

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

Ecole Nationale Polytechnique

Département d'Electronique

Mémoire du Projet de Fin d'Etudes d'Ingénieur

Thème

**Utilisation de l'analyse en composantes principales et de
l'analyse discriminante linéaire pour la reconnaissance de
visages**

Présenté par :

Mr DZIRI Aziz

Dirigé par :

Pr L. HAMAMI

Melle A. MOUSSAOUI

Soutenu le 18 juin 2011 devant le jury composé de :

Président :	Dr M. ADNANE	ENP
Rapporteur :	Pr L. HAMAMI	ENP
Examineur :	Mr L. ABDELOUEL	ENP

Promotion : juin 2011

REMERCIEMENTS

Ce travail a été effectué dans le laboratoire « Signal & Communications » avec l'équipe 2 « Traitement d'images et reconnaissance de formes » à l'Ecole Nationale Polytechnique d'Alger.

Je tiens tout d'abord à exprimer ma reconnaissance et mes remerciements à ma promotrice, le professeur L. HAMAMI pour ses précieux conseils et ses encouragements le long de ce travail.

Je remercie profondément le docteur M. Adnane d'avoir accepté de présider le jury et de m'avoir procuré de la documentation.

Je remercie également Mr L. Abdelouel, maître-assistant à l'école nationale polytechnique, d'avoir accepté d'examiner ce travail.

Je tiens également à exprimer ma profonde gratitude à tous les enseignants de l'Ecole Nationale Polytechnique qui ont contribué à ma formation.

Je remercie Mr S. MAKHLOUFA et M. N. Nafaâ pour leur aide précieuse.

الملخص:

يندرج هذا العمل في مجال المقاييس الحيوية وخاصة في مجال التعرف على الأشخاص باستعمال الوجوه. ويهدف الى الدراسة والمقارنة بين تقنيتان عامتان وهما التحليل باستعمال المركبات الرئيسية والتحليل بالتمايز الخطي وتستند المقارنة على معدل التحديد الصحيح ، ونسبة الخطأ على قدم المساواة والوقت اللازم لتحديد هوية الشخص ، تم اقتراح عدة خوارزميات مختلفة : التحليل باستعمال المركبات الرئيسية ، وجوه فشر ، التحليل بالتمايز الخطي المباشر ، والتحليل بالتمايز الخطي المباشر بخطوة جزئية. محاكاة هذه الخوارزميات تم باستعمال البرنامج ماتلاب "Matlab".

الكلمات المفتاحية: القياسات الحيوية، التعرف على الوجه، التحليل باستعمال المركبات الرئيسية، التحليل بالتمايز الخطي، تخفيض الأبعاد، التحليل بالتمايز الخطي المباشر، التحليل بالتمايز الخطي المباشر بخطوة جزئية ، وجوه فشر.

Résumé :

Ce travail rentre dans le domaine de la biométrie et plus précisément dans la reconnaissance de visages. Il consiste à étudier et comparer deux techniques globales de reconnaissance de visages qui sont l'analyse en composantes principales (PCA) et l'analyse discriminante linéaire (LDA). La comparaison est basée sur le taux d'identifications correctes, le taux d'égale erreur et le temps de reconnaissance. Pour cela, différents algorithmes sont proposés : PCA, FisherFaces, Direct-LDA et Direct-fractional step-LDA. Les algorithmes sont simulés en utilisant le logiciel Matlab.

Mots-clés: Biométrie, Reconnaissance de visages, Analyse en Composantes Principales (ACP), Analyse Discriminante Linéaire (ADL), réduction de dimension, DLDA, DFLDA, FisherFaces.

Abstract:

This work returns in the field of biometrics and more precisely in the face recognition. It consists to study and compare two global approaches of face recognition which are Principals Components Analysis (PCA) and Linear Discriminant Analysis (LDA). The comparison is based over the rate of correct identifications, equal error rate and the time of recognition. For that, various algorithms are proposed: PCA, FisherFaces, Direct-LDA and Direct-fractional step-LDA. The algorithms are simulated by using the Matlab software.

Keys words: Biometric, Face recognition, Principal Component Analysis (PCA), Linear discriminant analysis (LDA), dimensionality reduction, DLDA, DFLDA, FisherFaces.

SOMMAIRE

INTRODUCTION GENERALE	1
------------------------------------	----------

CHAPITRE I : Introduction aux systèmes biométriques

1. Introduction.....	3
2. Définitions.....	3
3. Les systèmes biométriques	4
3.1. Architecture d'un système biométrique.....	4
3.2. Performances d'un système biométrique.....	7
3.2.1. Les systèmes biométriques en mode authentification	7
3.2.2. Les systèmes biométriques en mode identification.....	9
4. Les modalités biométriques.....	9
4.1. Les modalités morphologiques	10
4.1.1. L'empreinte digitale	10
4.1.2. L'iris	11
4.1.3. La rétine.....	12
4.1.4. Le visage	12
4.1.5. La géométrie de la main	13
4.2. Les modalités comportementales.....	14
4.2.1. La voix.....	14
4.2.2. La signature	15
4.2.3. La dynamique de la frappe	16
4.2.4. La démarche	16
5. Comparaison des performances des différentes modalités	16
6. La multi-modalité.....	18
7. Conclusion	19

SOMMAIRE

CHAPITRE II : Etat de l'art sur la reconnaissance de visages

1. Introduction.....	20
2. Un système de reconnaissance de visages	20
3. Les différentes méthodes de reconnaissance de visages	22
3.1. Les méthodes globales.....	23
3.1.1. Analyse en composantes principales (PCA)	23
3.1.2. Réseaux de neurones	24
3.1.3. l'Analyse Discriminante Linéaire (LDA)	27
3.1.4. Les machines à vecteurs de support (SVM).....	28
3.1.5. Analyse en composantes indépendantes (ICA).....	30
3.2. Les méthodes locales	31
3.2.1. Elastic bunch graph matching (EBGM)	31
3.2.2. Les modèles cachés de Markov (HMM).....	33
3.3. Les méthodes hybrides	34
4. Conclusion	35

CHAPITRE III : Etude de la PCA et de la LDA

1. Introduction.....	36
2. Préparation de la base de données des images de visages	36
3. L'analyse en composantes Principales.....	37
3.1. Le principe de la PCA	38
3.2. Algorithme de la PCA	40
4. L'analyse discriminante linéaire	43
4.1. Principe de la LDA.....	44

SOMMAIRE

4.2.	La LDA pour la reconnaissance de visages	45
4.3.	FisherFaces (PCA+LDA)	46
4.4.	Direct-LDA.....	48
4.5.	Direct Fractional step LDA	51
5.	La Phase de reconnaissance	55
6.	Conclusion	58

CHAPITRE IV :

Simulation et résultats

1.	Introduction.....	59
2.	Présentation des bases de données utilisées	59
2.1.	La base de données ORL	59
2.2.	La base de données Faces94	60
2.3.	La base de données UMIST.....	61
3.	Performances des différents algorithmes.....	62
3.1.	Qualités des algorithmes en terme de séparation de classes.....	62
3.2.	Le taux d'identifications correctes	64
3.3.	Le taux de fausses acceptations (FAR) et le taux de faux rejets (FRR)	70
4.	Conclusion	81

CONCLUSION GENERALE	82
----------------------------------	-----------

BIBLIOGRAPHIE.....	83
---------------------------	-----------

LISTE DES FIGURES

FIG I.1 : Architecture d'un système biométrique	4
FIG I.2 : la phase d'apprentissage dans un système biométrique	5
FIG I.3 : l'identification d'une personne dans un système biométrique	6
FIG I.4 : l'authentification d'une personne dans un système biométrique	7
FIG I.5 : Illustration du FRR et du FAR	8
FIG I.6 : Illustration du EER.....	8
FIG I.7 : La courbe ROC	9
FIG I.8 : Processus de reconnaissance par empreinte digitale	10
FIG I.9 : Les minuties les plus utilisés en reconnaissance par empreinte digitale.....	10
FIG I.10 : Les différents types de minuties.....	11
FIG I.11 :L'iris et le dispositif de capture.....	11
FIG I.12 : Capteur de la rétine	12
FIG I.13 : La rétine	12
FIG I.14 : Processus de reconnaissance de visages	13
FIG I.15 : Dispositifs de capture de la géométrie de la main	13
FIG I.16 : Spectre d'un signal de voix	14
FIG I.17 : La capture d'une signature	15
FIG I.18 : Comparaison des différentes modalités	17
FIG I.19 : Architecture d'un système multimodal pour l'authentification	19
FIG II.1 : Système de reconnaissance automatique de visages.....	21
FIG II.2 : Méthodes de reconnaissance de visages	22
FIG II.3 : Schéma de base d'un neurone.....	24
FIG II.4 : Architecture d'un réseau de neurones multicouches	25
FIG II.5 : Illustration du principe de séparation optimale des classes par la LDA.....	27
FIG II.6 : Principe des SVM	28
FIG II.7 : Problème de classification linéaire et non linéaire	29

LISTE DES FIGURES

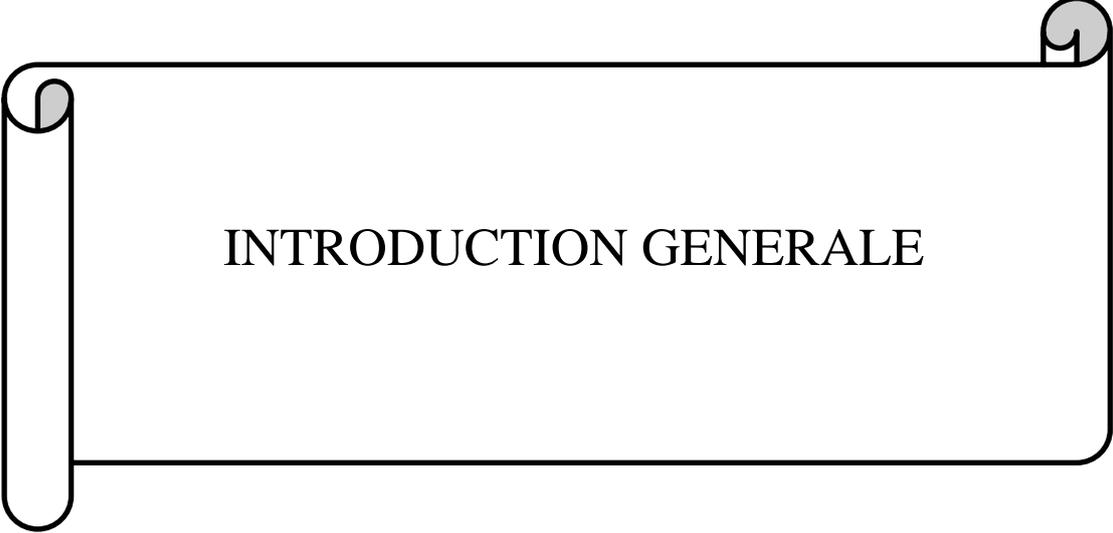
FIG II.8 : Les vecteurs de support	29
FIG II.9 : la représentation des images par l'ICA.....	30
FIG II.10 : Sélection des points caractéristiques et création du treillis.....	32
FIG II.11: Face Bunch Graph (FBG)	33
FIG II.12 : Exemples de modèles HMM	34
FIG III.1 : Le passage d'une image à un vecteur	37
FIG III.2 : Les données à traiter par la PCA	38
FIG III.3 : L'image moyenne.....	40
FIG III.4 : Les cinq premiers visages propres correspondant aux plus grandes valeurs propres	42
FIG III.5 : Schéma de l'algorithme D-LDA	48
FIG IV.1 : Quelques images de la base de données ORL.....	60
FIG IV.2 : Quelques images de la base de données Faces94.....	60
FIG IV.3 : Quelques images de la base de données UMIST	61
FIG IV.4 : Qualité des différents algorithmes pour la séparation de classes	63
FIG IV.5 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données ORL.....	65
FIG IV.6 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données Faces94	65
FIG IV.7 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données UMIST	66
FIG IV.8 : Taux d'identifications correctes en fonction du paramètre p et du nombre de vecteurs pour la DF-LDA sur la base données ORL.....	68
FIG IV.9 : Taux d'identifications correctes en fonction du paramètre r et du nombre de vecteurs pour la DF-LDA sur la base données ORL.....	68
FIG IV.8 : FAR et FRR en fonction du seuil pour 15 vecteurs gardés pour l'ORL	72
FIG IV.9 : Courbe ROC des différents algorithmes pour 15 vecteurs gardés pour l'ORL	73
FIG IV.10 : FAR et FRR en fonction du seuil pour 15 vecteurs gardés pour Faces94	74

LISTE DES FIGURES

FIG IV.11 : Courbe ROC des différents algorithmes pour 15 vecteurs gardés pour Faces94..	75
FIG IV.12 : FAR et FRR en fonction du seuil pour 12 vecteurs gardés pour l'UMIST.....	76
FIG IV.13 : Courbe ROC des différents algorithmes pour 12 vecteurs gardés pour l'UMIST	77
FIG IV.14 : Courbe ROC de la PCA en fonction de nombre de vecteurs gardés pour ORL ...	78
FIG IV.15 : Courbe ROC de la FisherFaces en fonction du nombre de vecteurs gardés pour ORL.....	79
FIG IV.16 : Courbe ROC de la PCA en fonction du nombre de vecteurs gardés pour faces94	79
FIG IV.17 : Courbe ROC de la FisherFaces en fonction du nombre de vecteurs gardés pour faces94.....	80

LISTE DES TABLEAUX

TAB I.1 : Comparatif des caractéristiques des modalités biométriques.....	18
TAB IV.1 : Taux d'identifications correctes en fonction du nombre d'images de l'ensemble d'apprentissage pour la base de données ORL	69
TAB IV.2 : Taux d'identifications correctes en fonction de nombre d'image de l'ensemble d'apprentissage pour la base de données UMIST.....	70
TAB IV.3 : Valeurs d'ERR pour le seuil de décision optimal	72
TAB IV.4 : Valeurs d'ERR pour le seuil de décision optimal pour Faces94.....	74
TAB IV.5 : Valeurs de ERR pour le seuil de décision optimal	76
TAB IV.6 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la PCA.....	78
TAB IV.7 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la Fisherfaces	78
TAB IV.8 : Valeurs du seuil et de ERR en fonction de nombre de vecteurs gardés pour la PCA.....	80
TAB IV.9 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la FisherFaces ...	80



INTRODUCTION GENERALE

Dans un monde où la sécurité est devenue une priorité et où aucune erreur humaine n'est tolérable à cet égard, les ordinateurs et plus particulièrement les logiciels se sont développés dans le but de pouvoir reconnaître automatiquement les personnes. Ce développement a donné naissance à la biométrie qui a remplacé les anciennes méthodes de reconnaissance et de vérification d'individus.

La reconnaissance des visages s'avère une des techniques biométriques qui ont connu un grand développement durant ces dernières années. Bien que la reconnaissance de visages humains soit un domaine difficile à cause de la multitude des paramètres qu'il faut prendre en compte (variation de posture, éclairage, style de coiffure, port de lunettes, de barbes, de moustaches, vieillesse...), il est très important de s'y intéresser vu les nombreux champs d'applications (vérification de personnes, télésurveillance, interfaces homme-machine ...).

Dans ce travail, on exposera les deux techniques les plus utilisées en reconnaissance de visages : l'analyse en composantes principales et l'analyse discriminante linéaire. Les algorithmes de ces deux techniques seront étudiés, simulés et comparés entre eux, tout en variant les différents paramètres influant sur les performances du système : la rapidité et la précision.

