réseau de neurones qui existent, le type *LVQ* est le type le plus adapté à l'application de reconnaissance de visage, puisqu'au contraire des autres types qui donnent des performances basses, le *LVQ* quand à lui réalise de hautes performances de classification.

L'algorithme de base *LVQ* :

- ✓ application d'une entrée
- ✓ recherche du neurone vainqueur (sa sortie est égale à 1)
- ✓ modification des poids (rapprochement ou éloignement selon que le neurone vainqueur est le neurone désiré ou non)

Dans le langage de programmation *Matlab* le réseau de type *LVQ* est programmé au moyen de 2 couches : une couche de compétition pour la détermination du neurone vainqueur (la sortie de ce dernier est mise à 1, celle des autres neurones à 0), et une couche de neurones linéaires pour la classification. La 2^{ème} couche comporte un neurone par classe, la 1^{ère} comporte un neurone par sous-classe.

L'apprentissage modifie les poids de la première couche. Les neurones de la 2ème couche combinent certaines sorties des neurones de la 1ère au moyen de fonctions OU [47].

II.5.7. Machine à vecteurs de support

Une machine à vecteurs de support est une technique de discrimination, c'est une méthode d'apprentissage supervisé utilisée pour la classification et la régression. Elle consiste à séparer deux ou plusieurs ensembles de points par un hyperplan. Selon les cas et la configuration des points, la performance de la machine à vecteurs de support peut être supérieure à celle d'un réseau de neurones ou d'un modèle de mixture gaussienne [49].

II.5.7.1. Le principe de SVM

L'idée originale des SVMs est basée sur l'utilisation de fonctions noyau kernel, qui permettent une séparation optimale des points du plan en différentes catégories. La méthode fait appel à un ensemble de données d'apprentissage, qui permet d'établir un hyperplan séparant au mieux les points [49].

II.5.7.2. La séparation optimale

Dans un plan à deux dimensions dans lequel sont répartis deux groupes de points associés à deux classes : les points (+) pour $y > 0$ et les points (-) pour $y < 0$, on peut trouver un séparateur linéaire évident, qui est l'axe des abscisses et qui sépare parfaitement les points. Le problème est dit *linéairement séparable*. Pour des problèmes plus compliqués, la caractérisation d'un séparateur linéaire peut être très compliquée et tout à fait non optimale. Par exemple dans un plan dans lequel les points (+) sont regroupés en un cercle, avec des points (-) tout autour, aucun séparateur linéaire en deux dimensions ne pourra correctement séparer les groupes. Le problème dans ce cas n'est pas *linéairement séparable*. La SVM utilise les *trucs du noyau* "kernel trick" pour appliquer une classification linéaire à un problème de classification non linéaire [49].

II.5.7.3. Le concept d'hyperplan

Afin de remédier au problème de l'absence de séparateur linéaire, le concept des SVMs est de reconsidérer le problème dans un espace de dimension supérieure. Dans ce nouvel espace, il existe un séparateur linéaire qui permet de classer au mieux les points dans les deux groupes qui conviennent, avec une distance maximale entre l'hyperplan et les points les plus proches des deux classes, afin de pouvoir classifier plus exactement les nouveaux points, cette distance s'appelle la *marge*. Cet hyperplan est connu sous le nom de *l'hyperplan à marge maximum* ou *l'hyperplan optimal* et les vecteurs (points de dimension n) les plus proches de cet hyperplan s'appellent les *vecteurs de support*. On pourra ensuite projeter le séparateur linéaire dans l'espace d'origine pour visualiser le résultat de la classification. Le séparateur linéaire obtenu peut être un hyperplan, c'est à dire la généralisation à n dimensions d'une ligne (1D) séparant un espace 2D, ou d'un plan (2D) séparant un espace 3D. Le changement d'espace se fait au moyen d'une fonction répondant au critère de Mercer. Ce critère permet un changement dans les deux sens, ce qui permet à partir de l'expression de l'hyperplan dans l'espace complexe de classer les éléments dans l'espace de description initial [49].

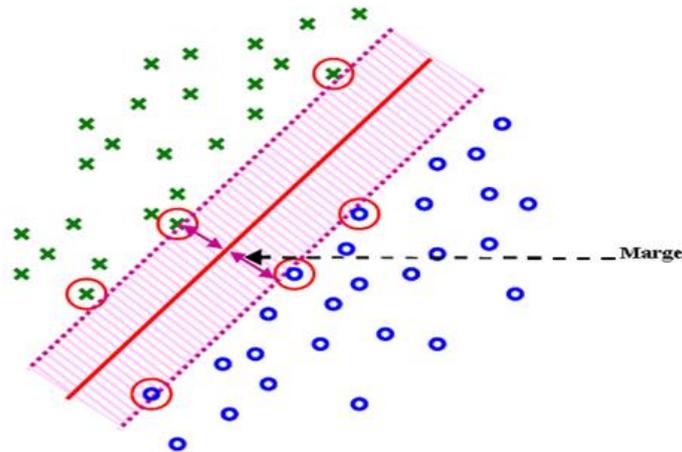


Figure 27 : Exemple de deux classes linéairement séparables. L'hyperplan déterminé par la SVM, maximisant la marge, permet de séparer les deux classes de manière optimale.

II.5.7.4. La classification linéaire :

On considère un ensemble de données $\{(x_1, c_1), (x_2, c_2), \dots, (x_p, c_p)\}$ où c_p est la classe à laquelle appartient le point x_p , c_p est égale à (-1) ou à $(+1)$ et x_p est un vecteur de dimension n de valeurs appartenant à l'intervalle $[0,1]$ ou $[-1,1]$. Ces *données d'apprentissage* désignent la classification correcte qu'on voudra que la SVM fasse par la suite, au moyen de l'hyperplan de division, qui obéit à la forme :

$$w \cdot x - b = 0 \quad (2.14)$$

On s'intéresse aussi aux hyperplans parallèles à l'hyperplan optimal qui sont tangents aux vecteurs de support des classes, et qui sont décrits par les équations : $w \cdot x - b = 1$ et $w \cdot x - b = -1$. On veut que ces hyperplans maximisent la distance avec l'hyperplan de division et qu'il n'y ait aucun point de données entre les deux. En employant la géométrie, on trouve que la distance entre les hyperplans est $2 / |w|$, ainsi le but est de réduire au minimum $|w|$.

Pour exclure les points de données, on doit s'assurer que pour chaque i on a $w \cdot x_i - b \geq 1$ ou $w \cdot x_i - b \leq -1$, ou sous une autre équation $c_i(w \cdot x_i - b) \geq 1$ avec $1 \leq i \leq p$. Le problème maintenant est de minimiser $|w|$ sous cette dernière contrainte, ceci est un problème d'optimisation de *programme quadratique QP*). Il existe plusieurs algorithmes spécialisés pour résoudre rapidement les problèmes *QP* et la méthode la plus commune est l'algorithme *SMO*.

Après l'apprentissage, la SVM peut classifier les nouvelles données (données de test) en utilisant la règle de décision suivante :

$$\hat{c} = \begin{cases} 1, & \text{if } w \cdot x + b \geq 0 \\ -1, & \text{if } w \cdot x + b \leq 0 \end{cases} \quad (2.15)$$

On peut constater d'après la règle de classification qu'elle est seulement en fonction des vecteurs de supports (les données d'apprentissage se trouvant sur la marge) [49].

II.5.7.5. La classification non linéaire

L'algorithme original de l'hyperplan optimal est un classificateur linéaire, et pour avoir un classificateur non linéaire on applique le *kernel trick* à l'*hyperplan à marge maximum*, l'algorithme résultant est formellement semblable, sauf que chaque produit scalaire est remplacé par une fonction noyau non linéaire.

Parmi les noyaux les plus communs, on trouve :

✓ le noyau polynomial (homogène) $k(x, \hat{x}) = (x \cdot \hat{x})^d$ (2.16)

✓ le noyau polynomial (inhomogène) : $k(x, \hat{x}) = (x \cdot \hat{x} + \mathbf{1})^d$ (2.17)

✓ la fonction radiale de base RBF : $k(x, \hat{x}) = e^{(-\gamma \|x - \hat{x}\|^2)}$ avec $\gamma > 0$, (2.18)

✓ la fonction radiale de base de type gaussien : $k(x, \hat{x}) = e^{-\left(\frac{\|x - \hat{x}\|^2}{2\sigma^2}\right)}$ (2.19)

✓ la fonction sigmoïde: $k(x, \hat{x}) = \tanh(kx \cdot \hat{x} + c)$ (2.20)

II.5.8. Modèle de Markov caché

Les modèles de Markov cachés *HMMs* sont massivement utilisés en reconnaissance de formes, en intelligence artificielle ou encore en traitement automatique du langage naturel. Les *HMMs* sont généralement employés pour la modélisation statistique de séries de vecteur de temps non stationnaire. En considérant l'information du visage comme étant une séquence variable dans le temps, le *HMM* peut être appliqué à la reconnaissance de visages [46]. Les caractéristiques faciales les plus significatives d'une image de visage frontale, à savoir les cheveux, le front, les yeux, le nez et la bouche, se présentent dans un ordre naturel de haut en bas, même dans le cas où il y'a des petites rotations. En se basant sur cette observation, l'image d'un visage peut être modélisée en utilisant un *HMM* unidimensionnel en assignant à chacune de ces régions un état. Le *HMM* classe une caractéristique par la propriété de la *chaîne de Markov*. Une séquence des valeurs des pixels forme une *chaîne de Markov*, si la probabilité que le système à l'instant $n+1$ soit à l'état x_{n+1} dépend uniquement de la probabilité que le système à l'instant n soit à l'état x_n .

Dans une *chaîne de Markov*, la transition d'un état à un autre est probabiliste ce qui donne une distribution de probabilité de l'ensemble des sorties de chaque état. Ce résultat est employé pour faire la comparaison entre deux visages [46] [48].

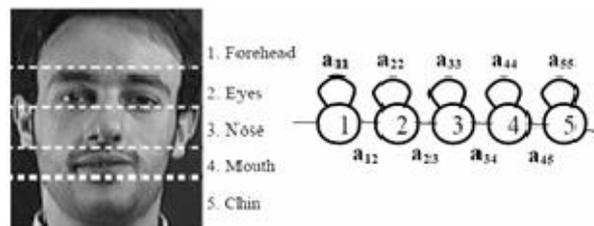


Figure 28 : Les 5 états du HMM (de haut en bas)

Pour une image de visage d'une personne donnée de l'ensemble d'apprentissage, le but de l'étape d'apprentissage est d'optimiser les paramètres pour mieux décrire l'observation. La reconnaissance est effectuée en assortissant l'image de test contre chaque modèle d'apprentissage (chaque *HMM* représente une personne différente). Et dans la fin de cette procédure, l'image est convertie en une séquence d'observation et la probabilité est calculée pour chaque modèle enregistré.

Le modèle avec la probabilité la plus élevée indique l'identité de la personne inconnue [46] [39]. L'approche *HMM* a des performances de reconnaissance satisfaisantes, cependant il nécessite des calculs intenses ce qui augmente les temps d'exécution.

II.5.9. Méthodes basées sur l'apparence locale (Local appearance-based methods)

Ces techniques sont utilisées de manière modulaire pour les différentes régions faciales. Un modèle global est alors défini à partir de la combinaison des différents modèles locaux. Ainsi, les différentes régions faciales ne seront plus affectées de la même manière par les différentes sources de variabilité. Par exemple, le port de lunettes de soleil change considérablement l'aspect des yeux, tandis qu'un sourire affecte plus la région de la bouche.

Deux paramètres sont utilisés pour définir les régions locales du visage : la forme (rectangle, ellipse) et la taille. Les caractéristiques des régions locales sont déterminées à partir d'une analyse des valeurs de niveau gris [54]. Des techniques comme les Ondelettes de Gabor [55] ou de Harr [56] et l'analyse fractale sont aussi utilisées pour l'extraction de caractéristiques. En général, les caractéristiques à base de valeurs de gris préservent l'information de texture, tandis que les caractéristiques de Gabor sont plus robustes face aux changements d'illumination et aux transformations géométriques [55].

II.5.9.1. ACP modulaire

Dans [12], Pentland et al. ont introduit l'approche des « Modular Eigenspace ». Les régions faciales utilisées sont: la totalité du visage, les yeux et le nez. Une ACP est appliquée sur chacune de ces régions et les résultats de classifications obtenus sont agrégés. La bouche étant trop sensible à des changements d'expression faciale, sa prise en compte engendre une baisse du taux de reconnaissance. Cette approche peut être qualifiée d'hybride, puisqu'elle utilise à la fois des caractéristiques globales et locales.

Les auteurs ont montré qu'elle est plus efficace que les techniques globales ou strictement locales appliquées séparément.

II.5.10. Template matching

Le template matching est une technique globale de reconnaissance de visage. La corrélation est généralement exploitée pour mesurer la similitude entre un template (masque) stocké et l'image à reconnaître. Les templates devraient être délibérément conçus pour couvrir la variété de variations possibles d'image. Pendant la recherche dans l'image, l'échelle et la rotation devraient également être soigneusement considérées pour accélérer le processus. Cette technique a aussi été exploitée pour localiser les caractéristiques saillantes "salient features" comme les yeux le nez et la bouche, dans une image de visage. D'affines transformations ont été introduites dans le processus de reconnaissance, lors de changement de pose [50].

II.5.10.1. L'algorithme du template matching :

On commence par l'acquisition, la lecture et la normalisation des images d'apprentissage. On calcule ensuite le visage moyen de chaque classe de personnes (les templates). Dans l'étape de reconnaissance, on fait l'acquisition, la lecture et la normalisation de l'image de vérification. Puis on calcule les différences entre l'image de vérification et les templates (les différences sont des images). On calcule après les sommes de ces différences (somme des pixels de l'image différence). Et finalement la somme minimale va référencer la classe auquel appartient l'image de vérification.

II.6. Discussion

Ces dernières années, de nombreuses approches pour la reconnaissance automatique de visages ont été proposées. Elles peuvent être regroupées en deux types d'approches principales. Les approches globales pour lesquelles les caractéristiques sont extraites directement de l'image du visage en entier et les approches locales basées sur l'extraction de traits locaux du visage, tel que les yeux, le nez ou la bouche.

Dans ce chapitre nous avons passé en revue les méthodes les plus connues de la détection de visage puis nous nous sommes focalisés sur les principaux algorithmes de la reconnaissance faciale, qui se divisent en trois catégories à savoir les méthodes globales, locales et hybrides.

Nous avons décrit également leurs avantages et leurs inconvénients puis dégager les problèmes liés à leur application dans un environnement réel. En effet, un système de reconnaissance automatique de visages est efficace dans le cas où les images faciales seraient bien cadrées et prises dans de bonnes conditions de luminosité. Sa performance se dégrade fortement lorsque ces images, utilisées pour l'apprentissage ou lors de la reconnaissance, présentent des variations importantes de luminosité et/ou de pose du visage.

Chapitre III

Reconnaissance par Eigenface

III.1. Préambule

La reconnaissance du visage a atteint un haut niveau de performances, même avec l'utilisation d'un seul exemple d'apprentissage. Dans les conditions actuelles des travaux sur la reconnaissance faciale, le système visuel humain reste encore le plus robuste face aux diverses variations pouvant altérer le processus d'identification : changement des conditions d'éclairage, variations de l'expression faciale, modifications de l'apparence du visage à travers la présence ou l'absence de lunettes, barbe, maquillage.

De nombreuses techniques ont été développées ces dernières années, nous avons passées en revue les plus connues dans le chapitre deux. Parmi elles, Eigenface, qui est une technique particulièrement prisée par les chercheurs de la communauté de la biométrie.

Nous commencerons d'abord par détailler le principe de la méthode Eigenface, puis nous présenterons les bases de données sur lesquelles les tests ont été effectués et nous suggérerons quelques étapes pour améliorer cette méthode, en fin nous terminerons par une discussion.