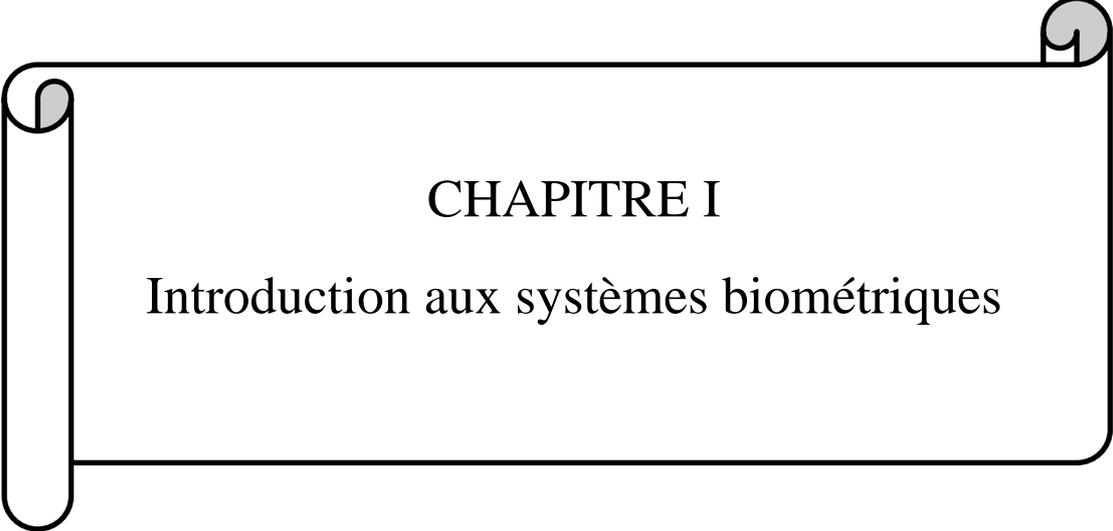
Notre travail est structuré comme suit :

Dans le premier chapitre, on présentera les systèmes biométriques en s'intéressant à leurs architecture et performances. Par la suite on fera une étude comparative des différentes modalités biométriques qui existent.

Dans le second chapitre, un état de l'art sur la reconnaissance de visages sera fourni. On commencera par la présentation des systèmes biométriques de reconnaissance de visages, par la suite on fera une étude des différentes méthodes et algorithmes dédiés à cette reconnaissance.

Ensuite, dans le troisième chapitre, on fera l'étude théorique de l'analyse en composantes principales et l'analyse discriminante linéaire, en exposant les différents algorithmes utilisés et leur principe.

Dans le dernier chapitre, on présentera les différentes bases de données utilisées. On exposera, par la suite, les résultats de la simulation, et on fera une comparaison entre les différents algorithmes utilisés.



CHAPITRE I

Introduction aux systèmes biométriques

1. Introduction :

Etablir l'identité d'une personne est devenue un problème critique dans la société. Des questions comme « est-elle la personne qui prétend être ? », « est-ce que cette personne est autorisée à entrer ? » ou « a-t-il un accès à cette information ? » sont habituellement posées pour différents scénarios. Le besoin de techniques fiables d'authentification d'individus n'a cessé d'augmenter à cause des problèmes liés à la sécurité et les avancements rapides dans les technologies de communication et des réseaux. [3]

Dans ce chapitre on présentera les systèmes biométriques en parlant de leurs architecture et performance. On entamera ensuite les différentes modalités biométriques qui existent, et on finira par la définition de la biométrie multimodale.

2. Définitions:

On entend par biométrie, la reconnaissance des individus en utilisant leurs caractéristiques comportementales (la signature, l'écriture, la démarche,...etc.) et morphologiques (empreinte digitale, visage, iris, contour de la main,...etc.). En se basant sur la présomption que les individus ne peuvent pas avoir les mêmes caractéristiques, la biométrie est utilisée comme un outil pour la reconnaissance (l'identification) ou la vérification (l'authentification) des personnes dans le but de contrôler l'accès à un espace physique, une information, des services ainsi qu'à d'autres droits ou avantages, y compris la capacité de franchir les frontières internationales.

Il existe deux manières traditionnelles qui permettent d'identifier un individu. La première méthode est basée sur une connaissance qui peut être un mot de passe, un code d'identification...etc. La seconde méthode est basée sur une possession. Cette possession peut correspondre à une carte magnétique, un badge, une clef...etc. Ces deux méthodes possèdent des faiblesses et des inconvénients. Dans le premier cas, le mot de passe peut être oublié par son utilisateur ou deviné par une autre personne. Dans le second cas, la carte et le badge peuvent être perdus ou volés. Cependant, quand il s'agit de la biométrie, les caractéristiques physiques et comportementales permettant l'identification de l'individu ne

peuvent pas être perdues et sont très difficiles à deviner ou à trafiquer. Donc la biométrie demeure, pour le moment, le meilleur moyen de sécurité que l'homme peut utiliser. [4]

Pour qu'une caractéristique biométrique puisse être utilisée pour identifier un individu, elle doit satisfaire plusieurs propriétés. Cette caractéristique doit être à la fois [4] :

- Universelle : possédée par tous les individus à identifier.
- Unique : deux individus différents ne peuvent pas avoir la même caractéristique.
- Permanente : elle ne doit pas varier au cours du temps.
- Impossible à dupliquer.
- Facilement et quantitativement mesurable.
- Acceptable : les utilisateurs acceptent d'utiliser la modalité.
- Résistante : elle ne doit pas être influencée par des éléments extérieurs.

3. Les systèmes biométriques :

3.1. Architecture d'un système biométrique :

Un système biométrique est un système de reconnaissance d'individus, il comporte deux modules, un pour l'apprentissage et l'autre pour le test (reconnaissance ou authentification).

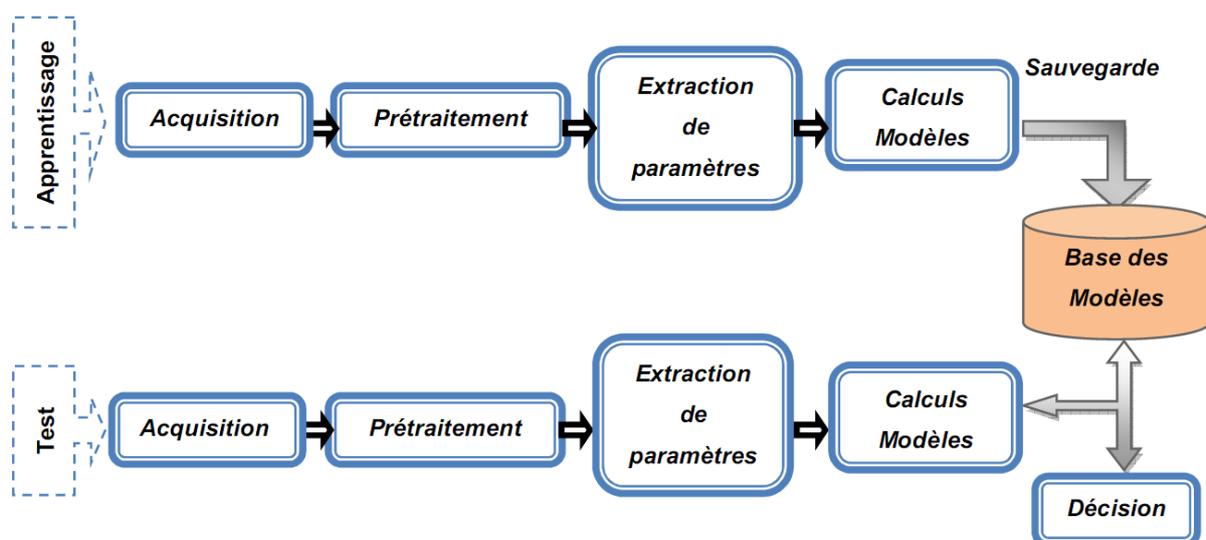


FIG.1 : Architecture d'un système biométrique.

Le module d'apprentissage permet de construire une base de données qui contient une représentation des informations des individus à reconnaître.

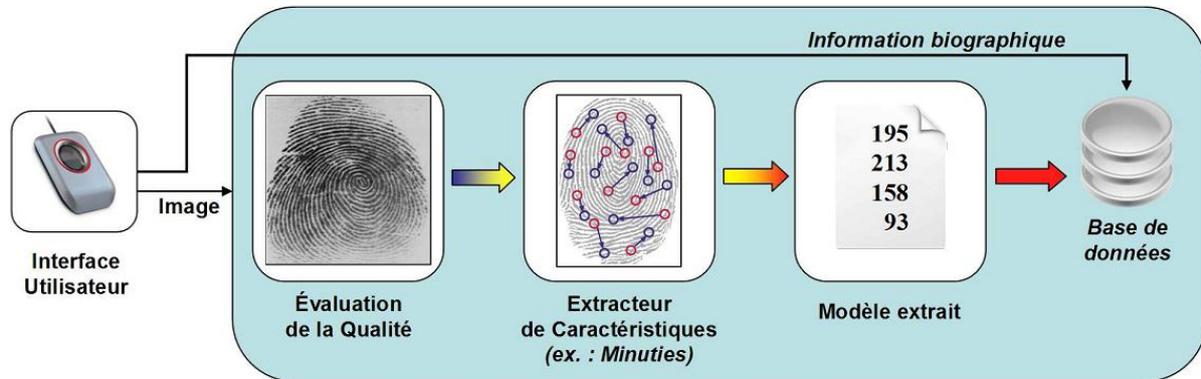


FIG I.2 : la phase d'apprentissage dans un système biométrique. [2]

Le module de test quant à lui permet de comparer les caractéristiques biométriques d'un individu avec ceux appartenant à la base de données puis de prendre une décision.

Les opérations effectuées par les deux modules ne sont pas différentes, à l'exception que l'opération de décision n'existe que dans le module de test. Ces opérations sont les suivantes :

L'acquisition : un capteur est utilisé pour l'acquisition des données biométriques des individus. Ce capteur peut être un appareil photo, un microphone, un lecteur d'empreintes digitales, une caméra de sécurité,...etc.

Le prétraitement : après acquisition des données, ce bloc élimine le bruit dû au capteur et à l'environnement d'acquisition.

L'extraction des paramètres : le rôle de ce bloc est d'extraire l'information utile du signal (extraction du visage, yeux,...etc.) afin de réduire sa taille et le temps de calcul dans la suite.

Le calcul des modèles : les paramètres extraits sont utilisés pour calculer le modèle de l'individu suivant l'approche et la méthode biométrique utilisées. Ces modèles sont stockés

lors de l'apprentissage et vont former la base de données. Lors de test, ces modèles seront comparés à ceux appartenant à la base de données.

La décision : ce bloc est disponible dans le module de test. Son rôle est de comparer le modèle obtenu lors du test à ceux de la base de données et de prendre une décision (l'individu figure ou non dans la base de données dans le cas d'une identification, confirmer ou nier l'identité de l'individu dans le cas d'une authentification).

Selon le module de test, le système biométrique peut servir pour l'identification ou à l'authentification de l'individu [1] :

L'identification (appelée aussi la reconnaissance): le système biométrique a pour rôle d'établir l'identité de l'individu en répondant à la question « qui est la personne X ? » et cela après comparaison de l'information biométrique de cet individu à toutes les informations de la base de données. Ces applications sont utilisées généralement pour identifier les criminels et les terroristes.

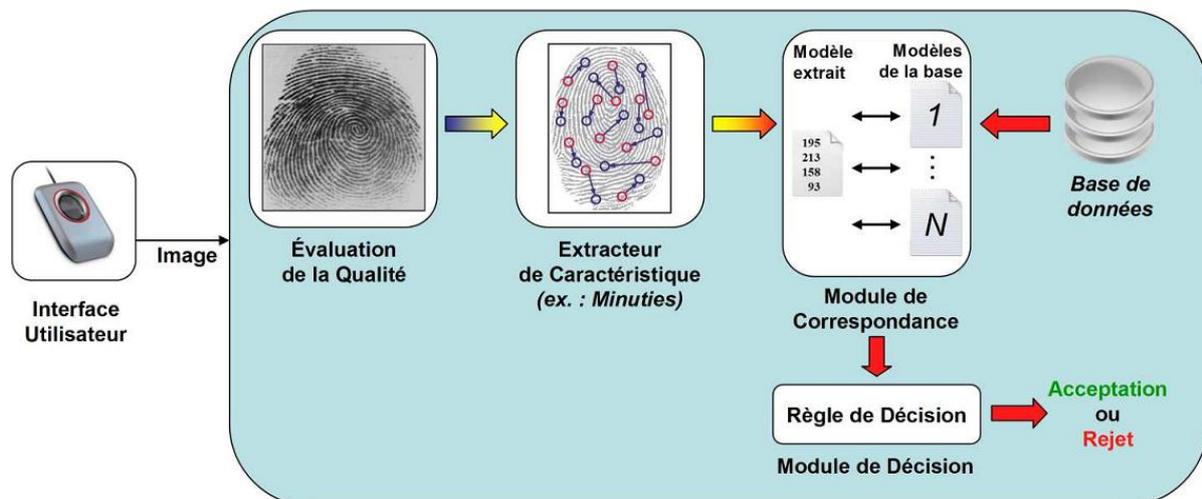


FIG I.3 : l'identification d'une personne dans un système biométrique. [2]

L'authentification (appelée aussi la vérification) : le rôle du système est de confirmer ou nier l'identité réclamée d'un individu en répondant à la question « est-ce la personne X? ». Pour cela, l'individu utilise un mot de passe, un nom d'utilisateur, un code d'identification ou une combinaison de ces trois, pour indiquer son identité au système. Ce dernier demande à

l'individu une information biométrique pour la comparer à l'information enregistrée correspondante à l'identité prétendue. Ces applications sont très utilisées dans le contrôle d'accès et le paiement par authentification.

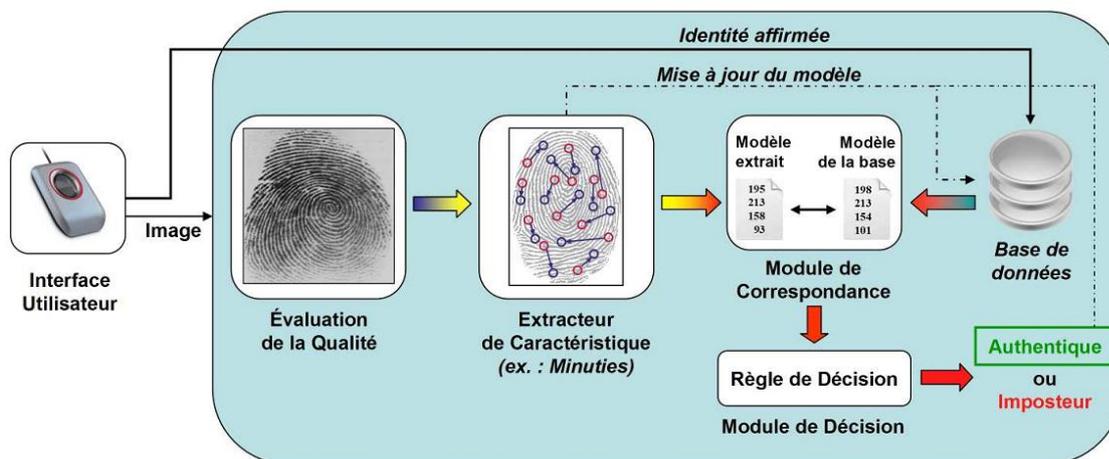


FIG I.4 : l'authentification d'une personne dans un système biométrique. [2]

3.2. Performances d'un système biométrique :

Les performances d'un système biométrique sont déterminées par le coût de sa conception et sa précision. Dans cette étude, on ne va s'intéresser qu'à la précision du système.

3.2.1. Les systèmes biométriques en mode authentification :

Dans ce type de systèmes la précision est mesurée en utilisant les critères suivants [2]:

- **Le taux de faux rejet FRR (False Reject Rate):** il représente le pourcentage des personnes qui sont censées être reconnues par le système (les personnes légitimes ou les clients) mais qui sont rejetées. Ce taux définit en partie le confort d'utilisation du système biométrique. Sa formule mathématique est la suivante :

$$FRR = \frac{\text{Le nombre de clients légitimes rejetés par le système}}{\text{Le nombre total de clients légitimes}} \quad (I.1)$$

- **Le taux de fausses acceptations FAR (False Accept Rate) :** il représente le pourcentage des personnes qui ne sont pas censées être reconnues par le système (appelées les imposteurs) mais qui sont tout de même acceptées. Ce taux définit la sécurité du système biométrique. Sa formule mathématique est la suivante :

$$FAR = \frac{\text{Le nombre des imposteurs acceptés par le système}}{\text{Le nombre total des imposteurs}} \quad (I.2)$$

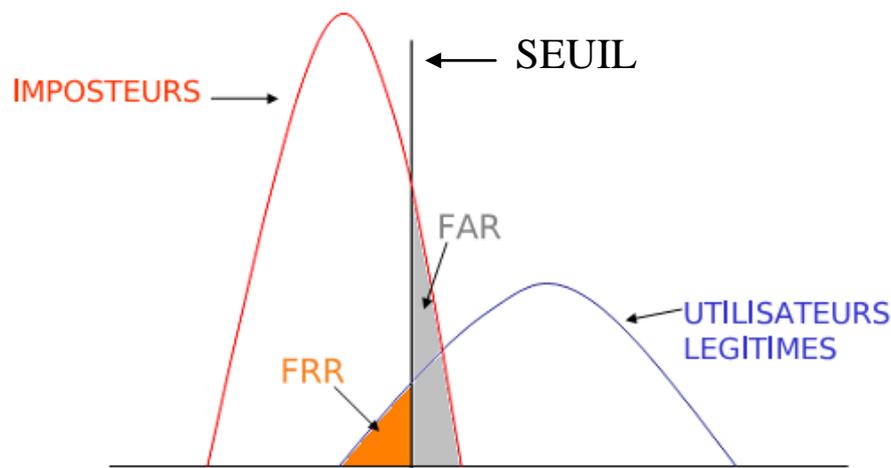


FIG I.5 : Illustration du FRR et du FAR.

- **Le taux d'égale erreur EER (Equal Error Rate) :** ce taux constitue un point de mesure de performance courant. Il est calculé à partir de FRR et FAR et il correspond plus exactement au point d'intersection des courbes FRR et FAR, c'est-à-dire le meilleur compromis entre les faux rejets et les fausses acceptations.

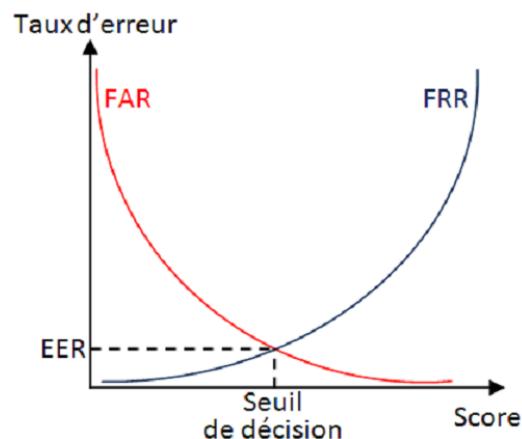


FIG I.6 : Illustration du EER.

On peut représenter graphiquement les performances d'un système biométrique opérant en mode authentification à l'aide de la courbe ROC (Receiver Operating Characteristic). Cette courbe trace le taux de faux rejets en fonction du taux de fausses acceptations pour les différentes valeurs du seuil. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé.

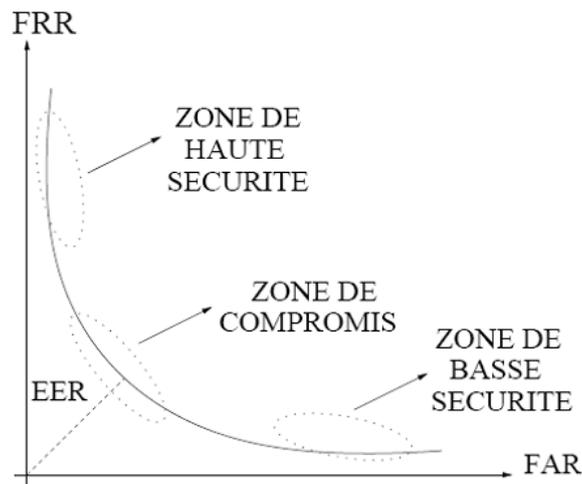


FIG I.7 : La courbe ROC.

3.2.2. Les systèmes biométriques en mode identification :

Dans ce type de systèmes la précision est mesurée à l'aide du taux d'identification correcte (TIC). Ce taux indique le pourcentage des identifications correctes. Sa formule mathématique est la suivante :

$$TIC = \frac{\text{Le nombre des tests correspondant à une identification correcte}}{\text{Le nombre total des tests effectués}} \quad (I.3)$$

4. Les modalités biométriques :

Les modalités biométriques se divisent en deux grandes familles : les modalités morphologiques (l'empreinte digitale, le visage, l'iris...etc.) et les modalités comportementales (la signature, la frappe au clavier...etc.).

4.1. Les modalités morphologiques :

Les modalités morphologiques sont basées sur la mesure des caractéristiques biologique et physique de l'individu. Dans la suite, on va présenter les modalités morphologiques les plus utilisées dans les systèmes biométriques.

4.1.1. L'empreinte digitale :

C'est la plus ancienne modalité utilisée dans les systèmes biométriques. Les empreintes digitales sont différentes d'une personne à une autre (même des jumeaux monozygotes n'ont pas les mêmes empreintes) et pour chaque doigt d'une même personne. La reconnaissance, dans un système à base d'empreinte digitale, se fait à partir des minuties qui représentent les propriétés locales des crêtes (les lignes en contact avec une surface au toucher). Des études ont montré que deux personnes ne peuvent avoir plus de huit minuties en commun, ce qui permet une discrimination entre les personnes. Lors de la reconnaissance, l'image contenant l'empreinte subit un prétraitement afin d'éliminer les bruits pour pouvoir identifier les minuties. Ensuite les minuties et leurs caractéristiques (position, orientation) sont extraites de l'image pour former un modèle de l'empreinte. Ce modèle va être comparé aux modèles de la base de données formés lors de l'apprentissage. [4]

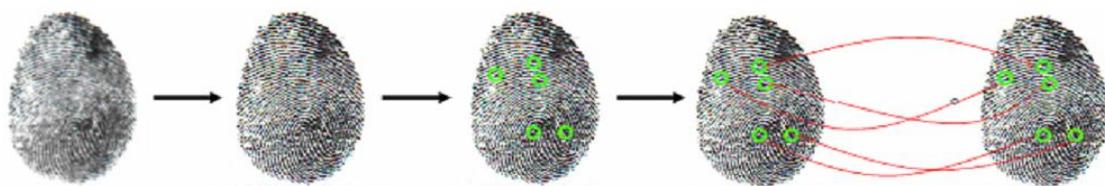


FIG I.8 : Processus de reconnaissance par empreinte digitale. [1]

Les appareils d'acquisition des empreintes digitales sont basés sur la capture optique, thermique, électromagnétique ou sur les ultrasons.

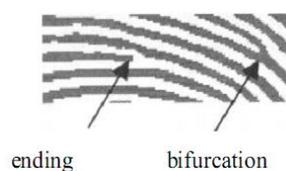


FIG I.9 : Les minuties les plus utilisés en reconnaissance par empreinte digitale.

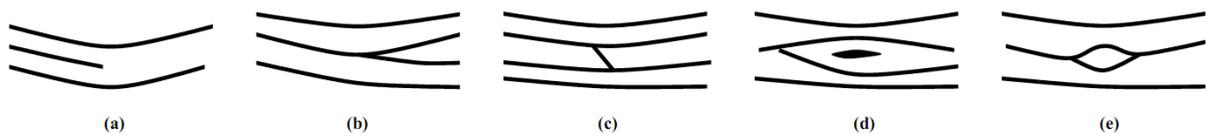


FIG I.10 : Les différents types de minuties :
(a) terminaison, (b) bifurcation, (c) pont, (d) lac et (e) île. [4]

La présence de blessures et de saletés sur les empreintes affecte considérablement la performance du système. Mais la stabilité et l'unicité de l'empreinte rendent son utilisation très vaste. Les systèmes biométriques à base de l'empreinte digitale sont employés dans les banques, les hôpitaux, chez la police et même dans les accessoires tels que les ordinateurs portables et les téléphones mobiles.

4.1.2. L'iris:

On appelle iris la zone colorée située entre la pupille et le blanc de l'œil. Grâce à sa texture détaillée stable et propre à chaque individu, l'iris est la modalité la plus précise pour l'identification et l'authentification. Pour la capture, on utilise une caméra qui doit être très proche de l'œil de l'utilisateur, ce qui rend cette modalité désagréable et mal acceptée par les utilisateurs. A cause de son coût élevé, elle n'est utilisée que dans les endroits où la sécurité est un enjeu critique, comme dans les bases nucléaires et les stations spatiales. Lors de la reconnaissance, l'iris est détecté dans l'image capturée par la caméra, et ses caractéristiques sont extraites et comparées à celles présentes dans la base de données [4] [5].



Iris

FIG I.11 :L'iris et le dispositif de capture.

4.1.3. La rétine :

La rétine est la surface d'à peine 0,5 mm située au fond de l'œil. Cette modalité se base sur le fait que la géométrie des vaisseaux sanguins de la rétine est unique pour chaque personne. Elle permet de distinguer des jumeaux grâce à la possible mesure de 400 caractéristiques distinctes [7]. Cela rend la reconnaissance biométrique par la rétine une méthode très précise. Pour sa capture, un faisceau lumineux est projeté sur l'œil de l'utilisateur qui doit se situer très proche du capteur. Cette méthode d'acquisition nécessite une collaboration étroite de la part de l'utilisateur, et présente des risques sur la santé de l'œil. C'est pourquoi, son utilisation est peu populaire, et se limite aux endroits très sécurisés comme les sites nucléaires militaires et les coffres forts des grandes banques.



FIG I.12 : Capteur de la rétine.

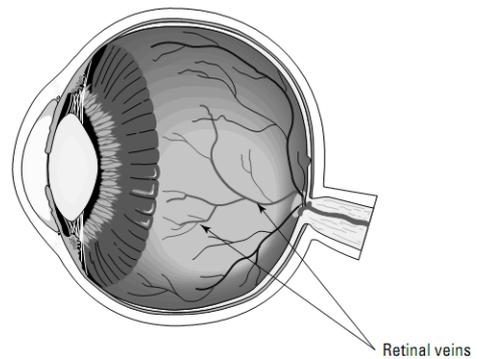


FIG I.13 : La rétine.

4.1.4. Le visage :

Cette modalité est très acceptée par les utilisateurs car le visage est la caractéristique biométrique que les humains utilisent le plus naturellement pour s'identifier entre eux [4]. Lors de la reconnaissance, une photo de visage de l'individu est capturée, volontairement ou involontairement, à l'aide d'une caméra ou un appareil photo. Les caractéristiques de l'individu sont extraites de l'image de visage et comparées à une base de données. Le problème de cette biométrie vient des possibles perturbations pouvant transformer le visage comme le maquillage, la faible luminosité, la présence d'une barbe ou d'une lunette, une expression faciale inhabituelle, le changement avec l'âge [6]...etc.

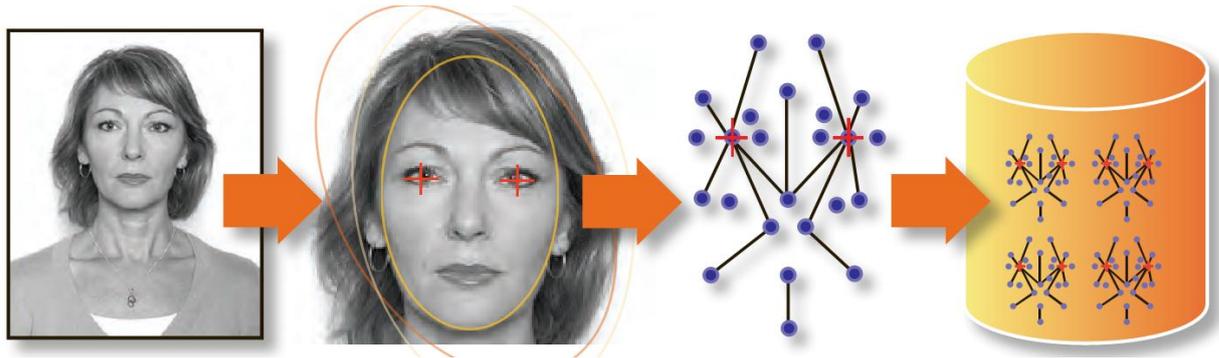


FIG I.14 : Processus de reconnaissance de visages.

Les avantages de cette modalité qui se présente par le faite qu'une personne peut être identifiée sans qu'elle le sache, et le faible coût des équipements qu'elle nécessite, la rend très utilisée dans beaucoup d'applications dont on cite les systèmes de surveillance dans les aéroports et les frontières, ou pour l'identification des personnes dans les laboratoires et les banques.

4.1.5. La géométrie de la main:

Cette modalité se base sur les caractéristiques de la main d'un individu comme sa forme, la longueur et la largeur des doigts, la longueur inter-articulations, les dessins des lignes de la paume...etc. Pour la capture, l'utilisateur doit poser la paume de sa main sur une plaque dont les emplacements des doigts sont indiqués. Une photo du dessus de la main sera prise par un appareil photo. Dans certain cas des captures de profil de la main sont faites dans le but de savoir son épaisseur. Les photos prises seront traitées afin d'extraire les caractéristiques de la main qui vont être comparées à la base de données. [4] [5]

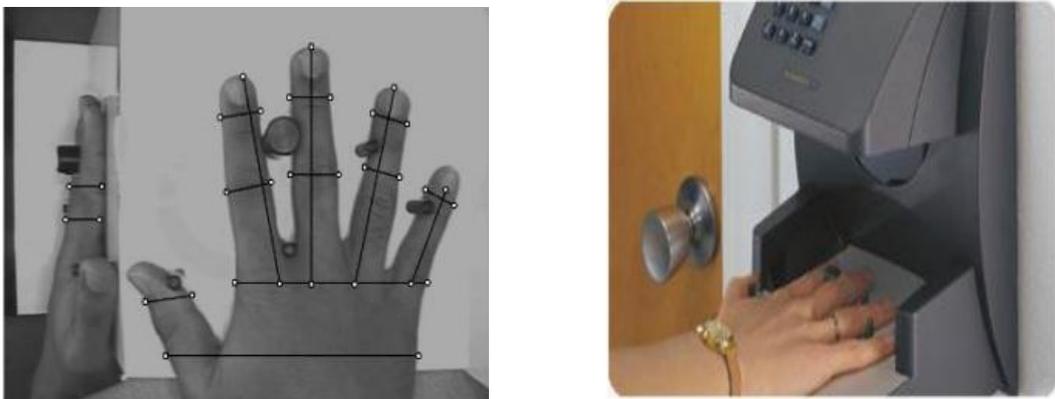


FIG I.15 : Dispositifs de capture de la géométrie de la main.

Cette modalité présente des taux d'erreur assez élevés et ne peut pas distinguer deux personnes de la même famille ou des jumeaux. Ce qui rend son utilisation pour l'identification inefficace. Elle est très bien acceptée par les utilisateurs et généralement employée pour l'authentification dans les contrôles d'accès physique. Cependant elle reste une modalité non stable dans le temps à cause des changements de la forme de la main avec le temps.

4.2. Les modalités comportementales :

Ces modalités reflètent le comportement d'un individu. Ce comportement peut changer au cours du temps. Dans cette section, on va citer quelques modalités comportementales.

4.2.1. La voix :

Plusieurs personnes croient que la voix est une modalité morphologique, hors que les caractéristiques de cette modalité sont une combinaison de facteurs comportementaux (vitesse, rythme...etc.) et physiologiques (tonalité, âge, sexe, fréquence, accent, harmoniques...etc.) [8]. La capture se fait à l'aide d'un microphone, et l'utilisateur peut prononcer, selon le système, une phrase arbitraire de son choix ou précisé par le système lui-même. Alors les caractéristiques physiologiques et comportementales sont extraites et comparées à celle enregistrées dans la base de données.



FIG I.16 : Spectre d'un signal de voix.

La performance d'un système utilisant la voix comme modalité est très sensible à l'état émotionnel de l'utilisateur, comme la joie et le stress, et peut être facilement influencée par une simple maladie comme un rhume. Cependant, cette modalité est très bien acceptée par les gens et sa simplicité d'acquisition la rend très utilisée dans beaucoup d'applications comme la téléphonie, l'espionnage...etc.

4.2.2. La signature:

Cette modalité se divise en deux : la signature statique et la signature dynamique.

Dans la signature statique seule la géométrie de la signature est prise en compte par le système. Par contre dans la signature dynamique, en plus de la caractéristique géométrique, le système prend en compte les caractéristiques dynamiques de l'utilisateur comme la vitesse d'écriture, la pression exercée, l'accélération...etc. Donc la signature dynamique offre plus d'informations que la signature statique. Ce qui rend la signature dynamique plus difficile à dupliquer, car même si l'imposteur peut reproduire la géométrie de la signature, il n'a pas accès à l'information dynamique. Pour la capture, on utilise un stylo électronique et une tablette graphique qui permet de numériser la signature de l'individu.