III.2. Présentation de la méthode Eigenface

L'algorithme ACP, PCA en anglais (Principal Component Analysis) est né des travaux de MA. Turk et AP. Pentland au MIT Media Lab, en 1991. Il est aussi connu sous le nom de Eigenfaces car il utilise des vecteurs propres et des valeurs propres. Cet algorithme s'appuie sur des propriétés statistiques bien connues et utilise l'algèbre linéaire. Il est relativement rapide à mettre en œuvre mais il est sensible aux problèmes d'éclairage [12], de pose et d'expression faciale. Il est à la base de nombreux algorithmes globaux actuels.

L'idée principale consiste à exprimer les M images d'apprentissage selon une base de vecteurs orthogonaux particuliers, contenant des informations indépendantes d'un vecteur à l'autre. Ces nouvelles données sont donc exprimées d'une manière plus appropriée à la reconnaissance du visage.

Nous voulons extraire l'information caractéristique d'une image de visage, pour l'encoder aussi efficacement que possible afin de la comparer à une base de données de modèles encodés de manière similaire. En termes mathématiques, cela revient à trouver les vecteurs propres de la matrice de covariance formée par les différentes images de notre base d'apprentissage [39].

Une image $I_i(m, n)$ est traitée comme un vecteur $\Gamma_i(m \times n, 1)$ dans un espace vectoriel de grande dimension ($N = m \times n$), par concaténation des colonnes.

Après avoir rassemblé nos M images dans une unique matrice, nous obtenons une matrice d'images Γ , où chaque colonne représente une image Γ_i :

$$\Gamma = \begin{pmatrix} a_{1,1} & b_{1,1} & \dots & z_{1,1} \\ \vdots & \vdots & \dots & \vdots \\ a_{n,1} & b_{n,1} & \dots & z_{n,1} \\ \vdots & \vdots & \dots & \vdots \\ a_{1,m} & b_{1,m} & \dots & z_{1,m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n,m} & b_{n,m} & \dots & z_{n,m} \end{pmatrix}$$

On calcule ensuite l'image moyenne Ψ de toutes les images collectées. Cette image peut être vue comme le centre de gravité du jeu d'images (Figure 29) :

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (3.1)$$

On ajuste ensuite les données par rapport à la moyenne.



Figure 29 : Image moyenne

L'image moyenne est alors soustraite de chaque image avec la formule suivante:

$$\Phi_i = \Gamma_i - \Psi, \quad i = 1 \dots M \quad (3.2)$$

On calcule ensuite la matrice de covariance du jeu de données. Cette matrice peut être vue comme une matrice de moments d'ordre 2 :

$$C = \sum_{i=1}^M \Phi_i \Phi_i^T = AA^T, \quad A = [\Phi_1 \Phi_2 \dots \Phi_M] \quad (3.3)$$

La prochaine étape consiste à calculer les vecteurs propres et les valeurs propres de cette matrice de covariance C de taille $(N \times N)$, c'est-à-dire de l'ordre de la résolution d'une image. Le problème est que cela peut parfois être très difficile et très long.

En effet, si $N > M$ (si la résolution est supérieure au nombre d'images), il y aura seulement $M - 1$ vecteurs propres qui contiendront de l'information [40] (les vecteurs propres restants auront des valeurs propres associées nulles).

Par exemple, pour 50 images de résolution 180×200 , nous pourrions résoudre une matrice L de 50×50 au lieu d'une matrice de 36000×36000 pour ensuite prendre les combinaisons linéaires appropriées des images Φ_i . Le gain de temps de calcul serait considérable, nous passerions d'une complexité de l'ordre du nombre de pixels dans une image à celle de l'ordre du nombre d'images.

Les étapes du processus qui nous permettent d'accélérer les calculs sont décrites ci-dessous :

Considérons les vecteurs propres e_i de $C = AA^T$, associés aux valeurs propres λ_i .

On a :

$$C e_i = \lambda_i e_i \quad (3.4)$$

Les vecteurs propres v_i de $L = A^T A$, associés aux valeurs propres μ_i sont tels que :

$$L v_i = \mu_i v_i$$

Soit :

$$A^T A v_i = \mu_i v_i$$

(3.5)

En multipliant à gauche par A des deux côtés de l'égalité, nous obtenons :

$$A A^T A v_i = A \mu_i v_i$$

Puisque $C = A A^T$, nous pouvons simplifier :

$$C(A v_i) = \mu_i (A v_i) \quad (3.6)$$

De (3.4) et (3.5), nous voyons que $A v_i$ et μ_i sont respectivement les vecteurs propres et les valeurs propres de C :

$$\begin{cases} e_i = A v_i \\ \lambda_i = \mu_i \end{cases} \quad (3.7)$$

Nous pouvons donc trouver les valeurs propres de cette énorme matrice C en trouvant les valeurs propres d'une matrice L beaucoup plus petite. Pour trouver les vecteurs propres de C , il suffit juste de multiplier les vecteurs propres de L par la matrice A .

Les vecteurs propres trouvés sont ensuite ordonnés selon leurs valeurs propres correspondantes, de manière décroissante. Plus une valeur propre est grande, plus la variance capturée par le vecteur propre est importante. Cela implique que la majeure partie des informations est contenue dans les premiers vecteurs propres.

III.2.1. Choix du nombre de vecteurs propres

L'étape suivante consiste à ne sélectionner que les « k » meilleurs vecteurs propres (ceux avec les « k » plus grandes valeurs propres).

Le nombre de vecteurs propres associés aux plus grandes valeurs propres à retenir est un grand défaut de cette technique. Pour le choisir, les chercheurs ont adopté différentes solutions:

- ✓ Pour un ensemble de 115 images, Sirovitch et Kirby ont trouvé que 40 eigenpictures sont suffisantes pour représenter efficacement cet ensemble,
- ✓ Turk et Pentland [12] l'ont choisi heuristiquement. Pour leurs tests, sur une base de 16 individus, 7 vecteurs propres ont été retenus,
- ✓ Moghaddam [15] a préservé, pour comparer différentes approches de reconnaissance de visages 20 vecteurs propres en justifiant son choix par une erreur de reconstruction raisonnable (0.0012) et un taux de reconnaissance 80% obtenu par eigenfaces sur une base de 1829 images,
- ✓ Zhao et al ont retenu 300 vecteurs propres pour une base de 1038 images après avoir observé que pour un nombre très élevé, les eigenfaces ne représentent pas des visages, donc leur choix était basé sur l'allure des eigenfaces au lieu des valeurs propres.

Pratiquement, le choix de « k » peut se faire soit en fixant un certain pourcentage D tel que:

$$\frac{\sum_{i=k+1}^n \lambda_i}{\sum_{i=1}^n \lambda_i} < D$$

(3.8)

avec n le nombre total des valeurs propres ou on se basant sur le fait que la $i^{\text{ème}}$ valeur propre est égale à la variance le long de la $i^{\text{ème}}$ composante principale et on cherche, donc « k » le long du spectre des valeurs propres juste avant qu'il s'annule [12].

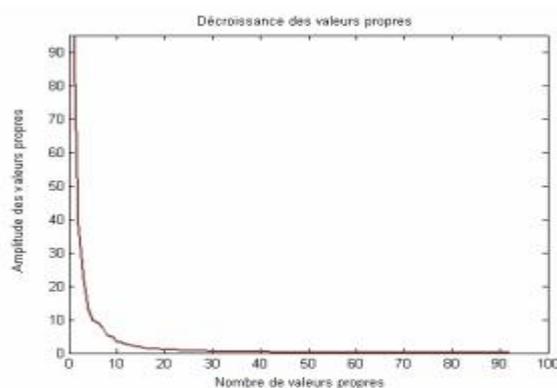


Figure 30 : Décroissance des valeurs propres de la base de données utilisée en fonction du nombre de vecteurs propres retenu.

Donc, les « k » premiers vecteurs propres correspondant aux « k » plus grandes valeurs propres sont un paramètre critique sur lequel dépend la performance du système de reconnaissance de visages (temps de calcul et taux de reconnaissance) et aussi un critère d'évaluation [12] des méthodes basées sur l'ACP.

Son choix dépend des contraintes d'application liées au temps de calcul et au nombre des images de visages disponibles et aussi de la qualité des images de visages reconstruites.

A partir de là, on définit un espace vectoriel engendré par ces k vecteurs propres, que l'on appelle l'espace des visages E_v (Face Space). Les images originales peuvent être reconstituées par combinaison linéaire de ces vecteurs propres. Les représentations graphiques de ces vecteurs rappellent un peu des images fantômes, chacune mettant en avant une partie du visage, on les appelle Eigenfaces (**figure 31**).



Figure 31 : Images représentant les Eigenface

III.2.2. Utilisation des Eigenfaces pour la classification des images visage

Nous allons maintenant projeter nos images de départ sur E_v . Une image Γ_i est alors transformée en ses composantes Eigenfaces par une simple opération de projection vectorielle :

$$\omega_k = e_k^T (\Gamma_i - \Psi) \quad , \quad k = 1, \dots, M' \quad (3.9)$$

Les vecteurs ω_k sont appelés poids et forment une matrice [5] $\Omega_T = [\omega_1, \omega_2, \dots, \omega_{M'}]$ qui décrit la contribution de chaque eigenface dans la représentation de l'image d'entrée. La matrice Ω_T est alors utilisée pour trouver quelle est, parmi un nombre prédéfini de classes, celle qui décrit le mieux une image d'entrée.

La méthode la plus simple pour déterminer quelle classe de visage fournit la meilleure description d'une image d'entrée est de trouver la classe de visage k qui minimise la distance Euclidienne.

$$\varepsilon_k^2 = \|\Omega - \Omega_k\|^2 \quad (3.10)$$

Où Ω_k est un vecteur qui décrit la $k^{\text{ème}}$ classe de visage.

Un visage appartient à une classe k quand le minimum ε_k est en dessous d'un certain seuil θ_ε . Dans le cas contraire, le visage est classé comme étant inconnu et peut éventuellement être utilisé pour créer une nouvelle classe de visage.

Il y a alors, quatre possibilités (figure 32) pour une image d'entrée d'être reconnue ou non :

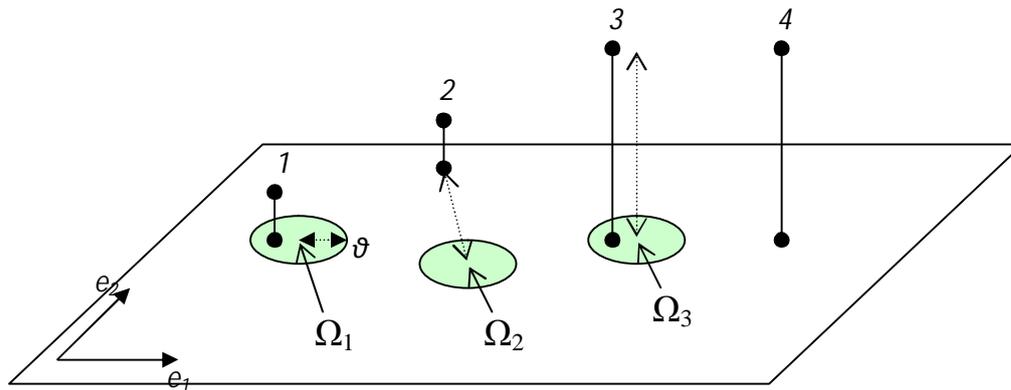


Figure 32 : Illustration des cas possibles de classification d'une image.

- ✓ Cas 1 : un individu est reconnu et identifié.
- ✓ Cas 2 : un individu inconnu du système est présent.
- ✓ Cas 3 et 4 : indiquent que l'image n'est pas une image de visage.
 - Cas 3, l'image est éloignée de E_v mais la projection est proche d'une classe connue, risque de fausse acceptation.

III.2.3. Résumé de la méthode

L'apprentissage des visages propres s'effectue selon les étapes suivantes :

- 1- Collecte des M images faciales et construction de la matrice T de taille M , par concaténation des colonnes des images faciales. Prétraitement des images collectées.
- 2- Calcul du visage moyen en sommant les colonnes de la matrice T et en divisant le vecteur résultant par le nombre d'image d'entrée (M).
- 3- Soustraction du visage moyen de la matrice T pour obtenir la matrice A ; où chaque élément représente la variance des valeurs d'intensité de chaque pixel.
- 4- Calcul de la matrice \hat{C} , $\hat{C} = A^T A$.
- 5- Calcul des vecteurs propres de C et les triés dans un ordre descendant selon les valeurs propres associées.
- 6- Calcul des vecteurs propres de la matrice de covariance C et obtention des visages propres en multipliant les vecteurs propres de C par la matrice A .
- 7- Choix des K meilleur valeurs propres et les vecteurs propres associés.
- 8- Détermination du poids des images d'entrée en projetant chaque image dans l'espace visage.
- 9- Chaque visage est maintenant représenté par un vecteur qui est utilisé pour reconstruire les images. Et enfin sauvegarde des calculs du visage moyen, des eigenfaces et du poids des images.

Les neuf étapes décrites transformeront une base de données d'images faciales en un ensemble de projections dans l'espace visage (face space).

L'étape de reconnaissance peut être résumée comme suit :

- ✓ Prétraitement de l'image d'entrée et soustraction du visage moyen.
- ✓ Détermination du poids de l'image d'entrée par la projection de celle-ci dans l'espace visage en multipliant le vecteur résultant de l'étape (1) par les eigenfaces de la base de données.
- ✓ Comparaison des résultats obtenus en utilisant des métriques telles que la distance Euclidienne.

III.3. Mesures de Distance

Lorsqu'on souhaite comparer deux vecteurs de caractéristiques issus du module d'extraction de caractéristiques d'un système biométrique, on peut soit effectuer une mesure de similarité (ressemblance), soit une mesure de distance (divergence).

La première catégorie de distances est constituée de distances Euclidiennes et sont définies à partir de la distance de Minkowski d'ordre p dans un espace euclidien \mathbb{R}^N (N déterminant la dimension de l'espace euclidien).

Considérons deux vecteurs $X = (x_1, x_2, \dots, x_n)$ et $Y = (y_1, y_2, \dots, y_n)$, la distance de Minkowski d'ordre p notée L_p est définie par :

$$L_p = \left(\sum_{i=1}^N |x_i - y_i|^p \right)^{1/p} \quad (3.11)$$

Nous allons présenter quelques mesures de distance dans l'espace original des images puis dans l'espace de Mahalanobis.

III.3.1. Distances Euclidiennes

III.3.1.1 Distance City Block (L1)

Pour $p = 1$, on obtient la distance City-Block (ou distance de Manhattan) :

$$L_1(x, y) = \sum_{i=1}^N |x_i - y_i| \quad (3.12)$$

III.3.1.2 Distance Euclidienne (L2)

Pour $p = 2$, on obtient la distance euclidienne :

$$L_2(x, y) = \sqrt{\sum_{i=1}^N |x_i - y_i|^2} \quad (3.13)$$

Les objets peuvent alors apparaître de façons très différentes selon la mesure de distance choisie (Figure 33).

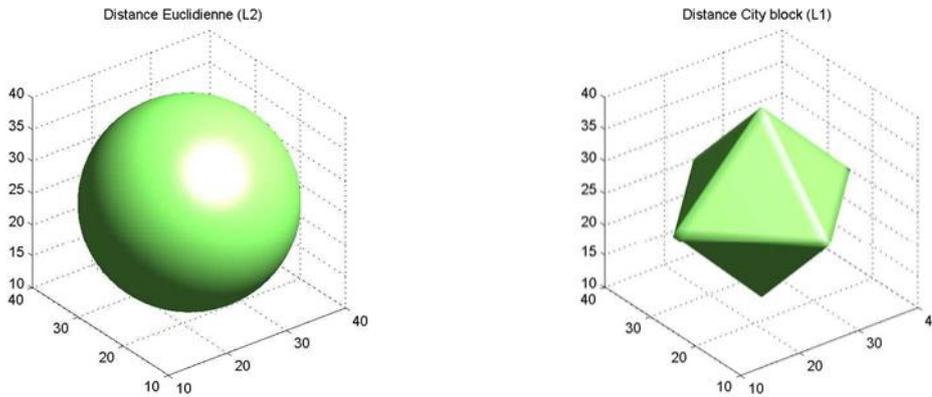


Figure 33 : Représentation d'une sphère avec la distance euclidienne (3.4.a) et la distance City-Block (3.4.b).