FIG I.17 : La capture d'une signature.

Le problème de cette modalité provient de la variabilité entre deux occurrences de la signature d'un même individu [4] et l'influence de la performance de système par l'état émotionnel de l'utilisateur comme la fatigue par exemple. Cependant, cette modalité reste avantageuse dans certaines applications telles que les transactions dans le commerce électronique.

4.2.3. La dynamique de la frappe :

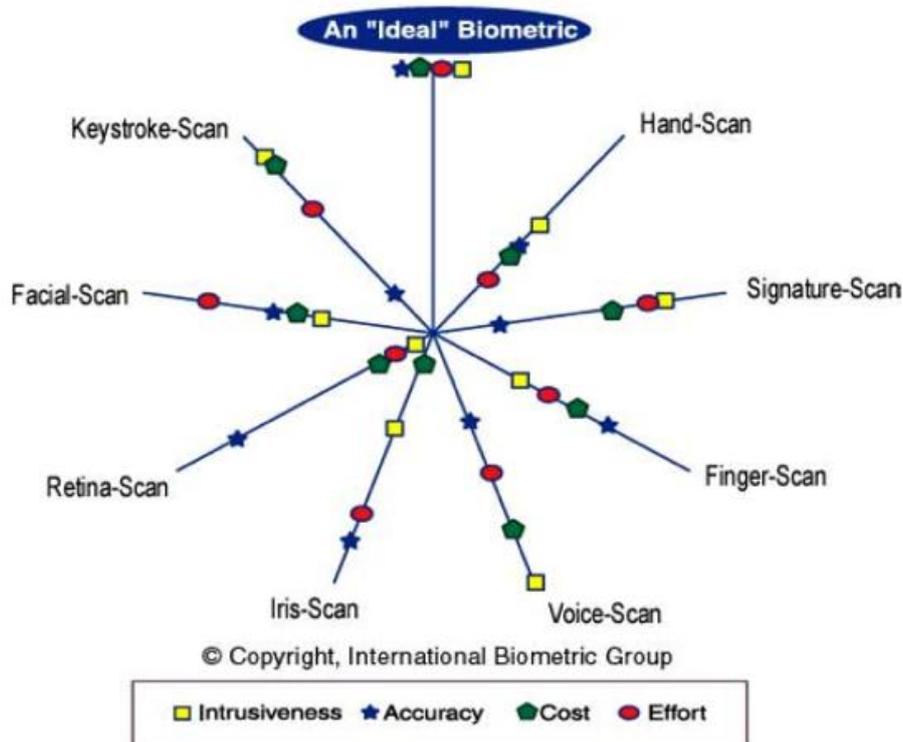
Cette modalité se base sur le fait que la dynamique de la frappe d'une personne sur un clavier est unique. Le système biométrique exploitant cette modalité n'a pas besoin d'un équipement spécial et tout ordinateur équipé d'un clavier pourra faire l'affaire. Il suffit juste d'avoir un logiciel qui permet de calculer le temps entre deux frappes au clavier et la durée pendant laquelle une touche est appuyée par l'utilisateur. La dynamique de la frappe est toujours reliée à un mot de passe. Pendant l'apprentissage, l'utilisateur tape plusieurs fois le mot de passe afin que le système crée un modèle de référence. Cette modalité est généralement utilisée pour les transactions de commerce électronique et pour l'accès aux bases de données.

4.2.4. La démarche :

Chaque personne possède sa propre démarche et c'est cette caractéristique qui est exploitée. A partir d'une séquence vidéo, le système extrait la silhouette de l'individu et analyse ses caractéristiques comme la vitesse, l'accélération...etc. le problème de cette modalité est sa sensibilité aux changements de l'environnement de l'individu comme ses habits, la surface de sol...etc. Ce qui donne des performances médiocres et limite son utilisation. Mais on la trouve, tout de même, employée dans certaines applications de la télésurveillance.

5. Comparaison des performances des différentes modalités :

D'après ce qu'on a vu dans la section précédente, chaque modalité possède des avantages et des inconvénients. Aucune modalité biométrique ne peut répondre aux besoins de toutes les applications. Le choix d'une modalité pour une application donnée dépend d'un certain nombre de paramètres comme le coût et la performance du système, l'acceptabilité de la modalité par les utilisateurs, la facilité de l'utilisation...etc. Une comparaison se basant sur le coût (Cost), l'effort de la conception (Effort), la précision (Accuracy) et l'intrusion (Intrusivness) est indiquée par la figure (FIG I.16).



Source: Zephyr Analysis

FIG I.18 : Comparaison des différentes modalités.

La comparaison fournie par cette figure permet de choisir la modalité la plus adaptée aux contraintes liées à l'application. On voit bien que si la précision est le facteur déterminant de notre application, l'utilisation de l'iris ou la rétine sera nécessaire même si elles sont les plus coûteuses. Quand il s'agit d'une application grand public où la précision n'est pas vraiment importante, les empreintes digitales sont souvent utilisées car elles assurent un très bon compromis entre les quatre critères de comparaison. Pour ce qui est du visage et de la géométrie de la main, la figure montre que le coût et la précision sont moyens, mais les systèmes se basant sur la géométrie de la main nécessitent une mise en œuvre rigoureuse contrairement à ceux utilisant le visage. La reconnaissance se basant sur la voix est très résistante aux intrusions et ne nécessite pas beaucoup d'effort, mais sa précision est faible.

Un autre comparatif des caractéristiques de toutes les modalités étudiées précédemment est indiqué par le tableau I.1. Les critères de comparaison sont cette fois-ci l'universalité, l'unicité, la permanence...etc.

Modalité	universalité	unicité	permanence	performance	acceptabilité	non reproductibilité
Empreinte digitale	moyen	élevé	élevé	élevé	moyen	élevé
Iris	élevé	élevé	élevé	élevé	faible	élevé
Rétine	élevé	élevé	moyen	élevé	faible	élevé
Visage	élevé	faible	moyen	faible	élevé	faible
Géométrie de la main	moyen	moyen	moyen	moyen	moyen	moyen
Voix	moyen	faible	faible	faible	élevé	faible
Signature	faible	faible	faible	faible	élevé	faible
Dynamique de frappe	faible	faible	faible	faible	moyen	moyen
Démarche	moyen	faible	faible	faible	élevé	moyen

TAB I.1 : Comparatif des caractéristiques des modalités biométriques.

6. La multi-modalité :

Les systèmes se basant sur une seule modalité présentent plusieurs problèmes, parmi lesquels on trouve le bruit engendré par le capteur lors de l'acquisition. Ce problème cause une dégradation des performances du système. Un autre inconvénient d'un système monomodal est la modalité en elle-même qui peut présenter des variabilités intra-classes (les données collectées lors de l'apprentissage ne correspondent pas aux données de la reconnaissance à cause, par exemple, d'un changement de la luminosité pour une reconnaissance de visage.).

Afin d'augmenter la fiabilité des systèmes biométriques et améliorer leurs performances, la multi-modalité est souvent employée. La multi-modalité signifie la combinaison de deux (ou plus) systèmes biométriques différents afin de former un seul système. Les deux systèmes combinés peuvent:

- Utiliser des modalités différentes (par exemple la voix et le visage, l'iris et l'empreinte digitale...etc.).
- Utiliser la même modalité, mais les algorithmes sont différents pour chaque système.
- Utiliser des capteurs d'acquisition différents.

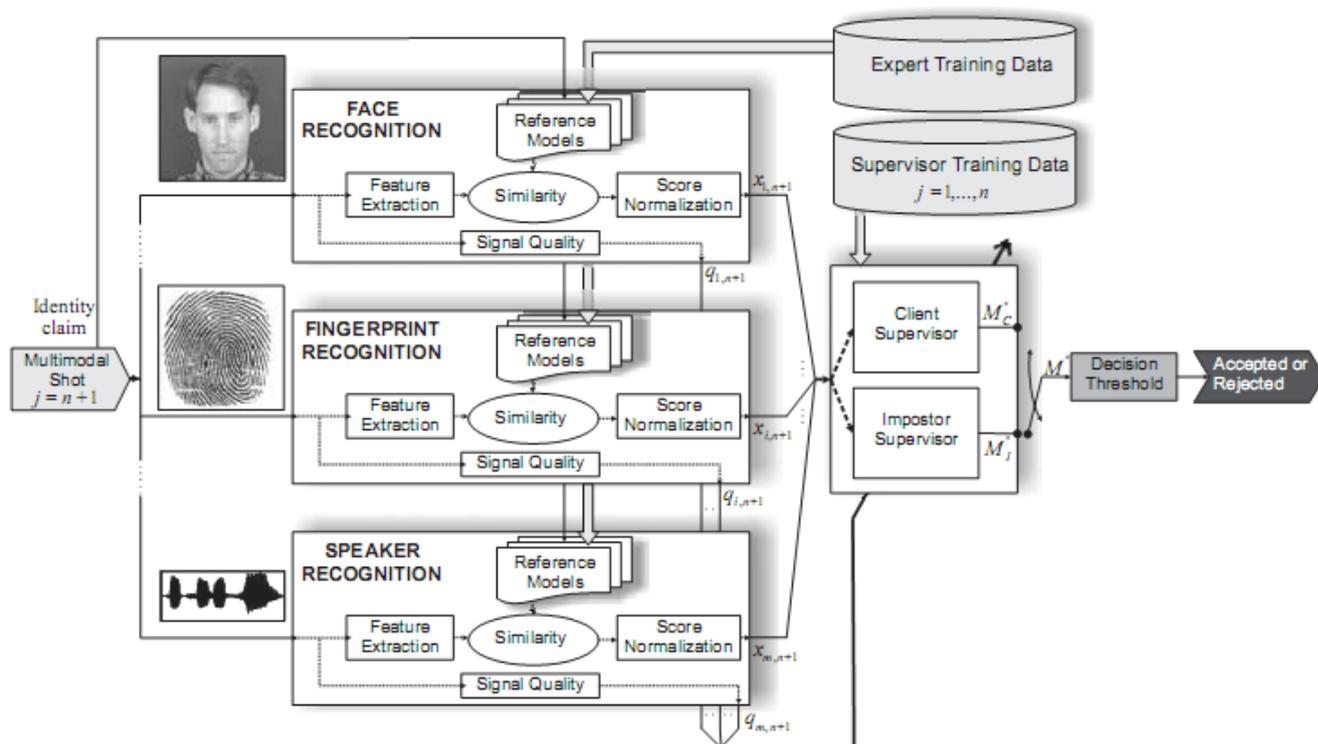
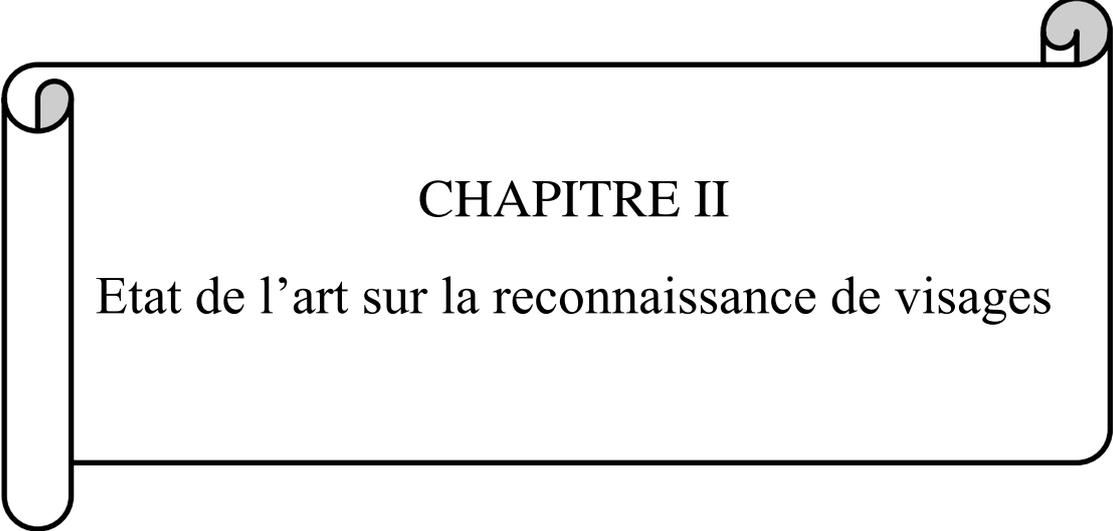


FIG I.19 : Architecture d'un système multimodal pour l'authentification.

7. Conclusion :

Dans ce chapitre, on a présenté une vue générale sur les systèmes biométriques et leur fonctionnement. Les modalités biométriques citées ont chacune leurs avantages et inconvénients en terme de précision, simplicité, coût et intrusion. Afin d'augmenter la fiabilité du système biométrique et d'assurer de bonnes performances, on fait appel à la multimodalité.

Parmi les modalités étudiées, le visage reste la plus naturelle pour reconnaître des personnes, et grâce à ses avantages dont le faible coût, non intrusive, acceptée par le public et assurant un niveau de sécurité satisfaisant, elle occupe une place importante dans le marché de la biométrie. Pour toutes ces raisons, on s'est intéressé dans notre travail à la reconnaissance du visage et on présentera dans ce qui suit quelques algorithmes de cette modalité.



CHAPITRE II

Etat de l'art sur la reconnaissance de visages

1. Introduction :

Les techniques biométriques sont de plus en plus utilisées dans différentes applications. En particulier, la reconnaissance de visages qui a vu un développement considérable ces dernières années. Grâce à sa facilité d'utilisation, son acceptabilité par les utilisateurs et son faible coût, cette technique est classée deuxième des plus utilisée après l'empreinte digitale. La recherche dans le domaine de reconnaissance de visages n'a cessé d'augmenter et grâce aux travaux de plusieurs chercheurs, les machines peuvent maintenant différencier un visage humain d'un autre et en particulier de chercher un visage dans une large base de visages [10].

La reconnaissance de visages est une tâche effectuée quotidiennement par les humains afin de s'identifier et se reconnaître entre eux. Cette caractéristique rend la reconnaissance automatique de visages très bien acceptée par les utilisateurs. Cette technique de reconnaissance trouve son utilité dans plusieurs applications de traitement numérique des images et des vidéos. La reconnaissance de visages peut se faire dans un environnement contrôlé où un certain nombre de contraintes sont imposées lors de l'acquisition (la luminosité, l'angle de capture, l'arrière-plan...etc.). Comme elle peut être effectuée dans un environnement non contrôlé ce qui nécessitera un prétraitement de l'image capturée.

Dans ce chapitre, on présentera, en détail, la reconnaissance de visages, en commençant par une description du système de reconnaissance automatique de visages et les différentes étapes de son processus, puis, on finira par une description des méthodes et des algorithmes utilisés dans la reconnaissance de visages.

2. Un système de reconnaissance de visages :

L'architecture d'un système de reconnaissance automatique de visages comprend, plus ou moins, les mêmes modules constituant un système biométrique. On va s'intéresser, dans la suite de cette section, à la reconnaissance de visages dans un environnement non contrôlé.

Un système de reconnaissance de visages a pour rôle d'identifier le visage présent dans une image ou une vidéo d'une manière automatique. Il est généralement constitué de

quatre blocs comme le montre la figure II.1 : la détection, l'alignement, l'extraction de caractéristiques et la comparaison. La détection et l'alignement constituent les étapes de prétraitement effectuées avant la reconnaissance.

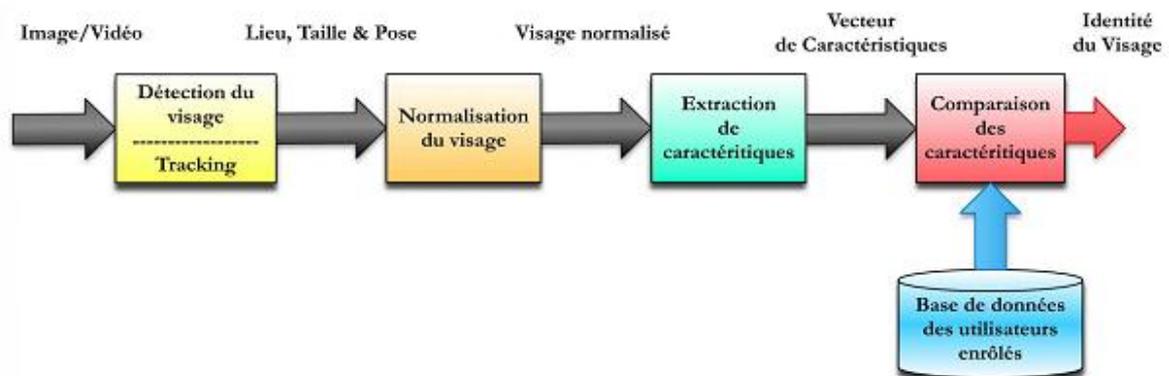


FIG II.1 : Système de reconnaissance automatique de visages.

L'acquisition : l'acquisition se fait à l'aide d'un appareil photo dans le cas d'une image et d'une caméra dans le cas d'une vidéo. Le système d'acquisition a pour rôle de représenter en 2D un objet qui est en 3D. L'image capturée sera numérisée, codée, et représentée par une matrice. Des prétraitements qui ont pour but d'éliminer le bruit sont effectués sur l'image.

La détection de visage : cette étape permet d'extraire le visage à partir de l'arrière-plan de l'image acquise. Dans le cas d'une vidéo, le visage détecté peut avoir besoin d'être suivi, pour ça, on utilise la composante « Tracking ». Dans la détection, des estimations non précises sur la position et la taille du visage sont fournies.

Normalisation du visage : Ce bloc a pour but de parvenir à une localisation plus précise et à normaliser le visage détecté. Les composantes de visages comme la bouche, le nez, les yeux sont localisées. En se basant sur ces localisations, l'image de visage est normalisée par rapport à ses propriétés géométriques telles que la taille et la pose, et par rapport à ses propriétés photométriques telles que l'illumination et le niveau de gris [11].

L'extraction des caractéristiques : une fois le visage normalisé, l'extraction de caractéristiques sera effectuée afin d'extraire les informations utiles nécessaires pour distinguer entre les visages des différentes personnes [11]. Lors de l'apprentissage ces

caractéristiques sont utilisées afin de former les modèles qui seront stockés dans la base de données.

Comparaison des caractéristiques : Pendant la phase de reconnaissance, les caractéristiques extraites précédemment seront comparées à celles de la base de données, et une décision sera prise par ce bloc. Pour l'authentification, la décision consiste à valider ou à refuser l'identité de la personne. Pour ce qui est de l'identification, la décision consiste à sélectionner l'individu dont le modèle, présent dans la base de données, est celui qui correspond le mieux aux caractéristiques extraites du visage.

3. Les différentes méthodes de reconnaissance de visages :

Ces dernières années, plusieurs techniques de reconnaissance de visage (voir la figure II.2) se sont développées, et elles se divisent, en général, en deux grandes familles : les méthodes globales qui s'intéressent à l'intégralité du visage et les méthodes locales qui exploitent les éléments du visage comme le nez, la bouche, les yeux...etc. Les techniques qui combinent ces deux méthodes sont appelées "méthodes hybrides".

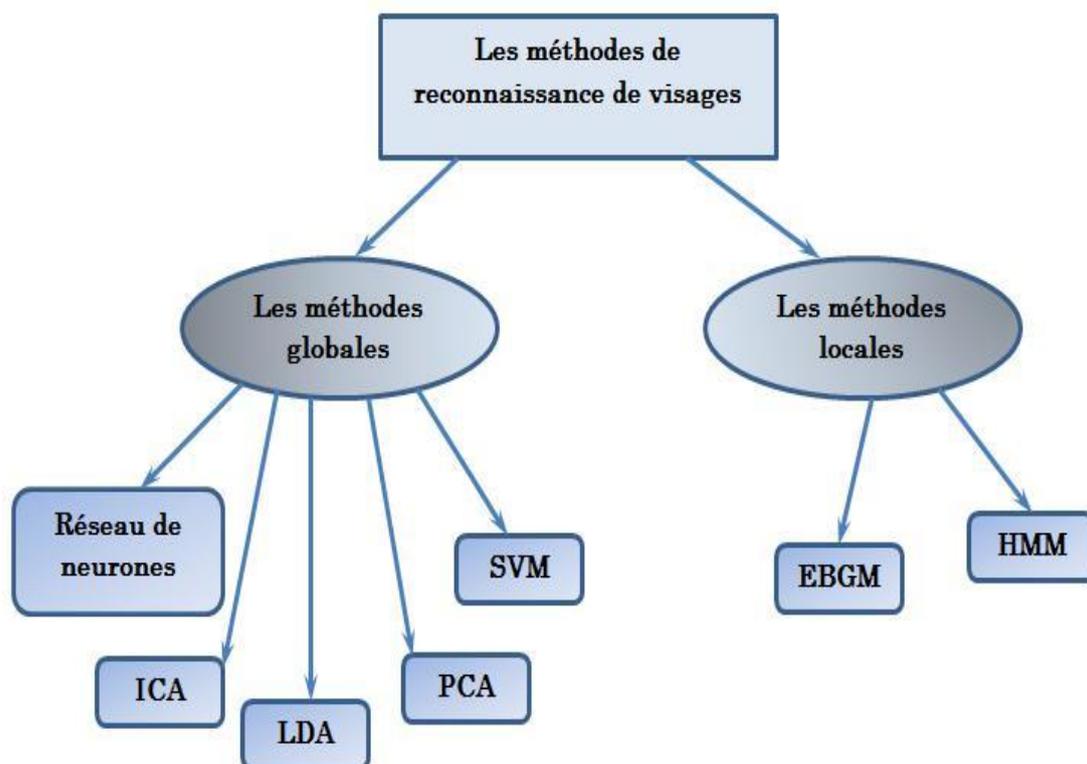


FIG II.2 : Méthodes de reconnaissance de visages.

3.1. Les méthodes globales :

Ces méthodes traitent le visage dans son intégralité sans tenir compte de caractéristiques locales telles que le nez, la bouche, les yeux...etc. Elles sont généralement basées sur des techniques statistiques bien connues. La matrice constituant l'image est injectée à l'entrée du système de reconnaissance et sera convertie en un ensemble de vecteurs, afin de faciliter les traitements. Ces méthodes sont rapides à mettre en œuvre et les calculs de base sont de complexité moyenne. En revanche, elles sont très sensibles aux variations d'éclairément, de pose et d'expression faciale [2]. On donne ci-après quelques techniques de reconnaissance de visages appartenant à cette famille :

3.1.1. Analyse en composantes principales (PCA):

Cette méthode a été utilisée, pour la première fois, dans la reconnaissance automatique des visages en 1991 par MA. Turk et AP. Pentland au MIT Media Lab. Elle est aussi connue sous le nom Eigenfaces. Dans cette approche, chaque image de l'ensemble d'apprentissage est représentée par un ensemble de coefficients (poids) arrangés sous forme de vecteurs. Les poids sont les coefficients de projection de l'image de visage sur l'espace engendré par les images propres, d'où le nom "Eigenfaces" [12]. Pour la phase de l'apprentissage, chaque vecteur contenant les poids représentant un visage est stocké dans la base de données.

Lors de la reconnaissance, l'image de test est représentée, elle aussi, par son vecteur de poids. L'identification consiste à trouver, dans la base de données, l'image dont les poids sont les plus proches, en terme de distance (par exemple la distance euclidienne), des poids de l'image de test.

Les images propres sont les composantes principales de la distribution des visages constituant l'ensemble d'apprentissage. Elles sont obtenues en calculant les vecteurs propres de la matrice de covariance formée par toutes les images de visage de l'ensemble d'apprentissage. Le nombre total des images propres est égal au nombre des images de l'ensemble d'apprentissage. Néanmoins, seulement les images propres les plus significatives (ayant les plus grandes valeurs propres) sont prises en compte, les autres sont négligées.

Cette méthode utilise le niveau de gris des pixels de l'image et elle est facile à mettre en œuvre, néanmoins, elle est très sensible aux variations d'éclairément, de pose et d'expression faciale. Le principe selon lequel on peut construire un sous-espace vectoriel en ne retenant que les "meilleurs" vecteurs propres, tout en conservant beaucoup d'information utile, fait de la PCA un algorithme efficace et couramment utilisé en réduction de dimensionnalité [2].

3.1.2. Réseaux de neurones :

Un réseau de neurones artificiel est un modèle de calcul dont la conception est schématiquement inspirée du fonctionnement des neurones biologiques. Le neurone est l'unité de traitement de base dans un réseau de neurones artificiel. Un neurone artificiel est composé des éléments suivants : les nœuds d'entrée, les connections synaptiques dont les forces sont représentées par des poids et la fonction d'activation φ qui relie toutes les entrées à la sortie de neurone [14].

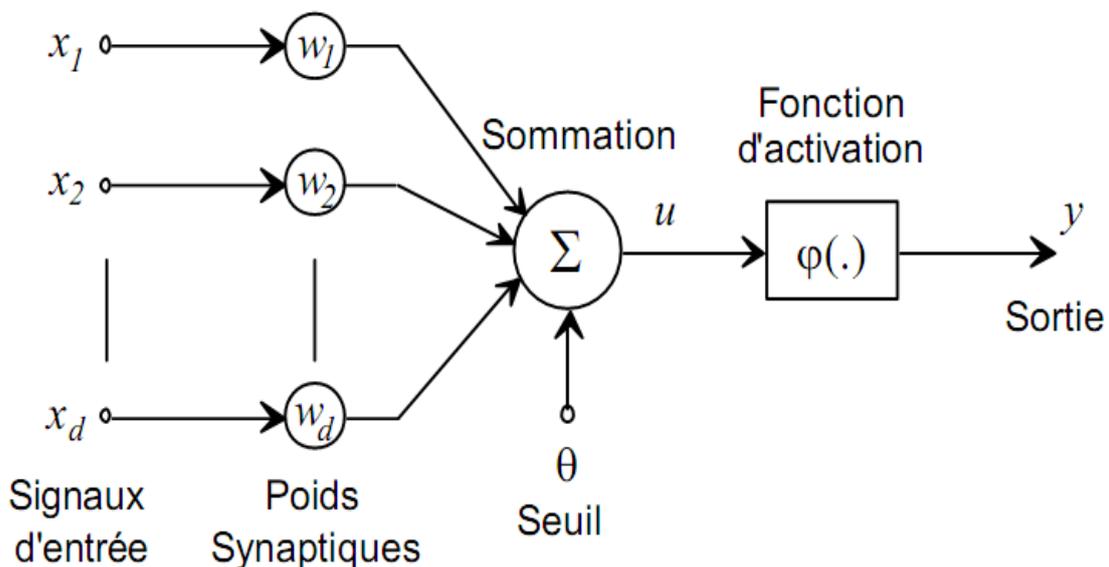


FIG II.3 : Schéma de base d'un neurone.

Chaque neurone est une fonction non linéaire paramétrée. Dans le réseau de neurones, les neurones sont fortement connectés entre eux et travaillent en parallèle. La fonction de transfert d'un neurone est la suivante :

$$y = \varphi(u) \quad (\text{II.1})$$

Avec

$$u = \sum_{i=1}^d x_i w_i - \theta = XW - \theta \quad (\text{II. 2})$$

$X = [x_1, x_2, x_3, \dots, x_d]$: Le vecteur d'entrée

$W = [w_1, w_2, w_3, \dots, w_d]$: Le vecteur de poids

La fonction d'activation φ peut être :

- La fonction signe :

$$\varphi(u) = \begin{cases} 1 & \text{si } u > 0 \\ -1 & \text{si } u \leq 0 \end{cases} \quad (\text{II. 3})$$

- La fonction sigmoïde de paramètre λ :

$$\varphi(u) = \frac{1}{1 + e^{-\lambda u}} \quad (\text{II. 4})$$

Il existe aussi d'autres fonctions pouvant être utilisées.

Plusieurs types de réseaux de neurones existent. Selon la connectivité, on a un réseau total (tous les neurones sont connectés entre eux) et un réseau par couche (les neurones d'une couche ne sont connectés qu'aux couches suivantes). Dans les réseaux par couches on peut trouver des réseaux monocouche et des réseaux multicouches. Les réseaux multicouches sont en général les plus utilisés pour les applications de la reconnaissance de visages. Voici quelques réseaux célèbres: Le Perceptron, le perceptron multicouches (PMC), Le réseau de Kohonen, Le réseaux de Hopfield...etc.

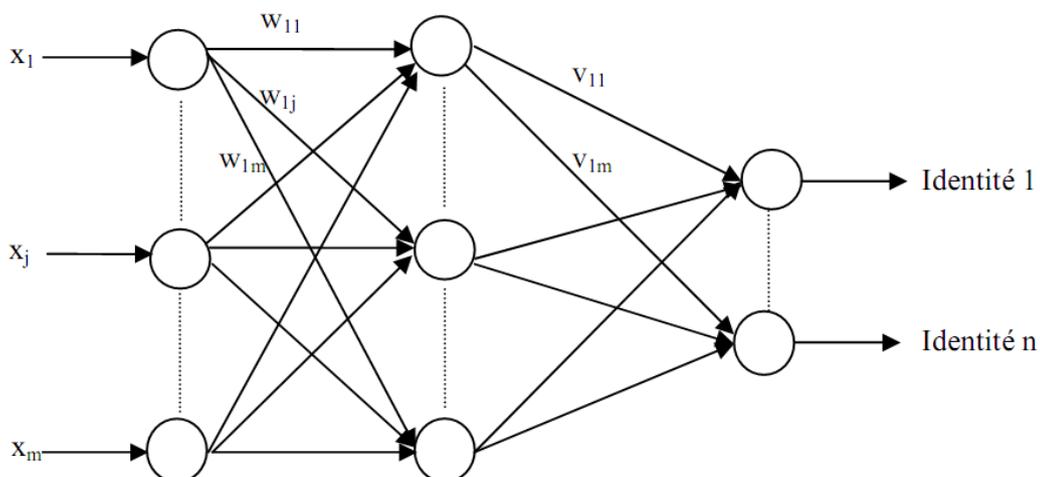


FIG II.4 : Architecture d'un réseau de neurones multicouches.

Pour la reconnaissance de visages, le nombre d'entrées de la première couche du réseau dépend directement du nombre de pixels de l'image de visages. Plus le nombre de pixels augmente, plus la complexité de calcul augmente. Pour ce qui est des sorties du réseau, leur nombre dépend directement du nombre de personnes à reconnaître présentes dans la base de données.

Dans la reconnaissance de visages avec les réseaux de neurones chaque image de visage est convertie en un vecteur colonne qui sera envoyé à l'entrée du réseau. Les étapes de la reconnaissance sont les suivantes :

La phase d'apprentissage : dans cette phase, le réseau de neurones utilise les images de l'ensemble d'apprentissage comme des entrées afin de calculer les poids des différents neurones. L'apprentissage peut se faire de deux façons différentes :

Apprentissage supervisé [15]: dans ce type d'apprentissage, la réponse désirée pour chaque entrée doit être connue. Les entrées X et la réponse désirée Y^* sont fournies au réseau. A partir de l'entrée X , le réseau calcule la sortie Y et la compare à la sortie désirée Y^* . Selon l'erreur obtenue, les poids de chaque neurone de réseau sont ajustés. La procédure est répétée jusqu'à ce que l'erreur entre la sortie obtenue Y et la sortie désirée Y^* soit suffisamment faible.

Apprentissage non supervisé [16]: dans ce type d'apprentissage, la sortie désirée n'est pas connue. En utilisant les entrées X , le réseau évolue jusqu'à la stabilité. Plusieurs algorithmes d'apprentissage existent. On peut citer : Propagation, Rétro-propagation, La méthode du gradient, Rétro-propagation du gradient...etc.

La phase de reconnaissance : après que l'apprentissage soit terminé, on pourra procéder à la reconnaissance. Dans cette phase, l'image à reconnaître est envoyée à l'entrée du réseau. La sortie Y du réseau est un vecteur qui contient K éléments (le nombre de personnes à reconnaître présentes dans la base de données). Le $i^{\text{ème}}$ élément du vecteur correspond à la probabilité P_i que le visage soumis au réseau appartienne à la $i^{\text{ème}}$ personne de la base de données. La personne dont la probabilité est élevée correspond à l'image du visage à reconnaître.