III.3.2. Distances dans l'Espace de Mahalanobis

III.3.2.1. De l'espace des images à l'espace de Mahalanobis

Avant de pouvoir effectuer des mesures de distance dans l'espace de Mahalanobis, il est essentiel de bien comprendre comment l'on passe de l'espace des images I_m à l'espace de Mahalanobis \mathcal{E}_{Mah} .

En sortie de l'algorithme PCA, nous obtenons des vecteurs propres associés à des valeurs propres (représentant la variance selon chaque dimension). Ces vecteurs propres définissent une rotation vers un espace dont la covariance entre les différentes dimensions est nulle. L'espace de Mahalanobis est un espace où la variance selon chaque dimension est égale à 1. On l'obtient à partir de l'espace des images I_m en divisant chaque vecteur propre par son écart-type correspondant.

Soit u et v deux vecteurs propres de I_m , issus de l'algorithme PCA, et m et n deux vecteurs de \mathcal{E}_{Mah} . Soit λ_i les valeurs propres associées aux vecteurs u et v , et σ_i l'écart-type, alors on définit $\lambda_i = \sigma_i$. Les vecteurs u et v sont reliés aux vecteurs m et n à partir des relations suivantes :

$$m_i = \frac{u_i}{\sigma_i} = \frac{u_i}{\sqrt{\lambda_i}} \quad \text{et} \quad n_i = \frac{v_i}{\sigma_i} = \frac{v_i}{\sqrt{\lambda_i}} \quad (3.14)$$

III.3.2.2. Mahalanobis L1 (MahL1)

Cette distance est exactement la même que la distance City-Block sauf que les vecteurs sont projetés dans l'espace de Mahalanobis. Ainsi, pour des vecteurs propres u et v de projections respectives m et n sur l'espace de Mahalanobis, la distance Mahalanobis L1 est définie par :

$$Mah_{L_1}(u, v) = \sum_{i=1}^N |m_i - n_i| \quad (3.15)$$

III.3.2.3. Mahalanobis L2 (MahL2)

Cette distance est identique à la distance euclidienne à part qu'elle est calculée dans l'espace de Mahalanobis. Ainsi, pour des vecteurs propres u et v de projections respectives m et n sur l'espace de Mahalanobis, la distance Mahalanobis L2 est définie par :

$$L_2(x, y) = \sqrt{\sum_{i=1}^N |x_i - y_i|^2} \quad (3.16)$$

Par défaut, lorsqu'on parle de distance de Mahalanobis, c'est à cette distance que l'on doit se référer.

III.3.2.4. Cosinus de Mahalanobis (MahCosine)

Il s'agit tout simplement du cosinus de l'angle entre les vecteurs u et v , une fois qu'ils ont été projetés sur \mathcal{E}_{Mah} et normalisés par des estimateurs de la variance (Figure 34).

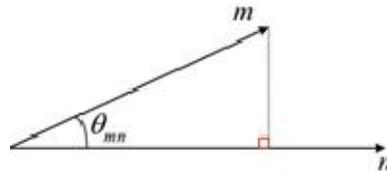


Figure 34 : Les deux vecteurs m et n dans l'espace de Mahalanobis

Nous avons donc par définition :

$$S_{MahCosine}(u, v) = \cos(\theta_{mn}) \quad (3.17)$$

De plus, on peut écrire :

$$\cos(\theta_{mn}) = \frac{|m||n|\cos(\theta_{mn})}{|m||n|} \quad (3.18)$$

D'où la formule finale de la mesure de similarité *MahCosine* :

$$S_{MahCosine}(u,v) = \frac{mn}{|m||n|}, \quad D_{MahCosine}(u,v) = -S_{MahCosine}(u,v) \quad (3.19)$$

où $D_{MahCosine}(u,v)$ est la mesure de distance équivalente. On peut enfin noter qu'il s'agit principalement de la covariance entre les vecteurs dans l'espace de Mahalanobis.

III.3.3 Choix de la distance de similarité

Dans la communauté de la reconnaissance des formes, le choix d'une distance de similarité est souvent expliqué et argumenté par rapport à l'espace d'attributs et aux points de données utilisés. Nous avons remarqué entre autre, avec les mêmes vecteurs attributs Eigenfaces, que la distance Euclidienne fournit de meilleurs résultats d'identification des individus qu'avec la distance de Manhattan. Ces performances nous ont poussés à utiliser la distance Euclidienne dans ce dernier processus.

III.4. Organigramme détaillé de l'approche Eigenface

Notre organigramme se divise en trois parties : la première est le prétraitement suivi de la phase d'apprentissage et se termine par la phase d'identification où la distance euclidienne est retenue pour le calcul de la différence entre les poids de l'image à identifier et les images de base de données puis le programme affiche la plus proche.

III.4.1. Organigramme du prétraitement

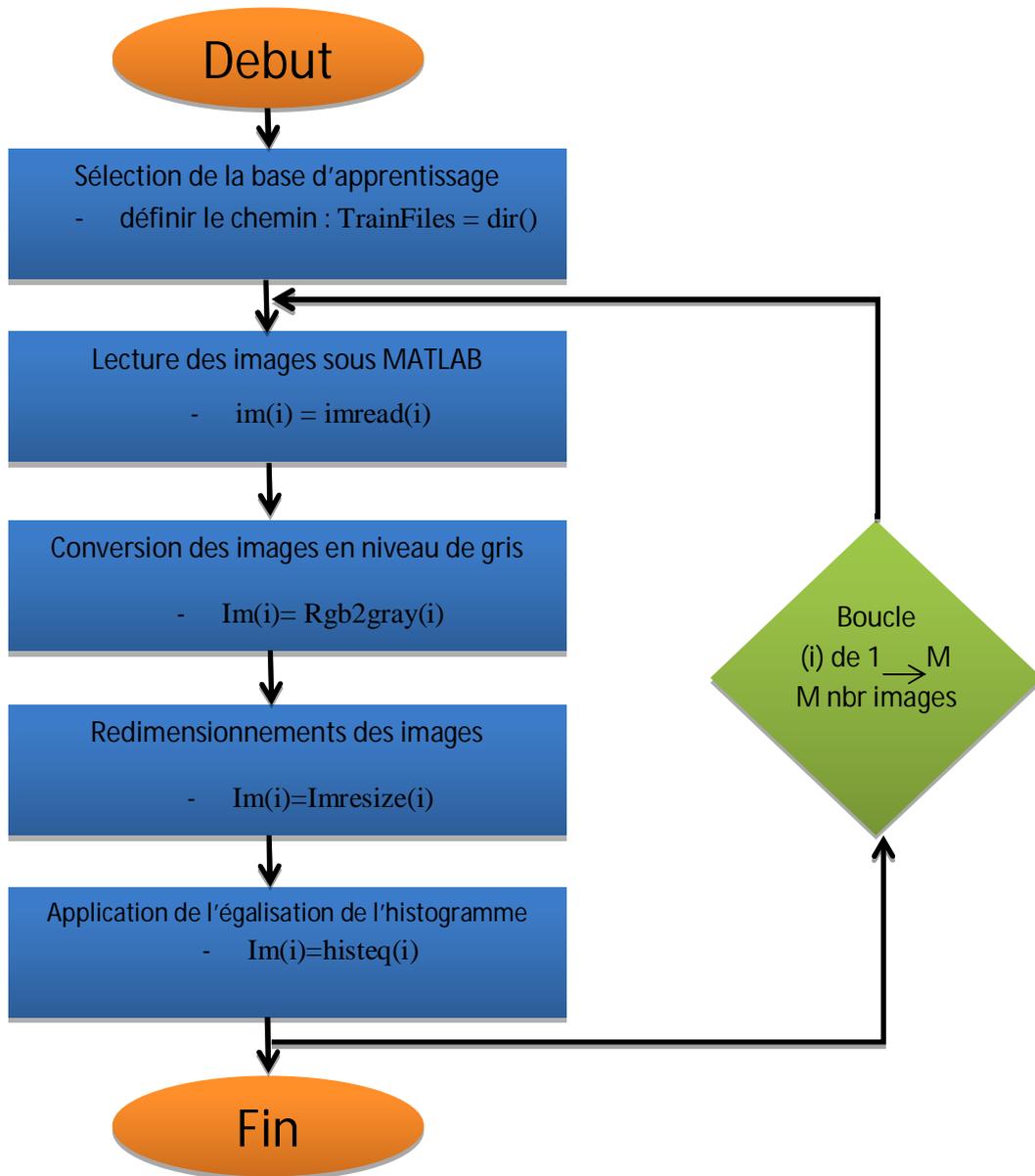


Figure 35 : prétraitements

III.4.2. Organigramme de la phase d'apprentissage

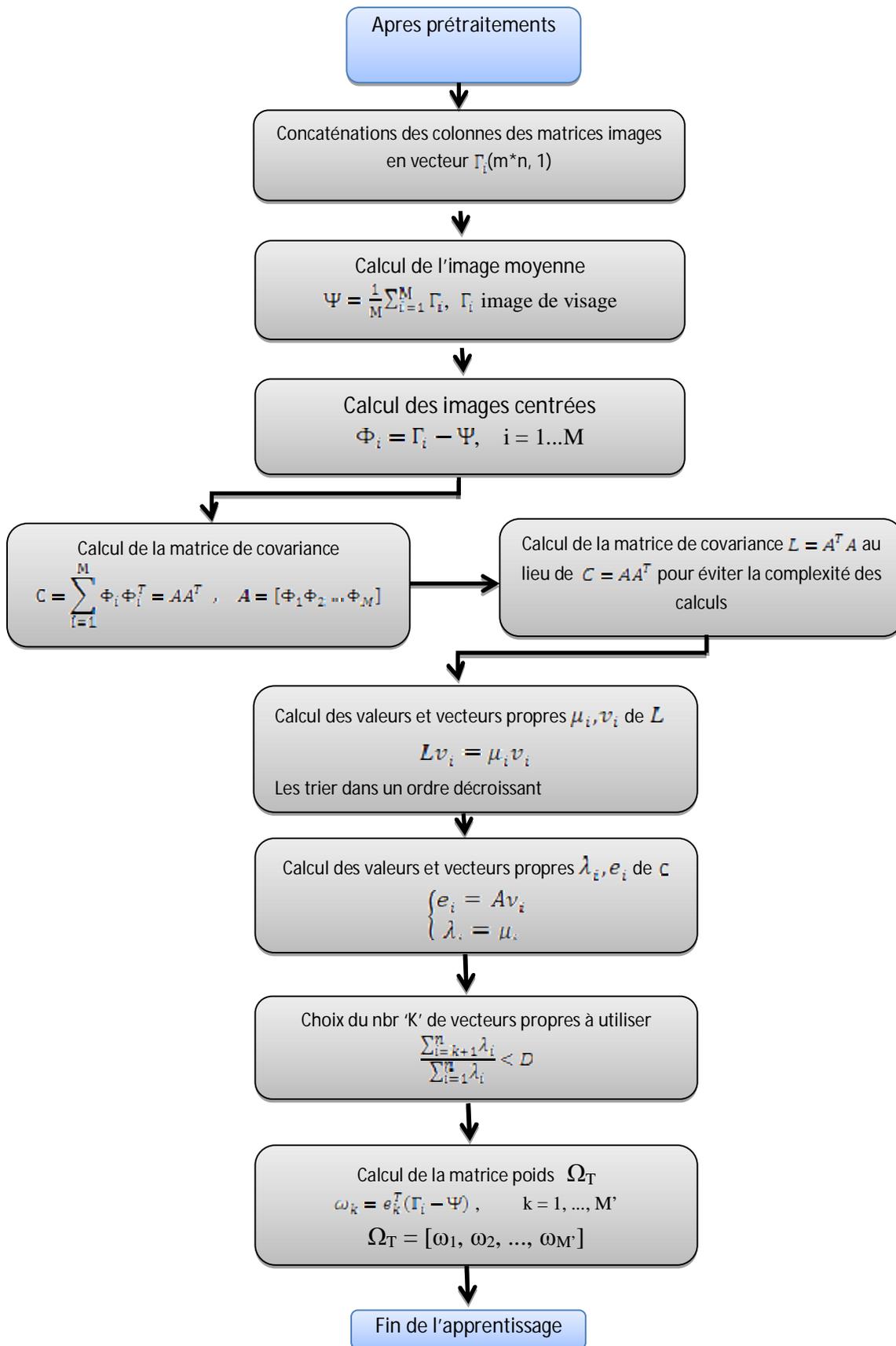


Figure 36 : phase d'apprentissage

III.4.3. Organigramme de la phase d'identification

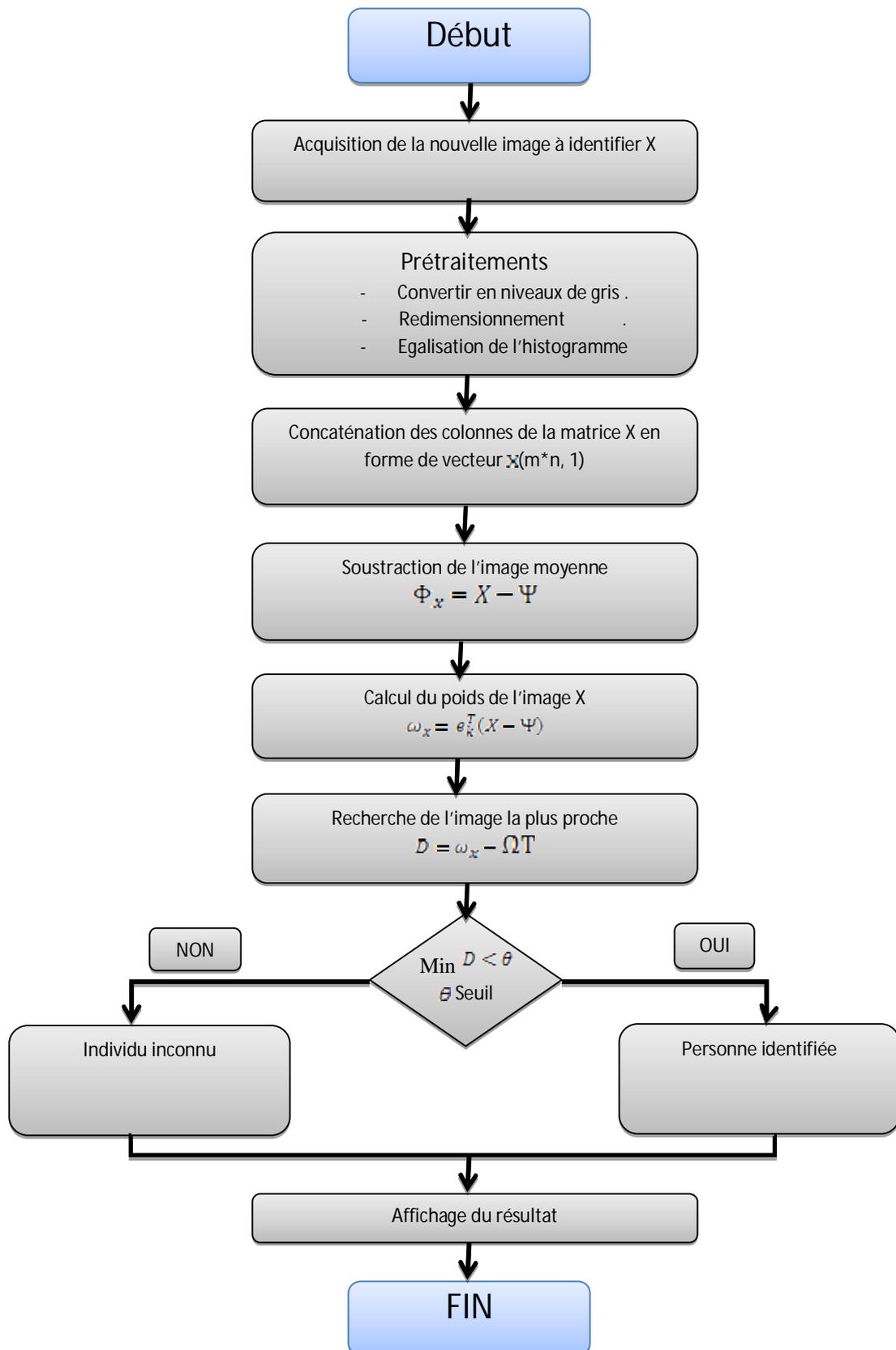


Figure 37 : phase d'identification

III.5. Les bases de données de visage

Plusieurs bases de données contenant des informations qui permettent l'évaluation des systèmes de reconnaissance de visages sont disponibles sur le marché. Toutefois, ces bases de données sont généralement adaptées aux besoins de quelques algorithmes spécifiques de reconnaissance, chacune d'elle a été construite avec des conditions d'acquisition d'images de visages diverse (changements d'illumination, de pose, d'expressions faciales) ainsi que le nombre de sessions pour chaque individu.