3.1.3. l'Analyse Discriminante Linéaire (LDA) :

L'algorithme LDA est né des travaux de Belhumeur et al. de la Yale Université (USA), en 1997. Il est aussi connu sous le nom de Fisherfaces. Contrairement à la PCA qui ne tient pas compte des différences entre les classes, l'algorithme LDA assure une véritable séparation des classes. Pour l'utiliser, il faut donc au préalable organiser l'ensemble des images d'apprentissage en plusieurs classes. Chaque personne est représentée par une classe qui contient toutes ses images [2].

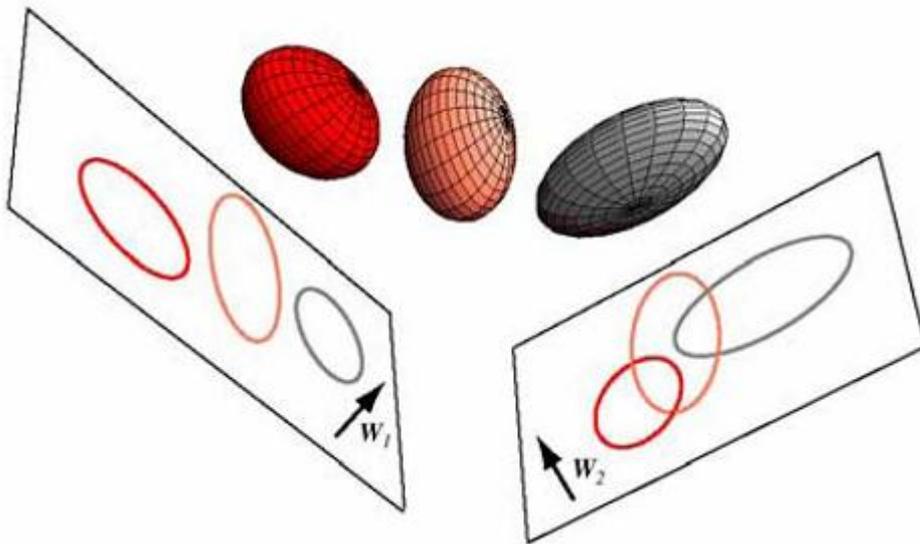


FIG II.5 : Illustration du principe de séparation optimale des classes par la LDA [2].

Pour utiliser cet algorithme dans la reconnaissance faciale, les images de l'ensemble d'apprentissage sont converties en vecteurs afin de faciliter leur manipulation. A partir de ces vecteurs, on calcule l'image moyenne (le vecteur moyen) de chaque classe et l'image moyenne de toutes les images de l'ensemble d'apprentissage. Les images de chaque classe sont centrées par rapport à la moyenne de la classe et les matrices de dispersion intra-classe (la dispersion entre les images appartenant à une classe et l'image moyenne de cette classe) et de dispersion inter-classe (la dispersion entre l'image moyenne de chaque classe et l'image moyenne de l'ensemble d'apprentissage) sont calculées. Par la suite, ces matrices seront utilisées pour trouver la base qui assure la projection la plus optimale de façon à maximiser la dispersion inter-classe tout en minimisant la dispersion intra-classe. Chaque classe sera alors représentée par les coefficients de sa projection sur la base trouvée.

L'image de test, convertie en vecteur, est projetée sur la base trouvée lors de l'apprentissage. La distance entre les coefficients de projection de l'image test et ceux des images de l'ensemble d'apprentissage est calculée. L'image test appartiendra alors à la classe dont la distance calculée est minimale par rapport aux autres classes.

Cet algorithme est très utilisé dans la reconnaissance de visages car il assure une véritable séparation entre les classes, en minimisant les variations entre les images d'un même individu (variations intra-classe) tout en maximisant les variations entre les images d'individus différents (variations inter-classes) [2].

3.1.4. Les machines à vecteurs de support (SVM) :

Les SVM ont été développées dans les années 1990 à partir des considérations théoriques de Vladimir Vapnik sur le développement d'une théorie statistique de l'apprentissage : la Théorie de Vapnik-Chervonenkis. Elles ont eu beaucoup de succès dans quasiment tous les domaines où elles ont été appliquées, grâce à leur fondement théorique et leurs bons résultats en pratique [17]. Les SVM sont, généralement, utilisées dans la classification ou reconnaissance de formes, mais elles permettent également de traiter les problèmes de régression non linéaire.

Les SVM font appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elles sont basées sur l'utilisation de fonctions qui permettent une séparation optimale de données. Ces fonctions sont dites noyau (kernel). La fonction noyau Φ permet de transformer un problème de classification non linéaire dans l'espace de représentation (l'espace d'entrée), en un problème de classification linéaire dans un espace de dimension plus grande appelé espace de re-description.

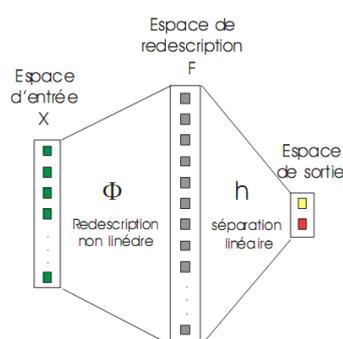


FIG II.6 : Principe des SVM.

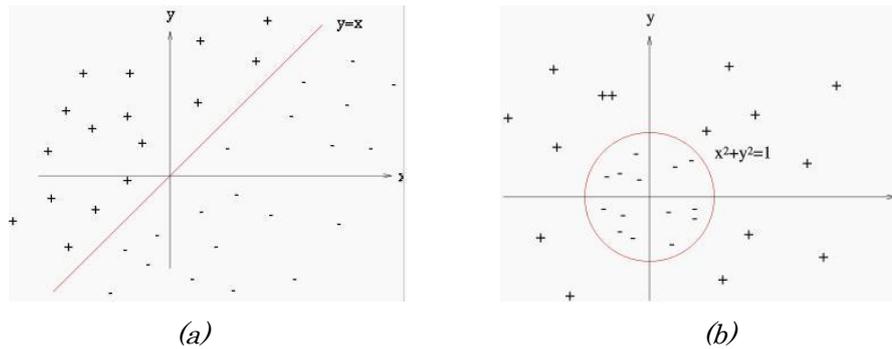


FIG II.7 : (a) problème de classification linéaire.

(b) problème de classification non linéaire.

Le but des SVM est de trouver un classificateur optimal permettant de séparer deux classes de données en maximisant la distance entre elles (la marge). Ce classificateur est un classificateur linéaire appelé hyperplan. Ce classificateur linéaire optimal peut être trouvé pour un problème de classification linéaire ou non linéaire. Les points les plus proches des deux classes sont les seuls utilisés pour déterminer l'hyperplan et ils sont appelés vecteurs de support.

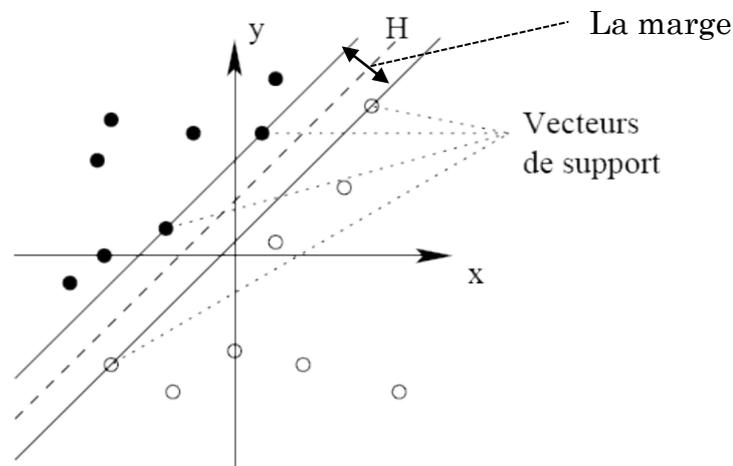


FIG II.8 : Les vecteurs de support.

Dans la reconnaissance de visages, les données d'apprentissage sont les images de visages constituant l'ensemble d'apprentissage et les classes à séparer avec l'hyperplan sont les individus. L'apprentissage consiste à trouver l'hyperplan en configurant autant de SVM que de classes (individus), en étiquetant +1 les images de la classe d'un individu et -1 toutes les images des autres individus [18].

3.1.5. Analyse en composantes indépendantes (ICA):

L'ICA a été utilisée pour la première fois dans la reconnaissance de visages par Bartlett et al. en 2002. Cette méthode consiste à rechercher une transformation linéaire qui permet d'obtenir des données sous forme de composantes statistiquement indépendantes. L'ICA est une extension de l'analyse en composantes principales (PCA) par le recours aux statistiques d'ordre supérieur à deux [19]. L'ICA est un outil d'analyse de données dérivée des techniques de traitement des signaux et plus exactement de séparation de sources. Elle a pour but de déterminer les signaux originaux à partir d'un ensemble d'observations, où chaque observation est un mélange des signaux originaux supposés statistiquement indépendants [20].

Dans la reconnaissance de visages, les images de l'ensemble d'apprentissage sont considérées comme des combinaisons linéaires d'images inconnues statistiquement indépendantes. Lors de l'apprentissage, l'algorithme ICA permet de déterminer la base de projection (les images inconnues statistiquement indépendantes) optimale et les coefficients de projection (poids) pour chaque image de l'ensemble d'apprentissage. Chaque image sera alors représentée par ses coefficients de projection.

Pour trouver un ensemble d'images composantes indépendantes, l'image moyenne est soustraite de toutes les images de l'ensemble d'apprentissage, puis les statistiques d'ordre 1 et 2 des données sont éliminées par la matrice de blanchement $W_z = 2 \times \sqrt{\text{cov}(X)}$. L'algorithme d'ICA permet alors de trouver une matrice W telle que $U = (W W_z)X$ où U est la matrice qui contient les images constituant la base de projection et X est la matrice qui contient les images de l'ensemble d'apprentissage avec une moyenne nulle. La matrice W est estimée de façon à ce que les images de la base de projection soient statistiquement indépendantes [21]. Une image s'écrit, alors, comme une combinaison linéaire de ces composantes indépendantes en utilisant la formule $X = (W W_z)^{-1}U$.

$$\text{Image} = b_1^* \mathbf{u}_1 + b_2^* \mathbf{u}_2 + \dots + b_n^* \mathbf{u}_n$$

ICA representation = (b_1, b_2, \dots, b_n)

FIG II.9 : la représentation des images par l'ICA.

Le nombre des composantes indépendantes trouvées par l'ICA est égal au nombre des images de l'ensemble d'apprentissage. Si on veut réduire ce nombre, on doit appliquer l'ICA sur un nombre de combinaisons linéaires des images de l'ensemble d'apprentissage. Ces combinaisons linéaires peuvent être trouvées par l'analyse des composantes principales (PCA), en choisissant les premiers visages propres de l'ensemble d'apprentissage.

Pour la reconnaissance, le même principe que celui de la PCA est appliqué. L'image de test est projetée à l'aide de la matrice $(W W_z)$, et à partir des coefficients de projection, on calcule la distance euclidienne entre cette image et les images de la base de données.

3.2. Les méthodes locales :

Les méthodes locales sont aussi appelées "méthodes géométriques". Elles consistent à appliquer des transformations en des endroits spécifiques de l'image, le plus souvent autour de points caractéristiques (coins des yeux, de la bouche, le nez...etc.). [23]

Les méthodes globales présentent un inconvénient qui est l'insuffisance de l'information extraite représentant le visage. En effet, dans ces méthodes, le maximum d'énergie est concentré pour représenter adéquatement l'ensemble d'images, tout en négligeant certains détails qui peuvent être déterminants dans la discrimination entre deux personnes possédant une physionomie faciale très ressemblante. Les méthodes globales sont aussi trop sensibles aux variations d'éclairément et de pose.

Les méthodes locales sont plus robustes aux problèmes posés par les variations d'éclairément, de pose et d'expression faciale. Elles accordent plus d'énergie aux détails locaux. Toutefois, elles sont plus lourdes à mettre en place [22] [23].

Dans cette section, on présentera les méthodes locales les plus utilisées dans la reconnaissance de visages.

3.2.1. Elastic bunch graph matching (EBGM) :

L'algorithme EBGM est né des travaux de Wiskott et al. de la Southern California University (USC - USA) et de la Ruhr University (Allemagne), en 1997 [22].

Cette méthode consiste à représenter le visage par un graphe. Pour cela, une localisation des points caractéristiques du visage, tels que les coins des yeux, du nez, de la bouche...etc. est nécessaire. Cette localisation peut se faire d'une manière manuelle ou automatique à l'aide d'un algorithme. Une fois les points caractéristiques localisés, un treillis élastique virtuel est appliqué sur l'image de visage à partir de ces points et un graphe sera obtenu.

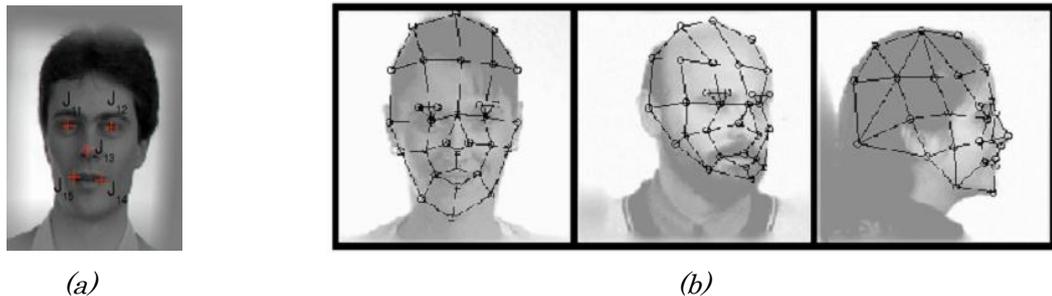


FIG II.10 : (a). Sélection des points caractéristiques.
(b). Création du treillis.

Chaque nœud du treillis est un point caractéristique représenté par un jeu de coefficients complexes d'ondelettes de Gabor, appelé Jet. Chaque arête du treillis reliant deux points caractéristiques est représentée par la distance entre ces deux points.

Les différents Jets sont calculés par la convolution de l'image, autour des points caractéristiques, avec plusieurs ondelettes de Gabor paramétrées. Un Jet peut être écrit comme un ensemble de coefficients complexes.

Au début de la phase d'apprentissage, la localisation des points caractéristiques de visage se fait manuellement. Les graphes obtenus seront classés selon leur pose (frontale, profil,...etc.). Une fois que le nombre de graphes par pose est jugé suffisant (environ 30~70), ils seront regroupés pour former un seul graphe, appelé Face Bunch Graph (FBG). Le FBG de chaque pose sera utilisé pour la localisation automatique des points caractéristiques des images de visages ayant la même pose. Tous les FBG et les graphes représentant les visages de l'ensemble d'apprentissage seront stockés pour former la base de données.

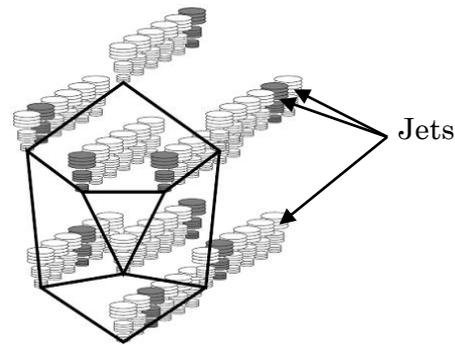


FIG II.11: Face Bunch Graph (FBG).

Les nœuds d'un FBG sont représentés par un jeu de Jets, et ses arêtes par les distances moyennes des différents graphes de visages constituant le FBG. Dans la figure II.11 chaque empilement représente un Jet. A partir d'un jeu de Jets relié à un nœud du treillis, seulement celui qui correspond le mieux est sélectionné pour localiser les points caractéristiques de l'image de visage (indiqué en gris à titre d'exemple). [22][23][24]

Durant la phase de reconnaissance, les points caractéristiques de l'image test sont trouvés à l'aide de FBG ayant la même pose, le treillis est mis en place et les Jets sont calculés. Le graphe de l'image, ainsi obtenu, sera comparé aux différents graphes présents dans la base de données. Une fonction de ressemblance sera utilisée pour déterminer quel graphe de la base de données correspond le mieux, et par ailleurs l'individu sera reconnu.