III.5.1 AT&T

Cette base contient des images de visage de 40 personnes, avec 10 images pour chacune. Pour la plupart des sujets, les 10 images ont été tirées à différentes heures et avec différentes conditions d'éclairage, mais toujours sur un fond foncé. La taille d'image est 92 x 112.

Les limitations de cette base de données sont : le nombre limité de personnes, la non-conformité des conditions d'éclairage d'une image à une image, et la non assignation de toutes les images par les différentes expressions faciales, rotation de la tête, ou conditions d'éclairage [58].

III.5.2 FEI

La base de données de visage de FEI est une base de données brésilienne de visage qui contient un ensemble d'images de visage prises entre juin 2005 et mars 2006 au laboratoire d'intelligence artificielle de FEI dans São Bernardo do Campo, São Paulo, Brésil. Elle contient des images de visage de 200 personnes, avec 14 images pour chacune d'entre elles, un total de 2800 images. Toutes les images sont en couleur et prises sur un fond blanc homogène en position frontale avec divers angles de pose de chaque personne, s'étendant du profil aux vues frontales. La variation d'éclairage est de 10% et la résolution de l'image est de 640x480 Pixel [58].

Le nombre de sujets masculins et féminin sont exactement identique et égale à 100. La figure 38 montre quelques exemples des variations d'image de la base de données de visage de FEI.



Figure 38 : Exemples d'images de visages de la base FEI

III.5.3. MUCT

La base de données de MUCT se compose de 3755 en couleur [58]. Elles ont été rassemblées en photographiant 345 sujets à de divers angles. La base de données a été créée pour fournir plus de diversité d'éclairage, d'âge, et d'appartenance ethnique, aussi pour développer, tester et évaluer des algorithmes de reconnaissance de visages. Les images couleurs ont une résolution : 512 x 768, en format JPG. Quelques exemples sont montrés dans la figure 39.



Figure 39 : Exemples d'images de visage de la base MUCT

III.5.4. CVSRP

- Nombre total des individus: 395
- Nombre d'images par individu: 20
- Nombre total des images: 7900
- Genre: contient des images des sujets masculins et féminins
- Ethnies : contient des images des personnes de diverses origines raciales
- Tranche D'Âge: les images sont principalement des étudiants de première année ainsi que des étudiants préparant une licence, la majorité des individus sont entre 18-20 ans, mais quelques individus plus âgés sont également présents.
- Lunettes : Oui
- Barbes: Oui
- Format d'image: JPEG de la couleur 24bit
- Appareil-photo utilisé: Camcorder de S-vhs
- Éclairage: artificiel.

Des exemples d'images de visage de cette base sont montrés dans la figure 40 [58]



Figure 40 exemples d'images de visage de la base CVSRP

III.6. Discussion

Ce chapitre a été consacré en premier lieu à la présentation de la méthode de reconnaissance faciale choisie qui est « Eigenface », de la structure des bases de données de visage et ensuite, nous avons décrit les différentes étapes de l'implémentation du système.

Dans le chapitre suivant, nous testerons le système dans son environnement afin d'évaluer ses performances pour en déduire les paramètres optimaux qui garantissent une meilleure efficacité du système. Enfin, nous proposerons quelques procédures pour améliorer les performances de notre système.

Chapitre IV

Tests et résultats

IV.1. Preambule

Ce chapitre représente les résultats des tests effectués avec la méthode des « Eigenface » sur les bases de données décrites dans le chapitre précédent. Rappelons que notre travail consiste à concevoir un système d'identification biométrique de personnes par reconnaissance faciale se basant sur une analyse en composante principale qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Plusieurs étapes sont nécessaires, l'étape d'extraction des caractéristiques est la plus importante car les performances du système en dépendent (résultats et robustesse, un temps de latence acceptable pour des applications « temps réel »). Nous évaluerons dans ce chapitre les résultats obtenus sur plusieurs bases de données sous différentes conditions et attributs de la méthodes.

IV.2 Environnement du travail

Dans cette section, nous présenterons les environnements matériel et logiciel de notre travail.

IV.2.1 Environnement matériel

Afin de mener à bien ce projet, il a été mis à notre disposition un ensemble de matériels dont les caractéristiques sont les suivantes :

Un ordinateur FUJITSU avec les caractéristiques suivantes :

- Processeur : Intel® Pentium® Dual CPU T3400 @ 2.16Ghz 2.17Ghz
- RAM : 2.00 Go de RAM
- Disque Dur : 250 Go
- OS : Microsoft Windows Seven

IV.2.2. Outils de développement :

Nous avons eu recours lors de l'élaboration de notre système à Matlab 7.12.0 que nous présenterons ci-dessous.

IV.2.2.1. Matlab 7.12.0 (R2011a) :

Matlab et son environnement interactif est un langage de haut niveau qui permet l'exécution de tâches nécessitant une grande puissance de calcul et dont la mise en œuvre sera bien plus simple et rapide qu'avec des langages de programmation traditionnels tels que le C, C++. Il dispose de plusieurs boîtes à outils en particulier celle du traitement d'images « ImageProcessing ToolBox » qui propose un ensemble d'algorithmes et d'outils graphiques de référence pour le traitement, l'analyse, la visualisation et le développement d'algorithmes de traitement d'images.

IV.3. Présentation de l'application :

On présente dans cette section les différents aspects du système de reconnaissance.

IV.3.1. Interface Présentation :

C'est une interface destinée aux utilisateurs, elle est simple et permet d'illustrer les principaux processus du système de reconnaissance (identification, visage moyen, reconstruction de

l'image, visage propre). Ces opérations sont effectuées sur les bases de données décrites dans le chapitre précédent.

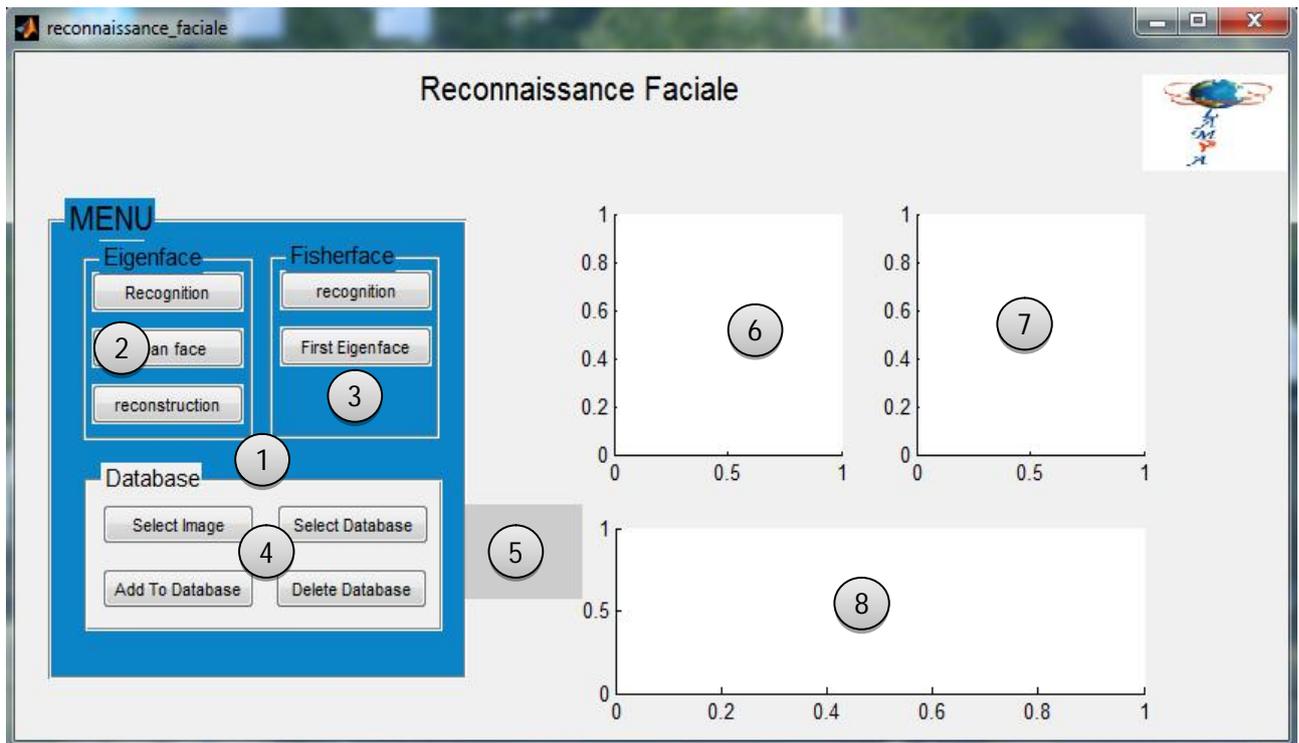


Figure 41 : Interface graphique de reconnaissance faciale

- 1- panel MENU : regroupant 3 sous panel (voir figure 42)
- ✓ Eigenface
 - ✓ Fisherface
 - ✓ Database

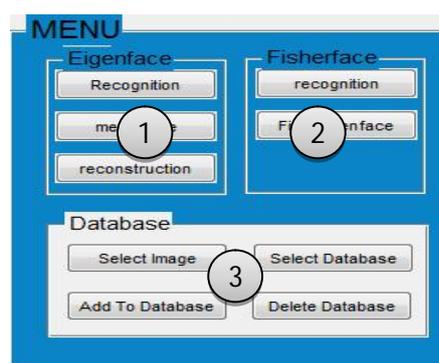


Figure 42 : panel MENU

- 2- Panel Eigenface : contient 3 boutons (voir figure 43) :
- ✓ Recognition : permet de lancer l'algorithme Eigenface pour la reconnaissance faciale ;
 - ✓ Mean Face : affiche le visage moyen de la base de données ;
 - ✓ Reconstruction : permet d'afficher l'image reconstruite du visage à identifier.

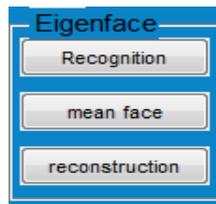


Figure 43 : panel Eigenface

- 3- Panel Fisherface : regroupant 2 boutons (voir figure 44) :
- ✓ Recognition : lance l'algorithme fisherface pour la reconnaissance faciale ;
 - ✓ First eigenface : affiche le premier visage propre de la base de données.



Figure 44 : panel Fishserface

- 4- Panel Database : regroupant 4 boutons (voir figure 4.5) :
- ✓ Select image : sélectionne l'image de visage à identifier ;
 - ✓ Select database : sélectionne la base de données à employer pour l'identification ;
 - ✓ Delete database : permet de supprimer la base de données ;
 - ✓ Add to database : permet d'ajouter une image de visage dans la base de données employée.

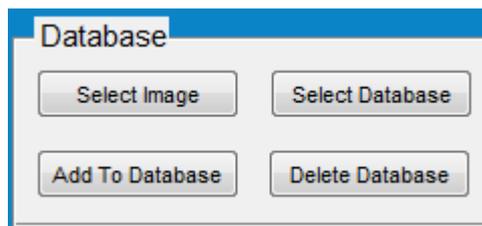


Figure 45 : panel Database

- 5- Static text : nous informe sur le nombre d'images de visage se trouvant dans la base de données ;
- 6- Axe 1 : affiche divers images de visages dont :
- L'image de visage à identifier ;
 - L'image de visage reconstruite ;
 - L'image moyenne de la base de données.
- 7- Axe 2 : affichage du résultat de la reconnaissance ;
- 8- Axe 3 : zone où s'affichent les différentes distances euclidiennes entre la projection du visage à identifier et les visages de la base de données.

IV.4. Démarche à suivre pour la reconnaissance faciale

1. Pour sélectionner une base de données, il faut cliquer sur le bouton **select database**, qui ouvre une fenêtre permettant de parcourir le disque et de choisir le dossier de la base de données du visage ciblé.



Figure 46 : recherche du dossier de la base de données

Un chiffre apparait dans le **static text** (voir figure 47) qui nous informe du nombre d'images existantes dans la base de données.

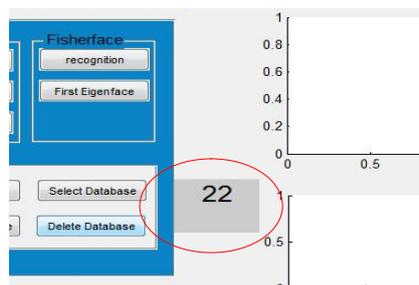


Figure 47 : Nbr d'image de visage dans la BD

2. pour sélectionner l'image de visage de la personne à identifier, il faut cliquer sur le bouton **select image** (voir figure 45), il permet d'ouvrir une fenêtre pour parcourir le disque et choisir l'image du visage à identifier.

3. lancer l'algorithme Eigenface en appuyant sur le bouton **Recognition** du panel Eigenface (voir figure 43), il apparaîtra un visage dans l'**Axe 2** qui sera le plus ressemblant (qui aura la plus petite distance euclidienne). Les distances euclidiennes entre la projection du visage à identifier dans un sous espace vectoriel et celles de la base de données seront affichées dans l'**Axe 3** (voir figure 4.8).

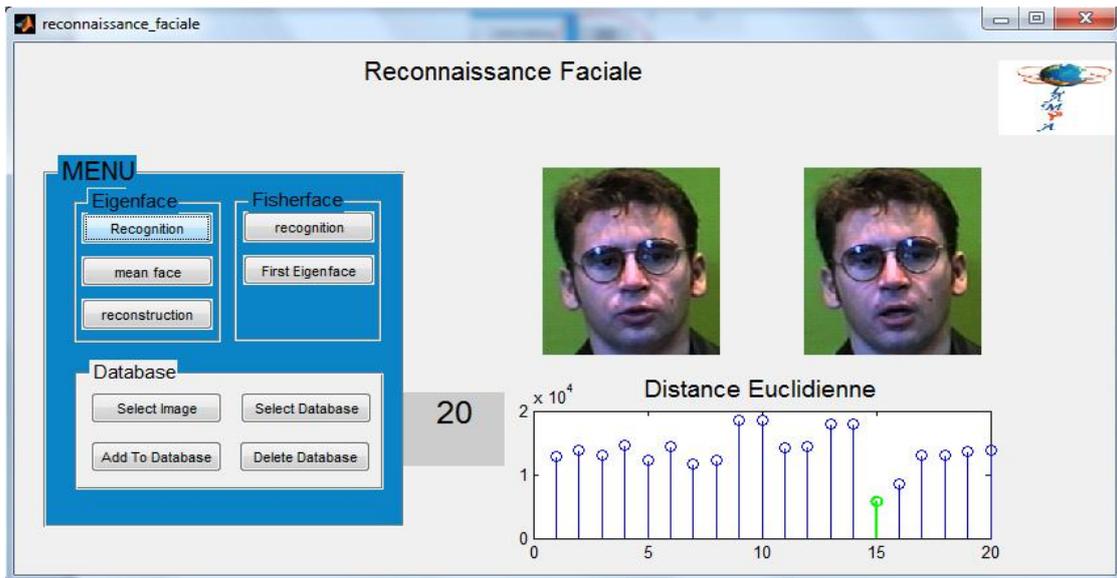


Figure 48 : Résultat de l’application de l’algorithme Eigenface

4. pour visualiser le visage moyen de la base de données, un clic sur le bouton **mean face** du panel Eigenface et il s’affichera sur l’Axe 1 (voir figure 43).

5. pour visualiser la qualité du visage reconstruit à partir du visage moyen et des visages propres de la base de données, cliquez sur le bouton **reconstruction** du panel Eigenface et il s’affichera dans l’Axe 1 (voir figure 49).

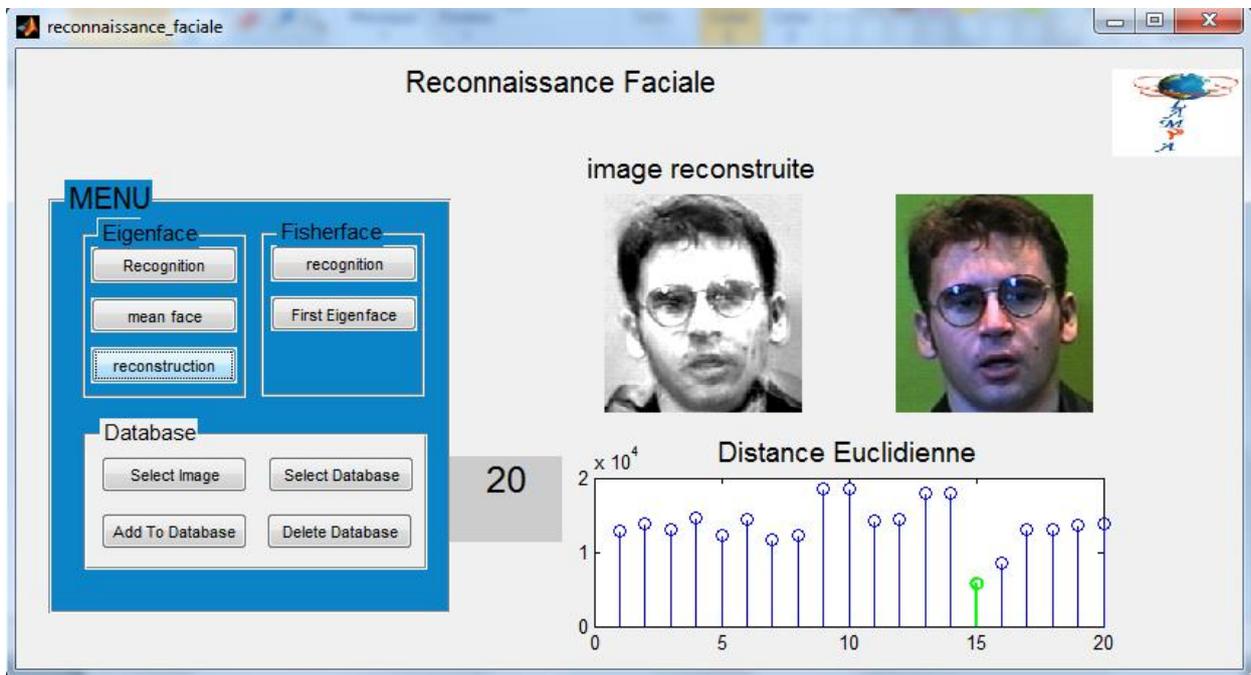


Figure 49 : Représentation de l’image reconstruite

6. pour lancer l'algorithme Fisherface après avoir choisi l'image du visage à identifier comme décrit dans l'étape 1, cliquez sur le bouton **recognition** du panel Fisherface et une fenêtre apparaîtra (voir figure 4.10) qui nous permettra de choisir le nombre de classe par personne existante dans la base de données, puis cliquer **OK** pour que l'algorithme poursuive son exécution jusqu'à ce qu'une image s'affiche sur l'**Axe 2** qui sera la plus proche (qui aura la plus petite distance euclidienne).

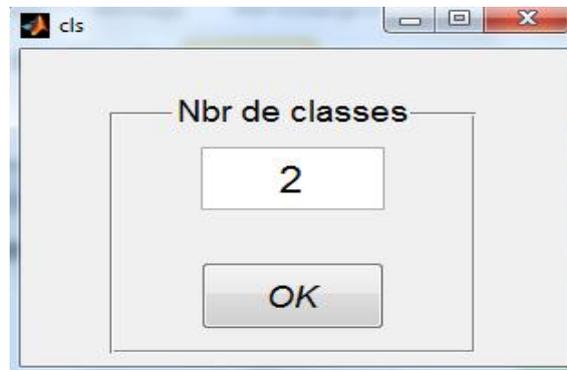


Figure 50 : Choix du nombre de classe

IV.5. Fonctionnement du système :

Notre système permet d'illustrer le processus d'identification. Nous allons décrire dans ce qui suit les différents événements qui s'y passent lors de ce processus.

Il comporte deux phases : la phase d'apprentissage pendant laquelle les modèles sont construits et la phase d'identification (phase de comparaison).

IV.5.1. Phase d'apprentissage :

Dans le cadre d'une méthode supervisée (les modèles doivent être connus à priori) pour l'identification des personnes par le visage, un modèle pour la représentation de chaque classe est construit. Le modèle peut être défini comme suit :

- Le vecteur correspondant à la distance minimale entre le vecteur recherché et l'ensemble de l'échantillon d'apprentissage, ceci permet d'aboutir à un modèle ayant les caractéristiques les plus proches du vecteur recherché au sens de la distance utilisée. Une fois les modèles des individus définis, l'opération d'identification peut s'effectuer, voici les étapes de la phase d'apprentissage :

- 1- Collecte des M images faciales et construction de la matrice T de taille M , par concaténation des colonnes des images faciales. Prétraitement des images collectées.
- 2- Calcul du visage moyen en sommant les colonnes de la matrice T et divisant le vecteur résultant par le nombre d'image d'entrée (M).
- 3- Soustraction du visage moyen de la matrice T pour obtenir la matrice A ; où chaque élément représente la variance des valeurs d'intensité de chaque pixel.

- 4- Calcul de la matrice \tilde{C} , $\tilde{C} = A^T A$.
- 5- Calcul des vecteurs propres de C' et les triés dans un ordre descendant selon les valeurs propres associées.
- 6- Calcul des vecteurs propres de la matrice de covariance C et obtention des visages propres en multipliant les vecteurs propres de C' par la matrice A .
- 7- Choix des K meilleurs valeurs propres et les vecteurs propres associés.
- 8- Détermination du poids des images d'entrée en projetant chaque image dans l'espace visage.
- 9- Chaque visage est maintenant représenté par un vecteur qui est utilisé pour reconstruire les images. Et enfin sauvegarde des calculs du visage moyen, des eigenfaces et du poids des images.

IV.5.2. Phase d'identification :

- ✓ Prétraitement de l'image d'entrée et soustraction du visage moyen.
 - ✓ Détermination du poids de l'image d'entrée par la projection de celle-ci dans l'espace visage en multipliant le vecteur résultant de l'étape (1) par les eigenfaces de la base de données.
 - ✓ Récupération du vecteur le plus proche au sens de la distance utilisée (euclidienne, Manhattan...).
 - ✓ Affichage du résultat (identité trouvée).
- **Phase d'identification :**

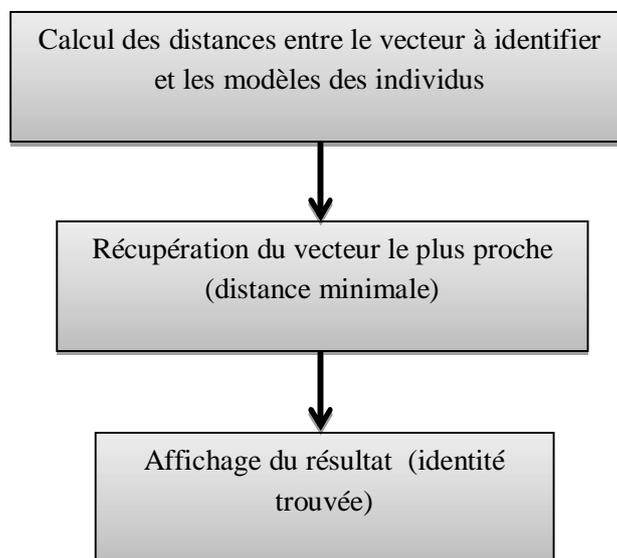


Figure 51 : Phase d'identification.

4.6. Détermination du seuil d'identification

Lors de la phase de décision, le système de reconnaissance associe à l'image à identifier, qu'elle soit d'une personne connue ou inconnue, une image de la base d'apprentissage (la plus ressemblante). D'où la nécessité d'utiliser un seuil de décision pour permettre l'accès seulement aux personnes connues du système.

La méthode de reconnaissance optimale est celle qui n'accepte pas les inconnus (taux de fausse acceptation FAR = 0) et qui ne rejette aucun des connus (taux de faux rejet FRR = 0). En réalité on ne peut pas avoir ceci, c'est pour cela qu'on cherche un seuil qui garantit à la fois un faible FRR et un faible FAR.

Pour trouver le seuil optimum, on calcule les FRR et FAR pour différentes valeurs du seuil. L'ensemble de ces résultats est représenté sous forme d'un graphe standard (courbe ROC).

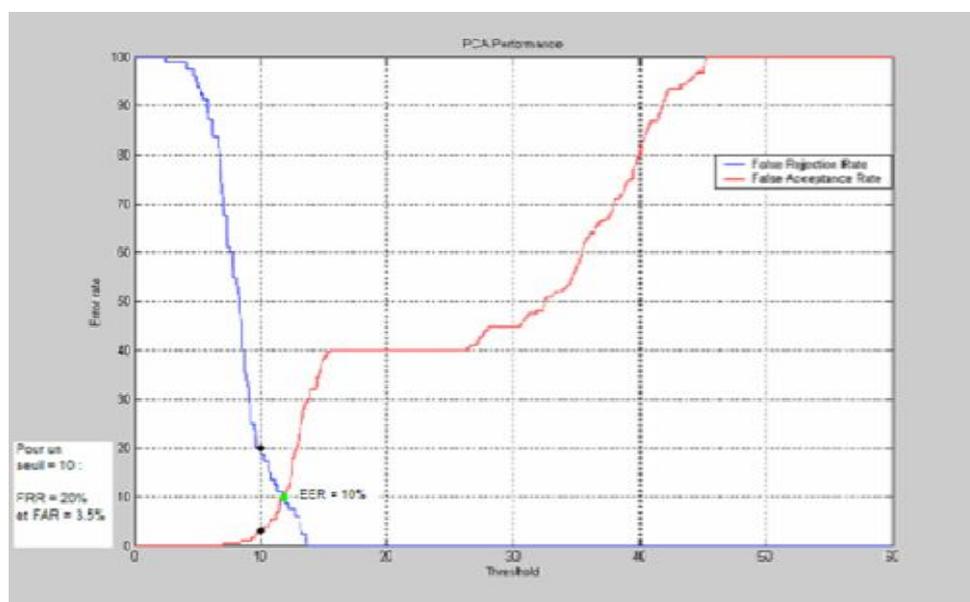


Figure 52 : Graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil

L'axe horizontal de ce graphe représente la valeur du seuil et l'axe vertical représente les valeurs de FRR et de FAR.

Toute base de données à utiliser est divisée en deux parties : les personnes incluses dans la base de données et les inconnues. Les images des personnes inclus dans la base de données sont elles aussi réparties en deux catégories : les images d'apprentissage et les images de test.

Dans la courbe du FAR et du FRR en fonction du seuil, on a l'information du taux du non rejet des sujets connus (acceptés par le système). C'est sujets là, peuvent être soit bien identifiés soit mal identifiés par le système. Et pour avoir l'information du taux de la bonne identification, on trace une courbe qui donne le taux de la bonne élimination (bon seuillage) des inconnus et le taux de la bonne classification des sujets connus, en fonction de la valeur du seuil.

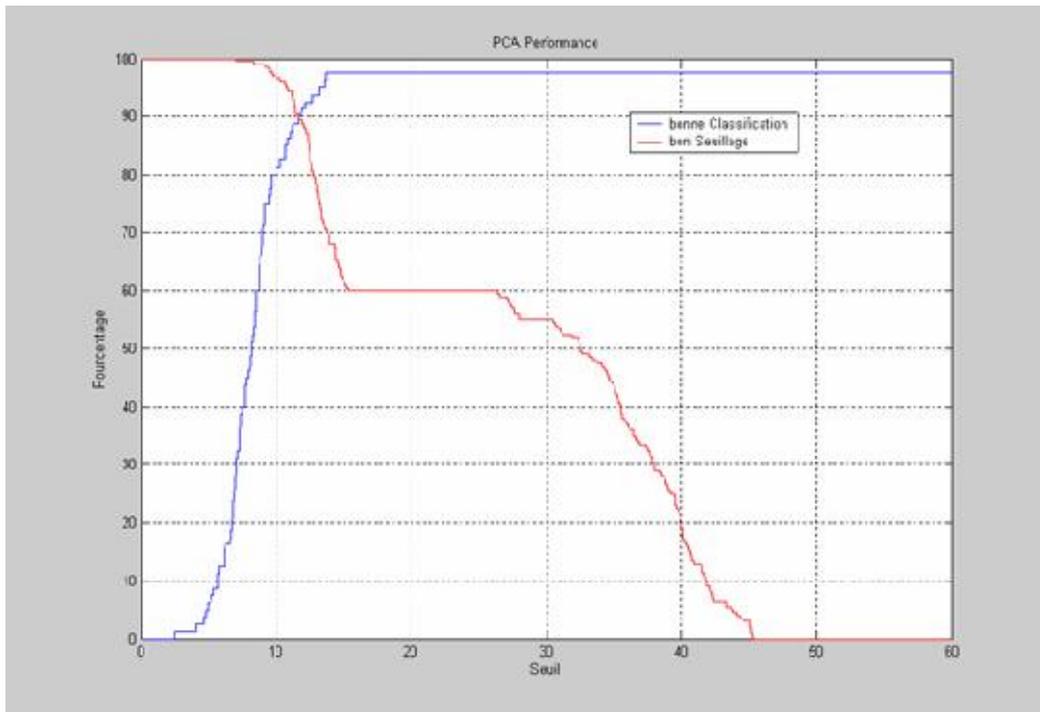


Figure 53: Graphe des taux du bon seuillage des inconnus et des taux de la bonne classification des sujets connus en fonction de la valeur du seuil

Pour les tests on a utilisé deux bases de données base, **FEI** et **MUCT**. On a pris 100 images d'apprentissage (20 personnes, avec 5 photos par personne), 60 images de test (20 personnes, avec 3 photos par personne), et 60 images d'inconnus (20 personnes, avec 3 photos par personne) de la base FEI et on a pris la même répartition pour la base MUCT.