Cet algorithme local ne traite pas directement les valeurs de niveaux de gris des pixels d'une image de visage, ce qui lui confère une plus grande robustesse aux changements d'éclairage, de pose et d'expression faciale. Cependant il est plus difficile à implémenter que les méthodes globales vues précédemment et le temps de calcul s'en retrouve augmenté.

3.2.2. Les modèles cachés de Markov (HMM):

Les HMM ont été utilisés pour la première fois dans la reconnaissance de visages par Samaria & Young 1994 à l'université de Cambridge (Angleterre). Cette méthode consiste à fabriquer pour chaque individu un modèle statistique (modèle de Markov).

Un modèle caché de Markov est un automate à temps discret se trouvant à chaque instant dans un état. Le nombre d'états du modèle est fini et la transition entre les états se

produit entre deux instants discrets consécutifs. Les probabilités de transition d'un état à un autre sont données par la matrice de transition A . Lorsque l'automate passe de l'état q_{t-1} (l'état occupé par le modèle à l'instant $t - 1$) à l'état q_t il émet une observation o_t avec une probabilité donnée par la matrice d'émission B . L'état initial du modèle est donné par le vecteur π qui contient la densité de probabilité de l'état initial.

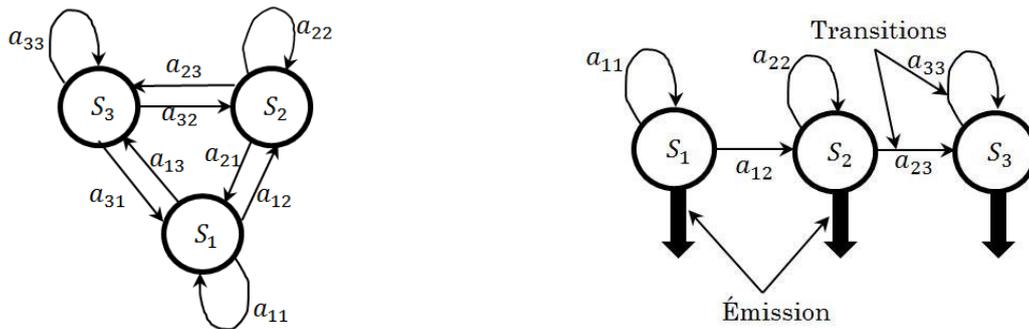


FIG II.12 : Exemples de modèles HMM.

Durant la phase d'apprentissage, des séquences d'observations sont générées à partir des images de chaque individu appartenant à l'ensemble d'apprentissage. Ces séquences d'observations sont utilisées pour construire les modèles qui seront stockés afin de former la base de données.

Lors de la reconnaissance, la séquence d'observation O générée à partir de l'image de test est utilisée afin de calculer la vraisemblance $P(O/\lambda)$ (λ est le modèle HMM) pour chaque modèle de la base de données. Le modèle possédant la plus grande vraisemblance correspond à l'individu à reconnaître.

Cette méthode assure de très bons résultats et est moins sensible aux variations d'éclairage et de pose. Cependant, elle est difficile à mettre en œuvre et nécessite des calculs complexes ce qui implique des temps de reconnaissance élevés. [25] [26]

3.3. Les méthodes hybrides :

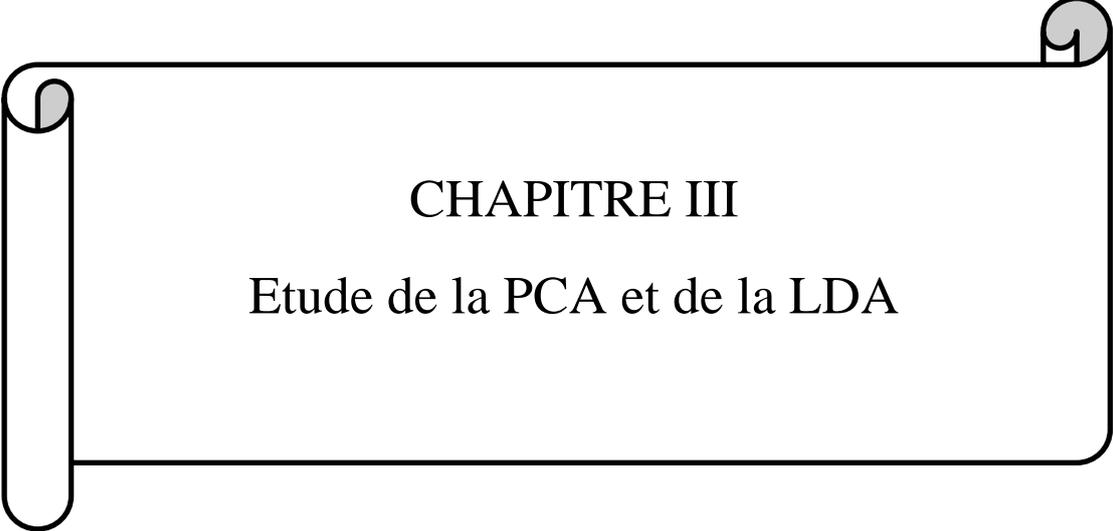
Les méthodes hybrides sont des combinaisons des deux méthodes vues précédemment. Elles permettent d'offrir au système plus de stabilité et de performance. Parmi ces méthodes,

on peut citer : L'analyse des caractéristiques locales (Local features analysis) et l'analyse en composantes principales modulaire (PCA modulaire) [28].

4. Conclusion :

Dans ce chapitre, on a fait le tour des méthodes utilisées pour la reconnaissance de visages. Vue la diversité des algorithmes qui existent, on ne s'est intéressé qu'aux méthodes qui ont apporté une vraie originalité et une avancée sensible dans le domaine de la reconnaissance automatique de visages.

Les méthodes globales sont les plus utilisées à cause de leur facilité de mise en œuvre et leur rapidité. Cependant, elles sont très sensibles aux variations d'éclairément et de pose. Les méthodes locales, quant à elles, sont très robustes et donnent de bons résultats, leurs défauts résident dans les temps de reconnaissance élevés et la complexité de conception.



CHAPITRE III

Etude de la PCA et de la LDA

1. Introduction :

Les approches statistiques globales s'avèrent les plus utilisées dans la reconnaissance de visages à cause de leur facilité d'implémentation. Dans ce chapitre on détaillera deux approches qui sont l'analyse en composantes principales (PCA) et l'analyse discriminante linéaire (LDA).

On commencera notre chapitre par la préparation de la base de données des images de visages, puis, on décrira et on expliquera le principe de la PCA et celui de la LDA ainsi que leurs algorithmes. Afin de faciliter la compréhension, on verra en premier lieu, les phases d'apprentissage de ces deux approches et en fin de chapitre on exposera la phase de test qui est la même pour les deux.

2. Préparation de la base de données des images de visages :

Dans la reconnaissance des visages, la base de données des images est divisée en deux ensembles :

- a. **L'ensemble d'apprentissage** : cet ensemble contient les images de visages qui vont être utilisées dans la phase d'apprentissage afin d'extraire leurs caractéristiques et de former la base de données des modèles qui servira lors de la reconnaissance.
- b. **L'ensemble de test** : cet ensemble, quant à lui, contient les images qui vont être utilisées dans la phase de test. Les images de cet ensemble sont injectées à l'entrée du système et comparées aux modèles formés lors de l'apprentissage, afin qu'une décision soit prise.

Supposons que notre ensemble d'apprentissage contient N images de résolution $L \times M$ et que le nombre d'individus dans cet ensemble soit égal à C . Chaque individu i de l'ensemble possède n_i images et représente une classe.

Avant de commencer la phase de l'apprentissage, les images de l'ensemble d'apprentissage doivent être converties de matrices en vecteurs afin de pouvoir les utiliser.

Soit I_j la matrice représentant une image de l'ensemble d'apprentissage ; la conversion se fait par concaténation et un vecteur de dimension n sera obtenu comme suit :

$$I_j = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1M} \\ x_{21} & x_{22} & \dots & x_{2M} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ x_{L1} & x_{L2} & \dots & x_{LM} \end{bmatrix} \quad \longrightarrow \quad X_j = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ \vdots \\ x_{LM} \end{bmatrix} = \begin{bmatrix} p_1^j \\ p_2^j \\ \vdots \\ \vdots \\ p_n^j \end{bmatrix}$$

FIG III.1 : Le passage d'une image à un vecteur.

Les coefficients p_i représentent les valeurs des pixels en niveaux de gris, codés de 0 à 255.

Une fois que toutes les images de l'ensemble d'apprentissage soient converties en vecteurs, elles seront rassemblées dans une seule matrice X de taille $n \times N$.

$$X = \begin{matrix} & X_1 & X_2 & \dots & X_j & \dots & X_N \\ \begin{bmatrix} p_1^1 & p_1^2 & \dots & p_1^j & \dots & p_1^N \\ p_2^1 & p_2^2 & \dots & p_2^j & \dots & p_2^N \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ p_n^1 & p_n^2 & \dots & p_n^j & \dots & p_n^N \end{bmatrix} \end{matrix}$$

↓
j^{ème} Image

Cette matrice X sera utilisée dans la suite de ce chapitre.

3. L'analyse en composantes Principales :

L'analyse en composantes principales (PCA) est l'une des méthodes de projection en sous-espaces des plus utilisées en reconnaissance de visages. Son but principal est de trouver la meilleure base permettant de dé-corréler des variables initialement corrélées, en utilisant les statistiques d'ordre deux (la moyenne et la covariance). Les composantes ou les vecteurs constituant cette nouvelle base sont appelés composantes principales.

Dans la reconnaissance de visages, la PCA est utilisée pour extraire les caractéristiques, dans le but de réduire la dimension des images de visages et d'accentuer leurs similitudes et différences. Ces caractéristiques extraites ne correspondent pas nécessairement aux caractéristiques faciales telles que les yeux, le nez ou les oreilles.

3.1. Le principe de la PCA :

La PCA s'applique sur un ensemble de données, représenté généralement par un tableau contenant n lignes et p colonnes. Les colonnes du tableau représentent les variables et les lignes représentent les individus. Par exemple dans la reconnaissance de visage, les colonnes représentent les images et les lignes les niveaux de gris (pixels).

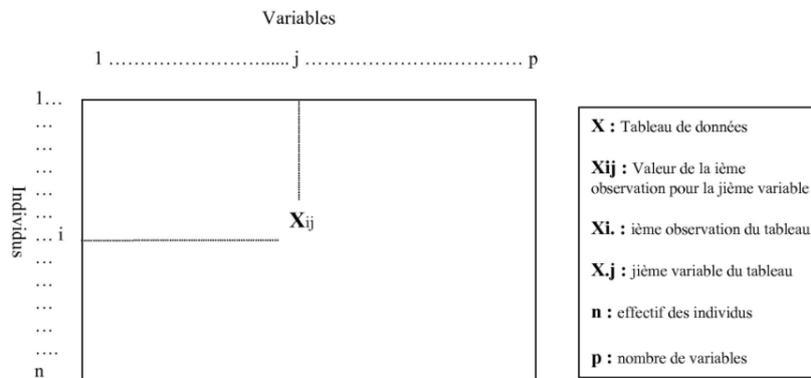


FIG III.2 : Les données à traiter par la PCA.

L'idée à la base de l'analyse en composantes principales est de pouvoir expliquer ou rendre compte de la variance observée dans la masse de données initiales en se limitant à un nombre réduit de composantes, appelées composantes principales, définies comme étant des transformations mathématiques pures et simples des variables initiales [31]. Cette réduction provoque généralement une perte d'information. Le rôle de la PCA est de procéder de façon à ce que cette perte d'information soit la plus faible possible. Pour cela, l'algorithme de PCA doit obéir à certaines contraintes.

Tout d'abord, la première composante extraite par la PCA doit correspondre à la plus grande proportion possible de la variance présente dans les variables initiales. Cette composante est une combinaison linéaire des variables initiales, et elle s'écrit comme suit :

$$C_1 = a_1X_1 + a_2X_2 + \dots + a_pX_p = \sum_{j=1}^p a_jX_j \tag{III. 1}$$

C_1 : la première composante.

a_j : les coefficients de pondération ou les poids de C_1 .

X_j : les variables initiales.

La variance restante, inexpliquée par la première composante, subit, elle aussi le même processus. Ce qui permet d'extraire une deuxième composante qui explique la plus grande proportion de variance possible parmi la variance laissée inexpliquée par la première composante. Cette deuxième composante qu'on appellera C_2 doit être indépendante de la première et elle peut s'écrire elle aussi comme suit :

$$C_2 = b_1X_1 + b_2X_2 + \dots + b_pX_p = \sum_{j=1}^p b_jX_j \quad (III.2)$$

b_j : les coefficients de pondération ou les poids de C_2 .

D'autres composantes seront extraites suivant le même principe. La totalité de la variance présente dans les variables initiales sera expliquée par un nombre de composantes égal au nombre de variables initiales. Chacune des composantes extraites explique une proportion de variance de moins en moins importante, la composante C_1 explique plus de variance que C_2 , C_2 plus que C_3 et ainsi de suite.

Comme le rôle de la PCA est de réduire la masse de données, alors seulement les composantes qui expliquent les plus grandes proportions de variance seront prises en compte. Les autres peuvent être négligées sans une grande perte de l'information.

En théorie de l'information, utilisé la PCA pour la reconnaissance de visages, revient à extraire l'information significative présente dans l'image de visage et l'encoder de la manière la plus efficace possible afin de la comparer à la base de données contenant les visages encodés de la même façon. En terme mathématique, cela revient à trouver les vecteurs propres de la matrice de covariance formée par les images de visages de l'ensemble d'apprentissage. Les vecteurs propres les plus significatifs formeront la base de projection. [13]

3.2. Algorithme de la PCA :

Dans cette section, on décrira l'algorithme PCA appliqué à la reconnaissance de visages. On s'intéressera seulement à la phase d'apprentissage.

A partir de la matrice X calculée précédemment, on calcule l'image moyenne de l'ensemble d'apprentissage qui peut être vue comme le centre de gravité de l'ensemble d'images.

$$I_{moy} = \frac{1}{N} \sum_{j=1}^N X_j \quad (III.3)$$



FIG III.3 : L'image moyenne.

L'image moyenne I_{moy} sera soustraite de chaque image de l'ensemble d'apprentissage et la matrice X devient Φ tel que :

$$\Phi_i = X_i - I_{moy} \quad (III.4)$$

$$\Phi = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_i \quad \dots \quad \Phi_N] \quad (III.5)$$

Par la suite, on calcule la matrice de covariance Cov de l'ensemble des images. Cette matrice est donnée par la formule suivante :

$$Cov = \sum_{i=1}^N \Phi_i \Phi_i^T = \Phi \Phi^T \quad (III.6)$$

La prochaine étape consiste à calculer les vecteurs et les valeurs propres de cette matrice. Etant donnée la grande taille de cette matrice $n \times n$ (n : la résolution ou le nombre de pixels de l'image), ce calcul s'avère difficile à effectuer et prend beaucoup de temps. Mais comme $N \ll n$ (le nombre d'images dans l'ensemble d'apprentissage est très inférieur à la

résolution de l'image), la matrice Cov aura seulement $N - 1$ valeurs propres non nulles. De ce fait, on peut passer d'une complexité de calcul de l'ordre du nombre de pixels dans une image à une complexité de l'ordre du nombre d'images par le procédé suivant :

Posons $A = \Phi^T \Phi$, qui est une matrice de taille $N \times N$, et soient e_i les vecteurs propres de cette matrice associés aux valeurs propres λ_i . Alors, on aura :

$$A e_i = \lambda_i e_i \quad (\text{III. 7})$$

Ce qui est équivalent à :

$$\Phi^T \Phi e_i = \lambda_i e_i \quad (\text{III. 8})$$

En multipliant chaque côté par Φ , on aura :

$$\Phi \Phi^T \Phi e_i = \Phi \lambda_i e_i \quad (\text{III. 9})$$

Ce qui est équivalent à :

$$\text{Cov } \Phi e_i = \lambda_i \Phi e_i \quad (\text{III. 10})$$

D'un autre côté, soient V_i les vecteurs propres de la matrice de covariance C associés aux valeurs propres μ_i . Cela permet d'écrire :

$$\text{Cov } V_i = \mu_i V_i \quad (\text{III. 11})$$

Par identification entre (III. 10) et (III. 11), on aura :

$$\bullet \quad V_i = \Phi e_i = \sum_{j=1}^N e_{ji} \Phi_j \quad \text{avec} \quad e_i = \begin{bmatrix} e_{1i} \\ e_{2i} \\ \vdots \\ e_{Ni} \end{bmatrix} \quad (\text{III. 12})$$

$$\bullet \quad \mu_i = \lambda_i \quad (\text{III. 13})$$

Comme on le voit, ce procédé permet de trouver les valeurs et les vecteurs propres de l'énorme matrice Cov en trouvant ceux de la petite matrice A . Par exemple, si l'ensemble d'apprentissage contient 100 images de résolution 112×92 chacune, ce procédé permettra de résoudre une matrice de taille 100×100 au lieu d'une matrice de taille 10304×10304 . Ce qui facilite énormément les calculs et permet un gain considérable de temps.