Les figures ci-dessous représentent les différents ensembles d'images utilisés et les graphes les résultats des tests effectués sur la base MUCT



Figure 54 : L'ensemble des images d'apprentissage de la base MUCT utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues

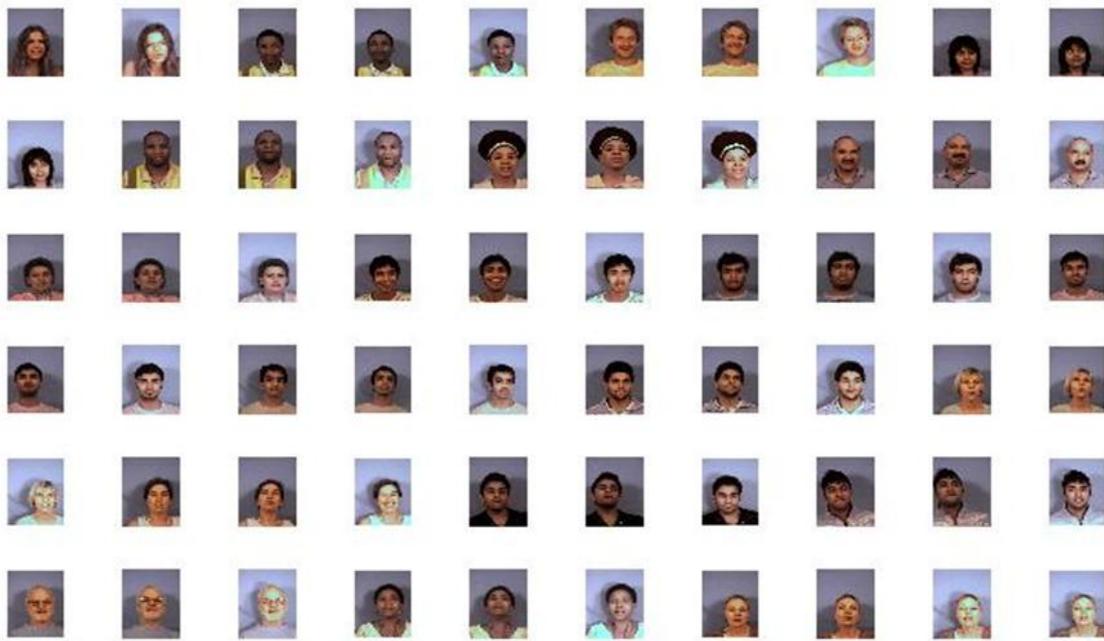


Figure 55 : L'ensemble des images de test de la base MUCT utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues



Figure 56 : L'ensemble des images des inconnus de la base MUCT utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues

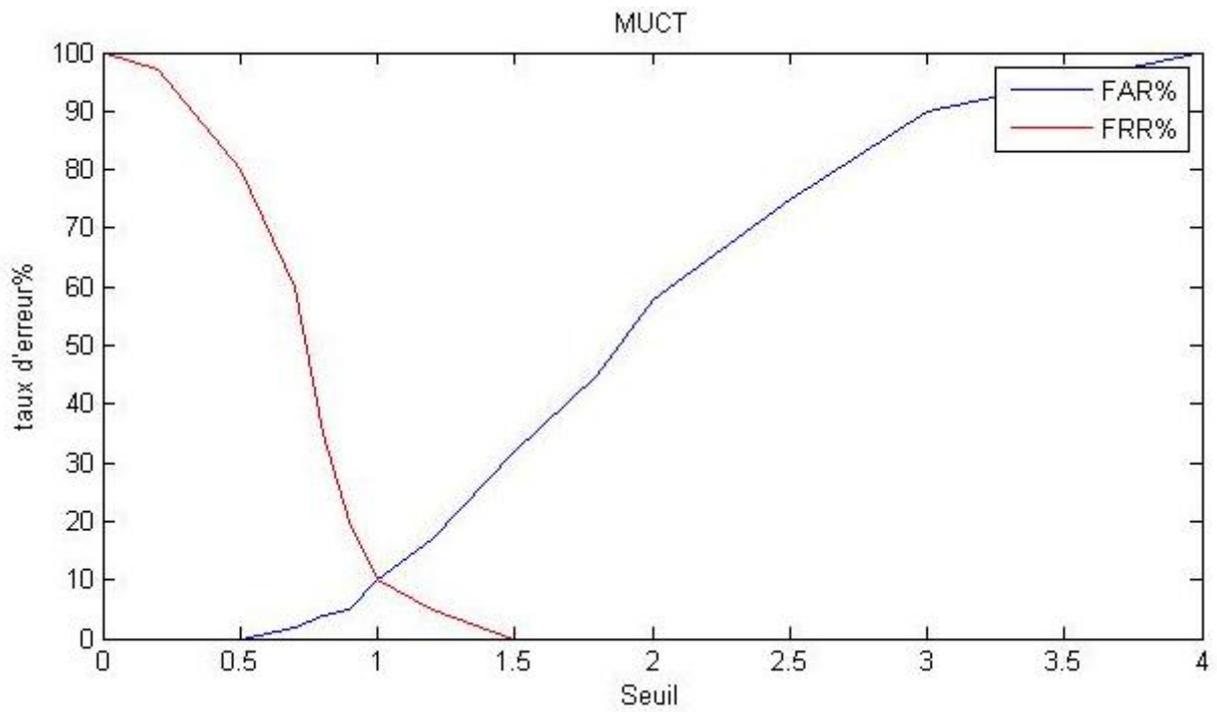


Figure 57 : Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil

Les figures ci-dessous représentent les différents ensembles d’images utilisés et les graphes les résultats des tests effectués sur la base FEI



Figure 58 : L’ensemble des images d’apprentissage de la base FEI utilisées pour la détermination d’un seuil pour l’élimination des personnes inconnues

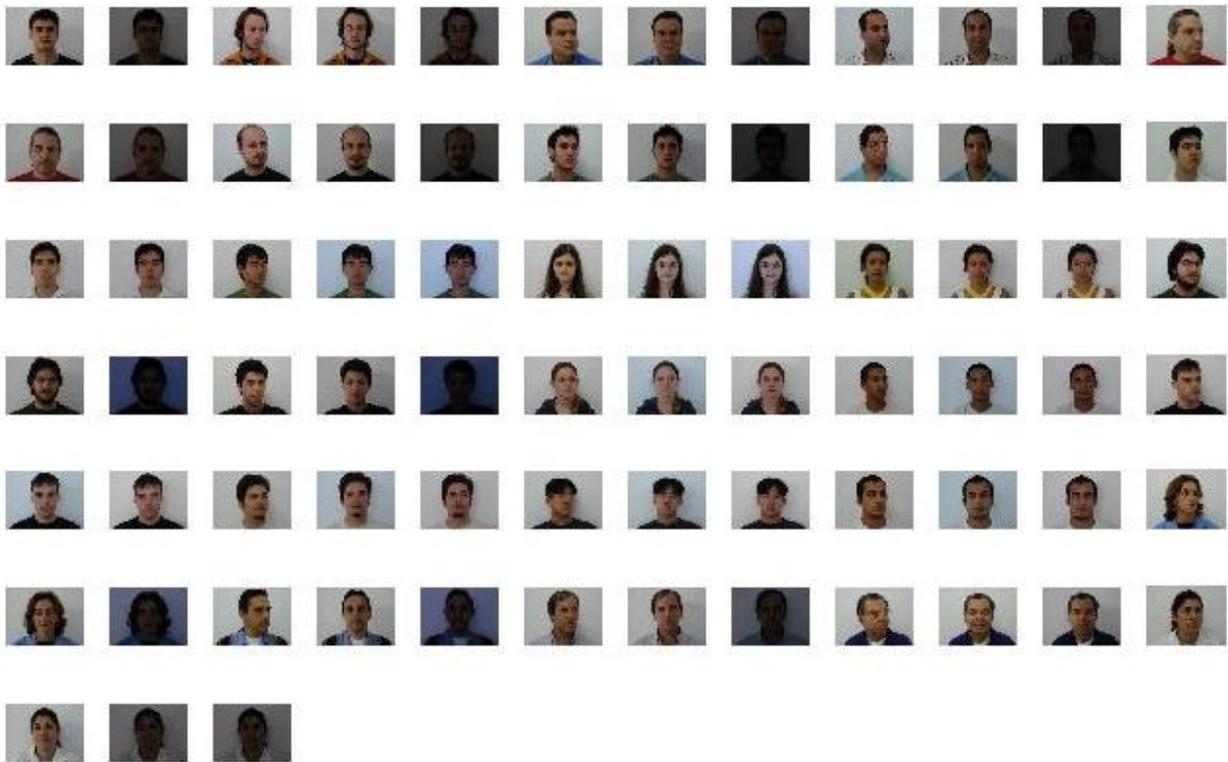


Figure 59 : L'ensemble des images de test de la base FEI utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues

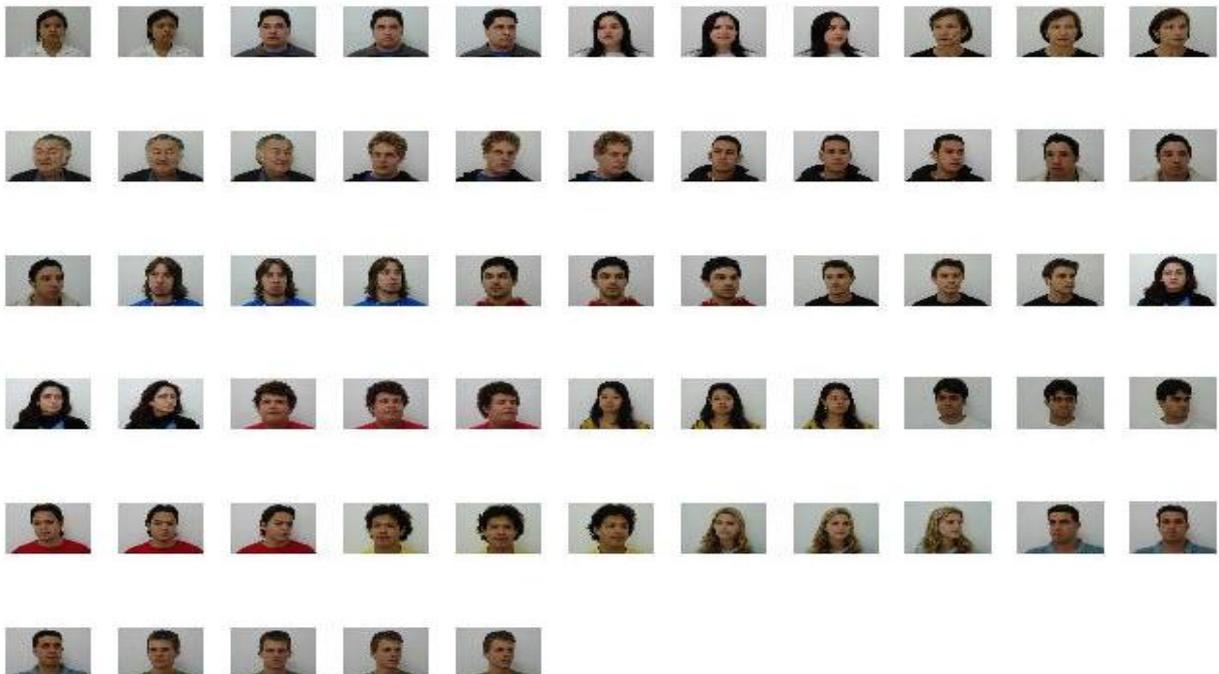


Figure 60 : L'ensemble des images des inconnus de la base FEI utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues

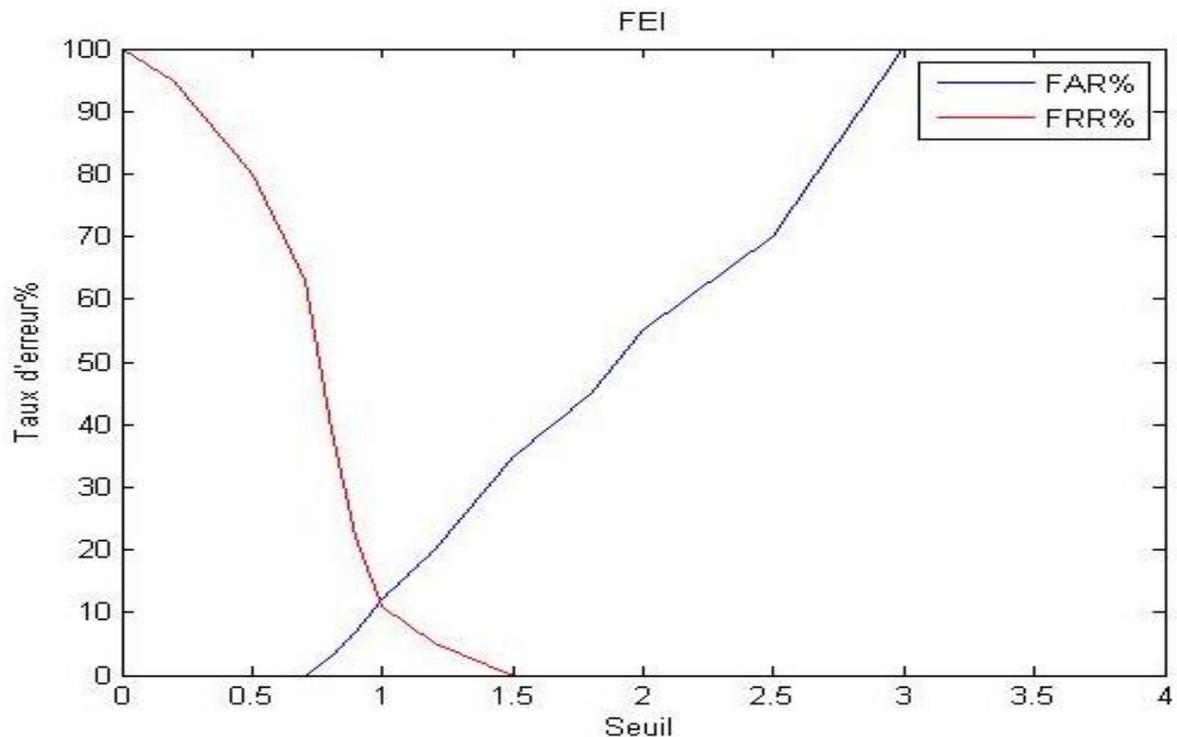


Figure 61 : Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil

On constate que lors de l'utilisation des deux bases de données, les seuils qui garantissent un bon compromis offrant à la fois un bas FAR et un bas FRR sont presque les mêmes pour les deux bases. Ces tests nous ont donnés des résultats assez concluants mais il est possible d'augmenter les performances de classification en fixant les meilleures conditions de poses et d'éclairages.

IV.7. Evolution du taux d'identification suivant le changement de poses et d'éclairages

Pour ce qui suit, nous allons voir les effets du changement de la pose et de l'éclairage sur le taux de bonne identification du système de reconnaissance :

Pour les tests on a utilisé deux bases de données : **AT&T** et **CVRP**. On a pris 90 images d'apprentissage (30 personnes avec 3 photos par personne), 60 images de test (30 personnes, avec 2 photos par personne), et 60 images d'inconnus (20 personnes, avec 3 photos par personne) de la base AT&T et on a pris la même répartition pour la base CVRP.

IV.7.1. Le changement de pose

Pour ce test on fera une comparaison entre une identification classique avec un seul exemple d'image de visage par individu (une seule classe) dans la base d'apprentissage et une identification avec une multitude d'image de visage dans divers poses, une du profile droite, une frontale et une du profile gauche (plusieurs classes).

Et voici les graphes relatifs aux deux bases de données **AT&T** et **CVRP**

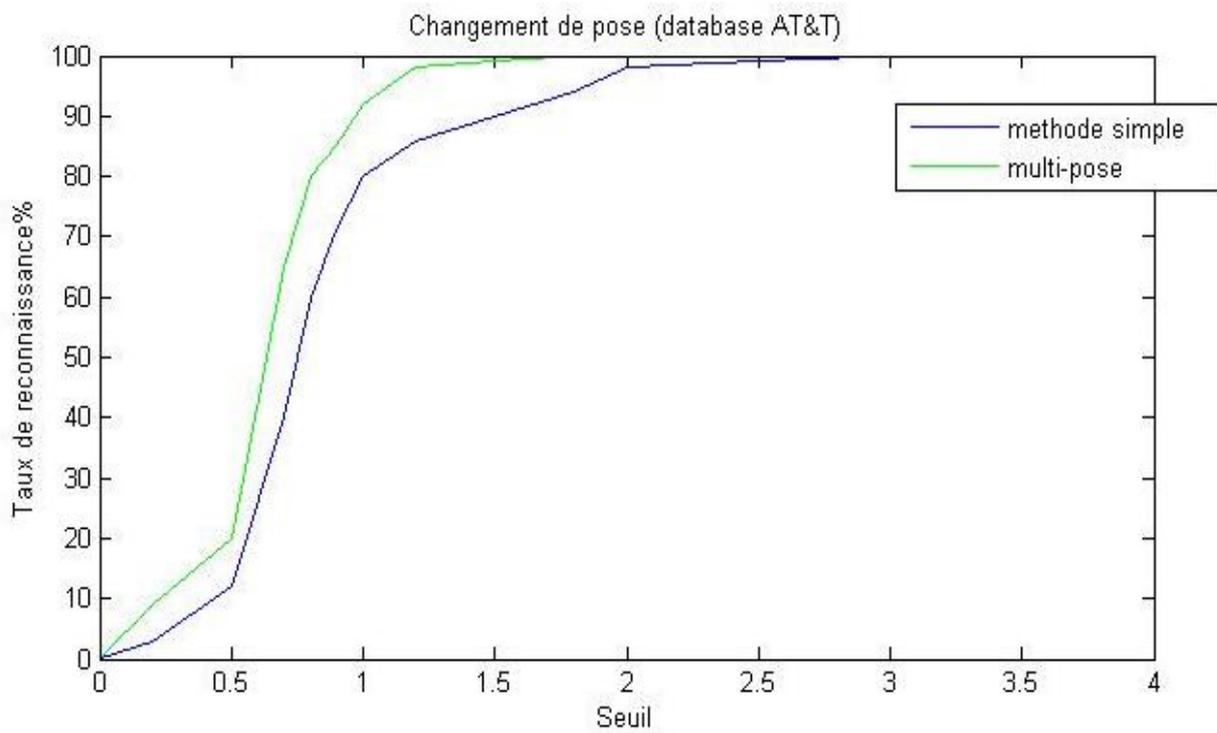


Figure 62 : Graphe du taux de reconnaissance avec la base AT&T

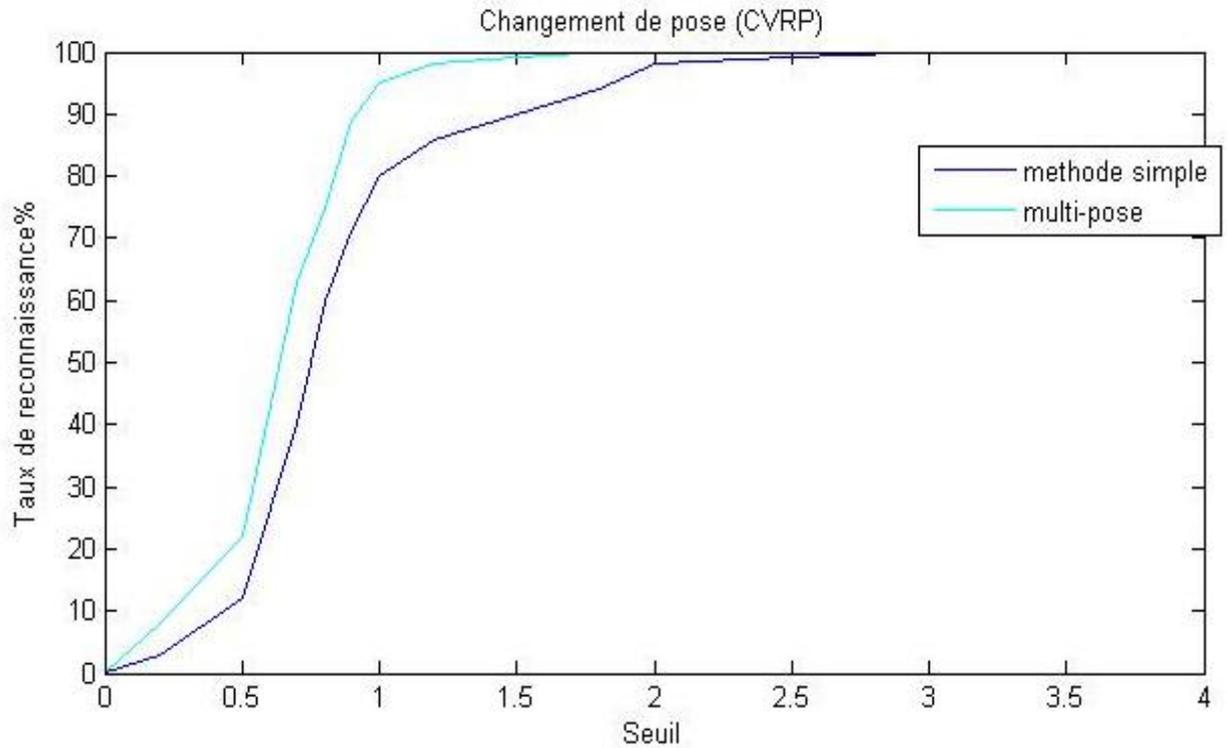


Figure 63 : Graphe du taux de reconnaissance avec la base CVRP

Et l'ensemble des résultats obtenus est résumé dans le tableau ci-dessous :

Bases de données	Identification simple	Identification à plusieurs classes
AT&T	83,7%	90,6%
CVRP	80,4%	91,2%

Tableau 1 : Comparaison du taux de reconnaissance de la méthode simple et celle avec divers poses

Dans le tableau 1, nous remarquons que la méthode utilisant plusieurs classes d'un même individu améliore le taux de reconnaissance par rapport à celle utilisant une seule, plus le nombre de classes augmente plus le taux de reconnaissance s'améliore, toute fois la pose doit respecter certaines règles : le profil du visage ne doit pas dépasser 35° par rapport à la pose frontale. L'inconvénient de cette méthode est qu'on se retrouve avec une base de données de taille importante d'où une augmentation du temps de calcul (exécution).

IV.7.2. Le changement des conditions d'éclairages

Pour ce test on fera une comparaison entre une identification avec et sans prétraitements en utilisant une base d'apprentissage contenant une multitude d'images de visage dans divers conditions d'éclairages, une à faible éclairages, une normal et une avec un fort éclairage.

Voici les graphes relatifs aux deux bases de données **AT&T** et **CVRP**

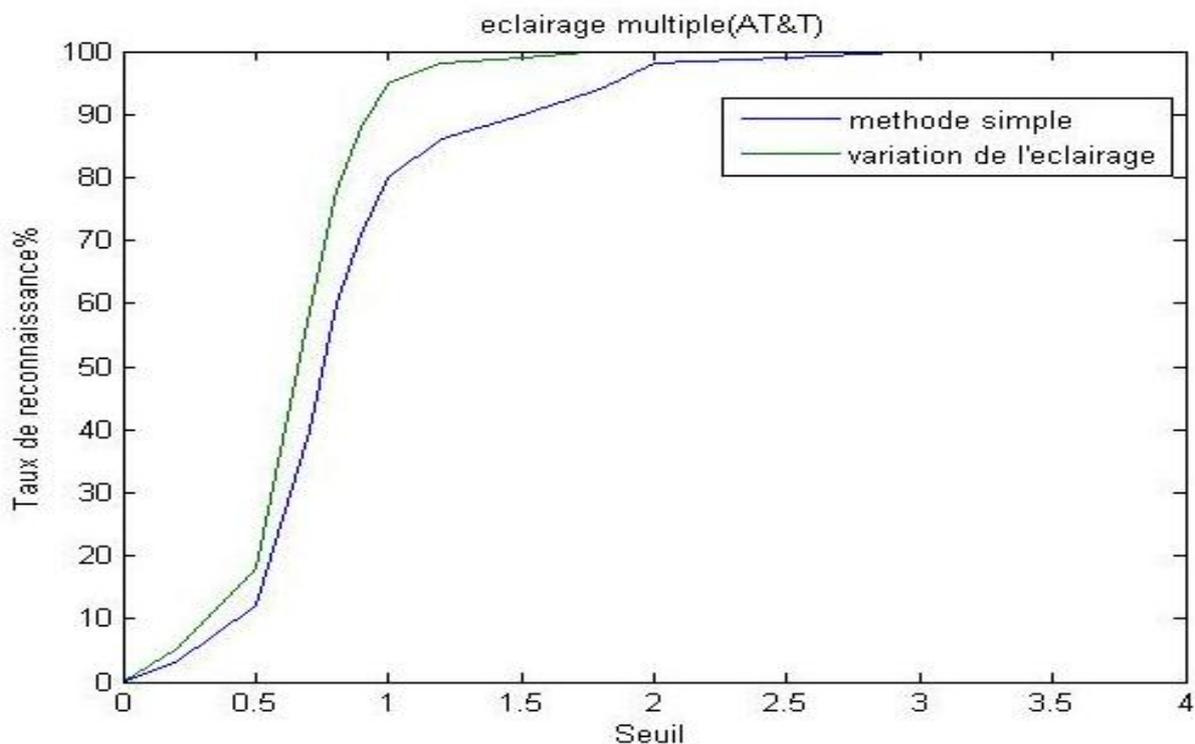


Figure 64 : Graphe du taux de reconnaissance avec la base AT&T

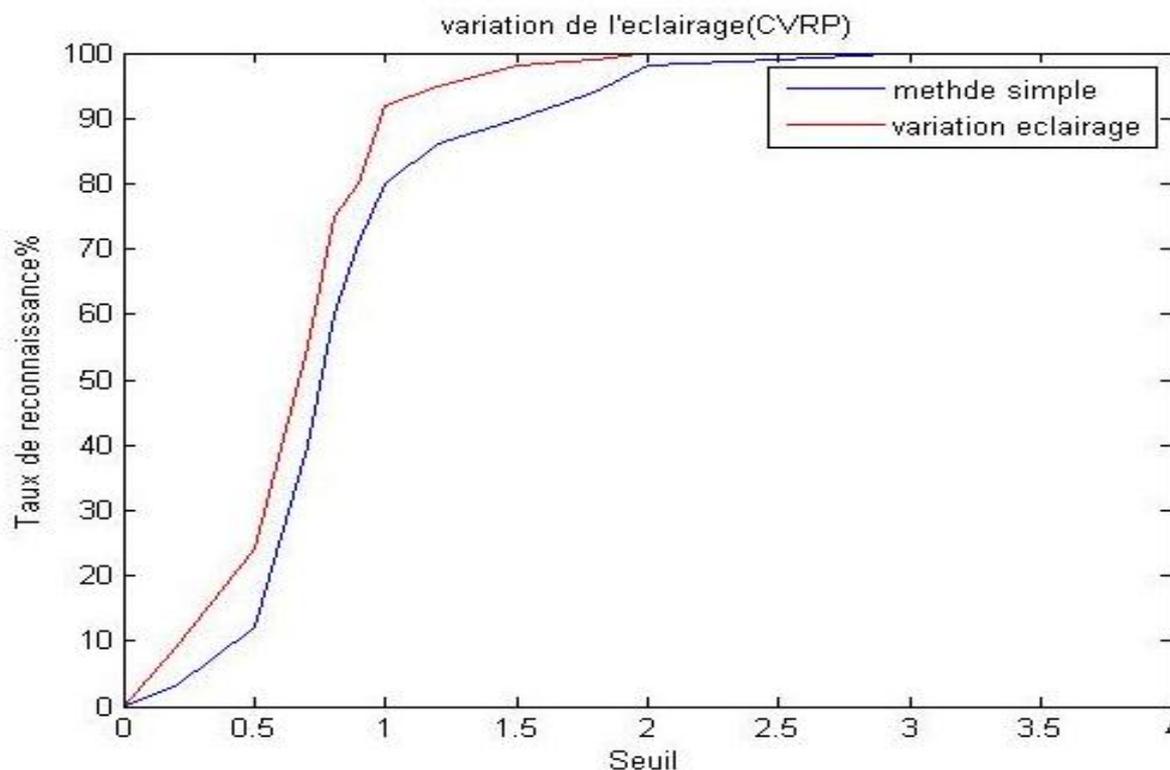


Figure 65 : Graphe du taux de reconnaissance avec la base CVRP

Et l'ensemble des résultats obtenus est résumé dans le tableau ci-dessous :

Bases de données	Sans prétraitements	Avec prétraitements
AT&T	83,2%	92,4%
CVRP	85,7%	90,1%

Tableau 2 : Comparaison du taux de reconnaissance de la méthode simple et celle avec divers conditions d'éclairages.

Ce qui ressort du tableau confirme que l'utilisation des prétraitements améliore les taux de reconnaissance surtout pour une base de données contenant des images prises sous de mauvaises conditions d'illuminations. Cependant, le temps de calcul est plus lent.

IV.8. Temps d'exécution

✓ Estimation du temps d'exécution en fonction de la taille de l'image

Pour ce test, nous avons redimensionné les images de visage normalisées de taille (480x640) de la base de données CVRP en différentes résolutions {25x50, 50x100, 100x150, 200x250, 300x400}. Pour chaque résolution, nous avons effectué une comparaison entre les signatures (projections) associées aux vecteurs image d'apprentissages et de tests selon la distance Euclidienne.

Le tableau 4.3 montre les résultats du taux d'identification obtenus pour les différentes valeurs de résolution ainsi que les temps de calcul. Nous remarquons qu'une diminution de la résolution, dans certain cas, n'engendre pas une baisse significative du taux d'identification. Nous constatons aussi que le taux d'identification le plus élevé correspond à une résolution de 300x400. Nous avons donc choisi de prendre une résolution de 200x250 pour avoir un meilleur compromis entre temps de calcul et précision.

	25x50	50x100	100x150	200x250	300x400
Taux d'identification	52%	72%	85%	90%	92%
Temps d'exécution	57s	93s	230	510s	1200s

Tableau 3 : Les résultats des EER, l'intervalle de confiance à 90% associé et le temps exécution

✓ Estimation du temps d'exécution avec et sans prétraitements

Nous comparons, dans ce paragraphe, le temps d'exécution de notre système d'identification avec et sans prétraitements. Le temps d'exécution pris par le processus d'identification sans prétraitements est relativement de faible coût par rapport à celui avec prétraitements, mais le taux d'identification se voit réduit aussi.

Base de données	Sans prétraitements		Avec prétraitements	
	Taux identifi	Temps execut	Taux identifi	Temps execut
AT&T	82,3%	406s	89,1%	492s
CVRP	80,7%	390s	90,6%	480s

Tableau 4 : Comparaison des temps d'exécution en seconde de l'approche avec et sans prétraitements

Le temps d'exécution varie suivant plusieurs paramètres. Tout d'abord, un nombre de classes important augmente la complexité de l'identification, la taille des vecteurs attributs (résolution de l'image) utilisés pour l'identification finale influe sur la vitesse d'exécution de la recherche. Des vecteurs attributs plus volumineux (nombre de classe) nécessitent plus d'espace mémoire et plus de temps de calcul. Et enfin, la phase de prétraitement qui est une étape cruciale pour l'obtention d'un bon taux d'identification génère des calculs qui risquent d'affecter visiblement le temps d'exécution du système.

Cette étude réalisée sur deux (02) base de données (AT&T, CVRP) de taille moyenne nous a permis de découvrir le potentiel de notre approche d'identification à la fois en termes de précision et de vitesse d'exécution. L'extension du corpus de travail à l'échelle d'une grande base de données est nécessaire pour évaluer correctement notre approche.

IV.9. L'amélioration des performances de PCA

Dans la littérature on parle toujours de l'effet néfaste à la fois de la pose et de l'éclairage, qui détériorent les performances de tout système biométrique basé sur le visage. Et dans notre approche, on s'est intéressé à rechercher des bonnes conditions de pose et d'éclairage à considérer pour améliorer les performances et exploiter ainsi toute la puissance de l'algorithme.

IV.9.1. Les conditions optimales de pose et d'éclairage

En général, il semble que tous les systèmes de reconnaissance basés sur le visage adoptent tous les mêmes conditions. Par exemple dans le système Face VACS-SDK Version 4.0.3 qui a été développé et commercialisé par la grande firme allemande Cognitec Systems GmbH, on a fixé les règles suivantes [60] :

✓ la procédure d'apprentissage

Une bonne procédure d'apprentissage, où on fait l'acquisition des images utilisées lors de cette phase, augmente sensiblement les performances de reconnaissance. Voici quelques procédés à suivre pour augmenter la qualité de l'apprentissage :

- Employez plusieurs images du visage, montrant des vues légèrement différentes. Ceci peut être accompli en demandant à l'utilisateur de faire de légers mouvements du visage dans la direction horizontale et verticale ou en employant un arrangement approprié de plusieurs appareils photo.
- Informez l'utilisateur que prendre des photos d'apprentissage n'est pas identique à faire un portrait chez le photographe. Ainsi il ne devrait ni poser ni sourire, mais montrer une expression journalière du visage, qui sera susceptible d'être la même lors de la vérification.
- Si le système éprouve des difficultés à reconnaître les personnes portant des lunettes, fusionnez des images avec et sans lunettes dans la phase d'apprentissage, ou demandez à la personne de faire l'apprentissage et la reconnaissance sans lunettes.

✓ l'influence et l'arrangement des conditions d'éclairage

Dans les circonstances où on peut contrôler les conditions d'éclairage, on devrait prêter l'attention aux recommandations suivantes :

- L'éclairage du visage devrait être une lumière frontale diffuse ou dirigée, pour éviter les ombres volatiles dans la région de visage.
- Éviter l'éclairage produisant de la lueur dans les lunettes, ou sur une peau brillante.
- L'éclairage du visage ne devrait pas changer avec la position ou la taille de l'utilisateur.
- Protégez la zone biométrique contre l'éclairage additionnel, particulièrement contre, la lumière du soleil.

Les conditions d'éclairage devraient être tout à fait semblables sinon identiques, à la fois dans la phase d'apprentissage et la phase de reconnaissance.

Dans beaucoup de cas, des tubes fluorescents placés à gauche et à droite de l'appareil photo, fournissent une illumination appropriée.

IV.10. Avantages et inconvénients du système :

Dans ce chapitre, nous avons décrit les étapes de la conception du système de reconnaissance, et l'interface développée pour lui. En vue de ce qui a été fait dans la conception et la réalisation, on peut dégager plusieurs avantages :

✓ Système très interactif :

C'est un système très interactif. En effet, en plus du fait qu'il permette l'identification d'individus avec de bons taux, il permet aussi de voir la reconstruction du visage après sa projection dans un sous espace vectoriel, le visage moyen, que la distance entre le vecteur projeté ainsi qu'il nous informe sur le nombre d'images de la base de donnée

✓ Système performant :

Le système est performant car les taux de reconnaissance sont bons.

✓ Interface simple et rapide :

C'est une interface que toute personne ayant peu de connaissances sur la biométrie peut utiliser.

Néanmoins, ce système présente quelques faiblesses, entre autres :

✓ Processus de normalisation long :

Cet inconvénient est dû aux prétraitements qui sont appliqués sur toutes les images de la base de données pour réduire l'impact néfaste du milieu extérieur sur le système en appliquant un redimensionnement et une égalisation de l'histogramme sur les images.

✓ Absence d'un système d'acquisition des images :

Faute de la disponibilité d'un système d'acquisition, les seules données utilisées par le système proviennent des bases. Cela rend les opérations d'identification une simple simulation.

✓ Liaison du système à l'environnement :

L'emplacement des données biométriques utilisées par le système le rend dépendant de l'environnement dans lequel le système s'exécute.

IV.11. Discussion

Dans ce chapitre nous avons testé notre système de reconnaissance sur plusieurs bases de données à savoir : CVRP, AT&T, FEI et MUCT. Le seuil de rejet des individus inconnus a été trouvé en minimisant le taux de l'EER. De plus, nous avons étudié l'influence des changements de poses et d'éclairages sur le taux de reconnaissance. L'ensemble des tests effectués a permis de conclure, qu'avec l'utilisation de PCA sous un éclairage uniforme tout en ayant plusieurs poses par personne permet d'obtenir des performances optimales.

En validant notre approche sur un corpus de données de grande taille et assez varié le rendant ainsi plus réaliste, nous avons apporté une amélioration considérable au taux d'identification grâce aux prétraitements effectués en amont sur les images d'apprentissages et aux tests, toute fois le temps de calcul s'est vu légèrement augmenté.

Plusieurs enseignements peuvent être tirés pour minimiser ce temps de calcul comme le choix de la distance Euclidienne L2 qui est beaucoup plus rapide que la distance de Mahalanobis L1 et de la résolution des images de la base de données adéquate.

Conclusion :

La biométrie est un domaine à la fois passionnant et complexe. Elle tente, par des outils mathématiques souvent très évolués, de faire la distinction entre des individus, nous obligeant à travailler dans un contexte de très grande diversité. Cette diversité se retrouve également dans le nombre considérable d'algorithmes qui ont été proposés en reconnaissance faciale.

Dans ce mémoire, nous nous sommes intéressés au problème de la reconnaissance faciale. Notre travail consiste à la mise au point d'un algorithme robuste destiné à reconnaître un individu par son visage en utilisant la méthode « Eigenface » qui se base sur une analyse en composante principale (ACP). L'ACP est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard.

Une analyse des différentes techniques de reconnaissance développées au cours de ces dernières années a été présentée, et cela pour mettre en évidence les particularités ainsi que les avantages et les inconvénients de chacune d'entre elles.

Nous avons souligné durant ce travail l'influence néfaste du changement de pose et d'éclairage sur le taux de reconnaissance du système et pour cela nous avons proposé quelques solutions qui ont été évaluées durant la phase de test. Ces solutions ont donné d'assez bons résultats. Malgré tous les progrès qui ont été réalisés, les problèmes de pose et d'éclairage et de l'identification dans des environnements extérieurs restent des challenges qui susciteront les efforts des chercheurs. L'ACP reste une méthode efficace et simple pour gérer ce type de problème. C'est pour toutes ces raisons que nous avons opté pour cette approche de reconnaissance du visage.

Nous estimons avoir réalisé un système répondant à l'objectif que nous nous sommes fixés au départ, à savoir la mise en œuvre d'un système permettant la reconnaissance d'individus et le contrôle d'accès.

En guise de perspectives, une extension de ce travail peut être réalisée en intégrant un système d'acquisition des images pour éviter que ces opérations d'identification et de vérification soient de simples simulations. Par ailleurs, il serait aussi intéressant d'appliquer, la méthode développée dans ce mémoire, dans un système de sécurité biométrique utilisant l'acquisition vidéo. Dans ce cas, cet algorithme doit répondre aux exigences du temps réel. Mais, cela ne peut se faire sans le développement d'algorithmes efficaces de détection et de poursuite des visages dans les séquences vidéo.

Bibliographie

- [1] John D. Woodward, Jr., Christopher Horn, Julius Gatune, and Aryn Thomas, “Biometrics A Look at Facial Recognition”, documented briefing by RAND Public Safety and Justice for the Virginia State Crime Commission, 2003.
- [2] Florent Perronnin, Jean-Luc Dugelay, « Introduction à la biométrie : Authentification des individus par traitement audio-vidéo », Institut Eurocom, Multimedia Communications Department, Revue Traitement du signal, Vol. 19, N° 4, 2002.
- [3] <http://www.biometricgroup.com>.
- [4] S. Liu, M. Silveanu, « A practical Guide to Biometric Security Technology », IEEE Computer Society, IT Pro-Security, Janvier-Février 2001.
- [5] A. K. Jain, L. Hong, S. Pankanti, « Biometrics : Promising Frontiers for Emerging Identification Market », Communications of the ACM, pp. 91-98, February 2000.
- [6] C. Fredouille, J. Mariethoz, C. Jaboulet, J. Hennebert, J.-F. Bonastre, C. Mokbel, F. Bimbot, « Behavior of a Bayesian Adaptation Method for Incremental Enrollment in Speaker Verification », International Conference on Acoustics, Speech, and Signal Processing, pp. 1197-1200, Istanbul, Turquie, 5-9 Juin 2000.
- [7] Y. Hori, M. Kusaka, and T. Kuroda. “A 0.79mm² 29mW Real-Time Face Detection Core”. Symposium on VLSI Circuits Digest of Technical Papers, pp. 188-189, June 2006.
- [8] D. Bolme, J. Beveridge, M. Teixeira, and B. Draper. “The CSU Face Identification Evaluation System : Its Purpose, Features, and Structure”. In : Proceedings of the 3rd International Conference on Computer Vision Systems (ICVS),
- [9] W. Bledsoe. “The model method in facial recognition”. Tech. Rep. PRI :15, Stanford University, Palo Alto, CA, USA, 1964.
- [10] T. Kanade. “Computer Recognition of Human Faces”. Interdisciplinary Systems Research, Vol. 47, 1977.
- [11] L. Sirovich and M. Kirby. “Low-dimensional procedure for the characterization of human faces”. Journal of Optical Society of America, Vol. 4, No. 3, pp. 519–524, 1987.

- [12] M. Turk and A. Pentland. “Eigenfaces for recognition”. *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71–86, 1991.
- [13] Y. Jian, D. Zhang, A. Frangi, and J.-Y. Yang. “Two-Dimensional PCA : A New Approach to Appearance-Based Face Representation and Recognition”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 1, pp. 131–137, January 2004.
- [14] P. Belhumeur, J. Hespanha, and D. Kriegman. “Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, pp. 711–720, 1997.
- [15] M. Bartlett, J. Movellan, and T. Sejnowski. “Face recognition by independent component analysis”. *IEEE Transactions on Neural Networks*, Vol. 13, No. 6, pp. 1450–1464, 2002.
- [16] G. Brown, S. Yamada, and T. Sejnowski. “Independent Component Analysis at the Neural Cocktail Party”. *Trends in Neuroscience*, Vol. 24, pp. 54–63, 2001.
- [17] B. Schölkopf, A. Smola, and K.-R. Müller. “Nonlinear component analysis as a kernel eigenvalue problem”. *Neural Computation*, Vol. 10, No. 5, pp. 1299–1319, 1998.
- [18] S. Mika, G. Ratsch, J. Weston, B. Schölkopf, and K.-R. Müller. “Fisher Discriminant Analysis With Kernels”. In : *Neural Networks for Signal Processing IX*, pp. 41–48, 1999.
- [19] C. Liu and H. Wechsler. “A Unified Bayesian Framework for Face Recognition”. In : *Proceedings of the IEEE International Conference on Image Processing*, pp. 151–155, 1998.
- [20] G. Guo, S. Li, and K. Chan. “Face Recognition by Support Vector Machines”. In : *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 196–201, 2000.
- [21] T. Ahonen, A. Hadid, and M. Pietikainen. *Face Recognition with Local Binary Patterns*. 2004.
- [22] S. Arca, P. Campadelli, and R. Lanzarotti. “A Face Recognition System Based On Automatically Determined Facial Fiducial Points”. *Pattern Recognition*, Vol. 39, No. 3, pp. 432–443, 2006.

- [23] P. Penev and J. Atick. "Local features analysis : A general statistical theory for object representation". *Neural Systems*, Vol. 7, No. 3, pp. 477–500, 1996.
- [24] V. Perlibakas. "Face recognition using Principal Component Analysis and Log-Gabor Filters". March 2005.
- [25] Y. Adini, Y. Moses, S. Ullman, Face recognition: The problem of compensating for changes in illumination direction. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 721–732, 1997.
- [26] D. Blackburn, M. Bone, P. J Phillips. "Face recognition vendor test 2000". Tech. rep. <http://www.frvt.org>, 2001.
- [27] R. Gross, J. Shi, J. Cohn. Quo Vadis Face Recognition?. Third Workshop on Empirical Evaluation Methods in Computer Vision, December, 2001.
- [28] A.M Martínez, R. Banavente. The AR face database. Tech. Report 24 CVC Barcelone, Espagne, June 1998.
- [29] D. Blackburn, M. Bone, P. J Phillips. "Face recognition vendor test 2000". Tech. rep. <http://www.frvt.org>, 2001.
- [30] Ming-Hsuan Yang, David J. Kriegman et Narendra Ahuja. Detecting faces in images : A survey. Dans *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24(1), pages 34–58, 2002.
- [31] H. A. Rowley, S. Baluja, et T. Kanade, "Neural Network based Face Detection", *IEEE Trans, Pattern Anal. Mach, Intell*, 23-38, January 1998
- [32] C. Kotropoulos and I. Pitas. Rule-Based Face Detection in Frontal Views. *Proc. Int'l Conf. Acoustics, Speech and Signal Processing*, vol. 4, pp. 2537-2540, 1997.
- [33] T. Kanade. "Picture Processing by Computer Complex and Recognition of Human Faces," PhD thesis, Kyoto Univ., 1973.
- [34] J.Yang, D. Zhang, A.F. Frangi, J.Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 131–137,2004.
- [35] P. Sinha, "Processing and Recognizing 3D Forms," PhD thesis, Massachusetts Inst. of Technology, 1995.

- [36] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces Using Deformable Templates," *Int'l J. Computer Vision*, vol. 8, no. 2, pp. 99-111, 1992.
- [37] R. Brunelli, T. Poggio. Face recognition: features versus templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, pp-1042–1062, 1993.
- [38] "Principal components analysis".
Document available at : http://en.wikipedia.org/wiki/Principal_components_analysis.
- [39] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby, "Face Recognition: A Literature Review", *INTERNATIONAL JOURNAL OF SIGNAL PROCESSING VOLUME 2 NUMBER 2 2005 ISSN 1304-4494*.
- [40] Cherng Jye Liou, "A Real Time Face Recognition System", DSP/IC Design Lab, Department of Electrical Engineering, National Taiwan University, June 1997.
- [41] Alex Chirokov, "Karhunen-Loeve, for face recognition". Matlab code available at : <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=6995&objectType=file>.
- [42] "Linear discriminant analysis".
Document available at : http://en.wikipedia.org/wiki/Linear_discriminant_analysis.
- [43] Dr. Andrzej Drygajlo, ELE 233, "BIOMETRICS".
Document available at : <http://scgwww.epfl.ch/courses>.
- [44] Juwei Lu, Kostantinos N. Plataniotis, and Anastasios N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms", *IEEE TRANSACTIONS ON NEURAL NETWORKS*, VOL. 14, NO. 1, JANUARY 2003.
- [45] M. Bartlett and H. Lades and T. Sejnowski, "Independent component representations for face recognition", in T. Rogowitz, B. & Pappas, editor, *Proceedings of the SPIE Symposium on Electronic Imaging: Science and Technology; Human Vision and Electronic Imaging III*, volume 3299, San Jose, CA, January 1998. SPIE Press. 1998.
- [46] Wen Gao, Shiguang Shan, Face Verification for Access Control, "Biometrics Solutions for Authentication in an E-World", Edited by David Zhang, Kluwer Academic Publishers, Chapter 13, pp339-376, 2002.
- [47] "Modules complémentaires à MATLAB".
Document disponible sur : <http://www.ryounes.net/cours/modules.pdf>.

- [48] L. Png, “Morphological Shared-Weight Neural Network For Face Recognition”, A dissertation submitted to the University of Manchester Institute of Science and Technology for the degree of MSc, August 2004.
- [49] “Support vector machine”.
Document available at : http://en.wikipedia.org/wiki/Support_vector_machine.
- [50] “Visual Object Recognition using Template Matching”.
Document available at : <http://www.araa.asn.au/acra/acra2004/papers/cole.pdf>.
- [51] Cotrell and Fleming. Face recognition using unsupervised feature extraction. In Proceedings of International conference on Neural Network, pages 322–325, Paris, France, 1990.
- [52] S. Lawrence, C.L. Giles, C. Tsoi, and A.D. Back. Face recognition : A convolutional neural network approach. IEEE Transactions on Neural Networks, 8 :98–113, 1997.
- [53] S.H. Lin, S.Y. Kung, and Lin. Face recognition / detection by probabilistic decision-based neural network. IEEE Transactions on Neural Networks, 8(1) :114–132, 1997
- [54] S.C. Chen, J. Liu, Z.-H. Zhou, MakinG. FLDA applicable to face recognition with one sample per person, Pattern Recognition 37 (7), 1553–1555, 2004.
- [55] 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, pp. 373–378, 1992.
- [56] 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, pp. 318–321, 2004
- [57] téléchargement des images de visage teste
http://en.pudn.com/downloads339/sourcecode/graph/detail1485237_en.html
- [58] Téléchargement de la base de données de visage
http://web.mit.edu/emeyers/www/face_databases.html#orl
- [59] Téléchargement de la base de données de visage
<http://cswww.essex.ac.uk/mv/otherprojects.html>
- [60] “FaceVACS-SDK Version 4.0.3”. Document available at :
http://www.cognitec-systems.de/documentation/frsdk_public/index.html