La $i^{\text{ème}}$ valeur propre μ_i de la matrice de covariance Cov représente la variance des visages contenus dans la base de données après leurs projections sur le visage propre V_i associé à cette valeur propre μ_i [9]. Plus les valeurs propres sont grandes, plus les caractéristiques des visages d'individus différents sont discriminées.

Les vecteurs propres de Cov doivent être ordonnés d'une manière décroissante suivant les valeurs propres qui leurs sont associées. La plus grande information est contenue dans les premiers vecteurs propres, ce qui permet de négliger les vecteurs propres dont les valeurs propres associées sont faibles. Les vecteurs propres gardés (dont les valeurs propres sont les plus grandes) constituent les visages propres « EigenFaces ». Après leur normalisation, ils vont former une base orthonormale pour l'espace vectoriel dit "espace de visage" ou "face space" [2].



FIG III.4 : Les cinq premiers visages propres correspondant aux plus grandes valeurs propres.

Chaque images de visage dans l'ensemble d'apprentissage peut être décrite comme une combinaison linéaire de ces visages propres.

$$X_i \cong \sum_{k=1}^{M'} w_k^i V_k \quad \text{avec} \quad i = 1, \dots, N \quad (\text{III. 14})$$

Avec :

V_k : les vecteurs (visages) propres gardés.

M' : le nombre de vecteurs (visages) propres gardés.

w_k^i : les poids des visages propres de la $i^{\text{ème}}$ image.

Pour trouver le poids de chaque visage propre, l'image originale est projetée sur l'espace engendré par les visages propres avec une simple opération vectorielle.

$$w_k^i = V_k^T \Phi_i \quad (\text{III. 15})$$

Donc :

$$w_k^i = V_k^T (X_i - I_{\text{moy}}) \quad (\text{III. 16})$$

$k = 1, \dots, M'$ (M' : le nombre de vecteurs propres gardés)
 $i = 1, \dots, N$

Chaque image I_j de l'ensemble d'apprentissage sera alors représentée par le vecteur W_j tel que :

$$W_j = \begin{bmatrix} w_1^j \\ w_2^j \\ \vdots \\ \vdots \\ w_{M'}^j \end{bmatrix} \quad (\text{III. 17})$$

La matrice W va contenir tous les coefficients de projection des images de l'ensemble d'apprentissage.

$$W = [W_1 \quad W_2 \quad \dots \quad \dots \quad \dots \quad W_N] \quad (\text{III. 18})$$

L'utilisation de l'algorithme de la PCA, tout seul, pour la reconnaissance de visages ne fournit pas de bons résultats surtout dans des conditions où la pose, l'éclairage ou les expressions faciales varient entre les images de la base de données et celle de tests. Pour cette raison, on trouve l'analyse en composantes principales utilisée pour le prétraitement des images de visages dans le but de réduire leur dimension et par ailleurs la complexité du système. Par la suite, un algorithme plus performant se chargera de la séparation des différentes classes.

Cependant, l'étude théorique de l'algorithme PCA reste très pédagogique et permet d'acquérir de solides bases pour la reconnaissance 2D du visage [2].

4. L'analyse discriminante linéaire :

L'analyse discriminante linéaire est une méthode statistique qui étudie des données provenant de groupes connus a priori appelés classes. L'idée de base de cette analyse est de

trouver la transformation linéaire permettant d'assurer la meilleure séparation possible de ces classes.

L'analyse discriminante linéaire a fourni de bons résultats dans plusieurs domaines, dont la reconnaissance de la parole et celle de visages. Contrairement à l'analyse en composantes principales qui encode l'information dans un espace orthogonal linéaire, la LDA encode l'information discriminatoire dans un espace linéaire séparable dont la base n'est pas nécessairement orthogonale [32].

Dans cette section, on verra les différents algorithmes de l'analyse discriminante linéaire utilisés dans la reconnaissance de visages tout en expliquant leurs principes. Là aussi, on ne s'intéressera qu'à la phase d'apprentissage.

4.1.Principe de la LDA :

Dans le but d'assurer une très bonne séparation des données appartenant à des classes différentes, l'algorithme LDA cherche à trouver la projection optimale W_{opt} permettant de maximiser le rapport de la dispersion inter-classes S_b contre la dispersion intra-classes S_w , c'est ce qu'on appelle le critère de Fisher. Cela revient à maximiser les variances des données appartenant à des classes différentes tout en minimisant les variances des données appartenant à la même classe.

$$W_{opt} = \arg \max_W \frac{|WS_bW^T|}{|WS_wW^T|} = [W_1 \quad W_2 \quad \dots \quad W_k] \quad (III.19)$$

S_b : la matrice de dispersion inter – classes.

S_w : la matrice de dispersion intra – classes.

Les colonnes W_1, W_2, \dots, W_m de W_{opt} représentent les vecteurs propres généralisés de S_b et S_w associés aux valeurs propres généralisées décroissantes $\lambda_1, \lambda_2, \dots, \lambda_m$. C'est-à-dire :

$$S_b W_i = \lambda_i S_w W_i \quad i = 1, \dots, k \quad (III.20)$$

On peut donc trouver la matrice W_{opt} en trouvant les vecteurs propres de la matrice $(S_w^{-1}S_b)$. Car en multipliant l'équation précédente par S_w^{-1} on aura :

$$S_w^{-1}S_b W_i = S_w^{-1}\lambda_i S_w W_i \quad (III.21)$$

Ce qui est équivalent à :

$$(S_w^{-1}S_b)W_i = \lambda_i W_i \quad (III.22)$$

On voit bien que W_i sont les vecteurs propres de la matrice $(S_w^{-1}S_b)$.

Il existe d'autres variantes du critère de Fisher [33] :

$$W_{opt} = \arg \max_W \frac{|WS_tW^T|}{|WS_wW^T|} \quad (III.23)$$

$$W_{opt} = \arg \max_W \frac{|WS_bW^T|}{|WS_tW^T|} \quad (III.24)$$

S_t : la matrice de dispersion totale $S_t = S_b + S_w$.

4.2. La LDA pour la reconnaissance de visages :

L'analyse en composantes principales va s'effectuer sur les images de visages qui sont des données de grande dimension. Cela engendre des matrices de grande taille et par ailleurs la solution nécessitant l'inversion directe de la matrice S_w afin d'obtenir les valeurs et les vecteurs propres sera numériquement instable car cette matrice peut être singulière, donc non inversible. Ceci provient du fait que le rang de S_w est toujours inférieur ou égal à $N - C$ et le nombre de pixels dans une image est très grand devant le nombre des images de l'ensemble d'apprentissage ($n \gg N$), ce problème s'appelle en anglais « small sample size problem ».

Afin d'éviter le problème de singularité de la matrice S_w , plusieurs approches peuvent être utilisées. Dans notre travail on va s'intéresser à deux de ces approches : FisherFaces (PCA+LDA) et Directe LDA (D-LDA).

Dans l'approche FisherFaces le problème est résolu en réduisant la dimension des images par l'utilisation de l'analyse en composantes principales. Par la suite, l'analyse discriminante linéaire est appliquée aux images dont la dimension est réduite.

L'approche directe LDA consiste à effectuer une diagonalisation simultanée de S_b et S_w afin de trouver la matrice W_{opt} , permettant de maximiser le rapport, donnée par l'équation (III.19). Cela peut se traduire par ces deux équations :

$$W_{opt} S_w W_{opt}^T = I \quad (III. 25)$$

$$W_{opt} S_b W_{opt}^T = \Lambda \quad (III. 26)$$

I : la matrice d'identité.

Λ : une matrice diagonale dont les éléments diagonaux sont ordonnés de façon décroissante.

Dans la suite de cette section on détaillera les deux algorithmes : FisherFaces et D-LDA.

4.3.FisherFaces (PCA+LDA) :

Dans l'algorithme FisherFaces, l'analyse en composantes principales est premièrement utilisée pour projeter les images de l'espace initial (espace image) au sous-espace de visages dans le but de réduire la dimension de ces images et d'avoir une matrice S_w de rang égal à sa taille (l'espace nul¹ de S_w sera éliminé). Chaque image sera alors représentée par ses coefficients de projections (les poids des visages propres).

$$Z = V^T \Phi \quad (III. 27)$$

Avec : $\Phi_i = X_i - I_{moy}$

Φ : la matrice contenant les images centrées.

V : la matrice de transformation de PCA.

Z : la matrice qui contient la représentation des images par ses coefficients de projection.

La matrice Z qui est de petite taille comparée à Φ servira comme entrée à l'algorithme LDA. Par la suite, on calcul les matrices de dispersion inter-classes S_b et intra-classes S_w en utilisant les formules suivantes :

$$S_b = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T = \Phi_b \Phi_b^T \quad (III. 28)$$

¹ L'espace nul de S_w est l'espace engendré par des vecteurs V_i tel que $S_w V_i = 0$.

$$S_w = \sum_{i=1}^c \sum_{j=1}^{n_i} (Z_j^i - m_i) (Z_j^i - m_i)^T = \Phi_w \Phi_w^T \quad (III.29)$$

$$m_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Z_j^i \quad (III.30)$$

$$m = \frac{1}{C \times n_i} \sum_{i=1}^c \sum_{j=1}^{n_i} Z_j^i = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{n_i} Z_j^i \quad (III.31)$$

$$\Phi_b = [\sqrt{n_1} (m_1 - m) \quad \sqrt{n_2} (m_2 - m) \quad \dots \quad \sqrt{n_c} (m_c - m)] \quad (III.32)$$

$$\Phi_w = [(Z_1^1 - m_1) \quad (Z_2^1 - m_1) \quad \dots \quad (Z_1^2 - m_2) \quad \dots \quad (Z_{n_c}^c - m_c)] \quad (III.33)$$

La taille de Φ_b est $k \times C$.

La taille de Φ_w est $k \times N$.

N : le nombre d'images dans l'ensemble d'apprentissage.

n_i : le nombre d'images dans la classe i

m_i : l'image moyenne de la classe i .

m : l'image moyenne de l'ensemble d'apprentissage.

Z_j^i : la $j^{\text{ème}}$ image de la classe i .

Après la réduction de dimension effectuée par la PCA, la taille de la matrice S_w sera petite ($k \times k$) et son rang égal à k (k représente le nombre de vecteurs propres qui forment l'espace de projection de la PCA). De ce fait cette matrice ne sera pas singulière ce qui permet son inversion.

Les vecteurs propres généralisés de S_w et S_b sont alors calculés et ordonnés de manière décroissante suivant les valeurs propres généralisées qui leurs sont associées. Les vecteurs propres correspondant aux plus grandes valeurs propres sont gardés (le même raisonnement que celui expliqué pour la PCA) et formeront la matrice de transformation W . Tel que les images dans ce nouvel espace s'obtiennent comme suit :

$$Y = W^T Z \quad (III.34)$$

Donc :

$$Y = W^T V^T \Phi = (VW)^T \Phi \quad (III.35)$$

La projection optimale qui maximise le rapport $\frac{|W_{opt} S_b W_{opt}^T|}{|W_{opt} S_w W_{opt}^T|}$ est :

$$W_{opt} = (VW)^T \tag{III.36}$$

$$Y = W_{opt} \Phi \tag{III.37}$$

Cette approche assure de bons résultats dans le cas où la projection des images par la PCA n'élimine pas une grande quantité de l'information discriminatoire. Dans le cas contraire, les résultats seront mauvais.

Afin d'obtenir de meilleurs résultats on utilise l'approche Directe LDA qui ne nécessite pas une réduction de dimension et qui a été décrite par H. Yu et J. Yang dans [33].

4.4.Direct-LDA :

Dans l'algorithme FisherFaces l'étape consistant à réduire la dimension par la PCA élimine à la fois l'espace nul de S_b et S_w . Cependant les études de Chen et al [34] ont montré que la plupart de l'information discriminatoire est contenue dans l'espace nul de S_w lorsque la projection de S_b obtenue par la PCA est différente de zéro, ce qui engendre une perte de l'information utile [32]. Pour éviter cela, L'algorithme D-LDA s'applique directement sur les données de grande dimension; il consiste à éliminer l'espace nul de S_b , qui ne contient aucune information utile plutôt que d'éliminer l'espace nul de S_w . Cela peut être accompli en diagonalisant S_b d'abord et ensuite, on diagonalise S_w .

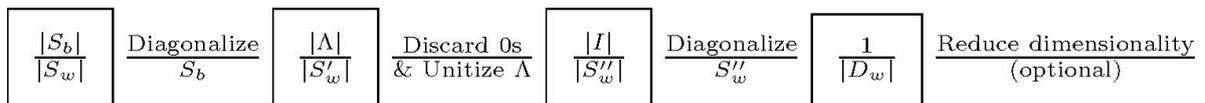


FIG III.5 : Schéma de l'agorithme D-LDA.

L'algorithme D-LDA peut être expliqué comme suit :

Tout d'abord il faut calculer les matrices de dispersion inter-classes S_b et intra-classes S_w en utilisant les équations (III.28) et (III.29). Une fois ceci effectué, on diagonalise S_b

afin d'éliminer son espace nul. Cette diagonalisation consiste à trouver la matrice V qui vérifie l'équation suivante :

$$V^T S_b V = \Lambda \quad (III. 38)$$

Avec $V^T V = I$ et Λ est une matrice diagonale dont les éléments diagonaux sont ordonnés de manière décroissante.

Cette diagonalisation peut s'effectuer en calculant les valeurs et vecteurs propres de S_b , mais comme S_b est de grande taille ($n \times n$), ce calcul sera difficile. Pour éviter cela, on utilisera le même procédé expliqué précédemment dans la PCA.

Chaque colonne de la matrice V est un vecteur propre de S_b et la matrice Λ contiendra toutes les valeurs propres de S_b . Comme la matrice S_b peut-être singulière alors certaines valeurs propres peuvent être égales ou proches de zéro. Il est alors nécessaire d'éliminer ces vecteurs et valeurs propres car ils ne contiennent aucune information discriminatoire.

Soit Y la matrice de taille $n \times k$ qui contient les k premières colonnes de V (n est le nombre de pixels dans une image) alors on aura :

$$Y^T S_b Y = D_b \quad (III. 39)$$

D_b : est la sous-matrice de Λ de taille $k \times k$ qui contient les valeurs propres strictement supérieures à zéros de S_b .

On calcule maintenant la matrice $Z = Y D_b^{-1/2}$ et comme :

$$(Y D_b^{-1/2})^T S_b (Y D_b^{-1/2}) = I \quad (III. 40)$$

Donc :

$$Z^T S_b Z = I \quad (III. 41)$$

On voit bien que cette matrice Z transforme la matrice S_b en une matrice identité et réduit la dimension de n à k .

Par la suite, la matrice S_w est projetée en utilisant la transformée Z et le résultat sera diagonalisé, ce qui donne :

$$U^T(Z^T S_w Z)U = D_w \quad (III.42)$$

Chaque colonne de U contient un vecteur propre de la matrice $(Z^T S_w Z)$.

D_w : contient toutes les valeurs propres de la matrice $(Z^T S_w Z)$.

La matrice D_w peut contenir des zéros dans sa diagonale. Comme l'objectif est de maximiser le rapport entre la dispersion inter-classes et la dispersion intra-classes, on peut alors ordonner les valeurs propres contenues dans D_w et éliminer les plus grandes valeurs propres, ainsi que les vecteurs propres correspondants.

On voit bien que la matrice $A = U^T Z^T$ diagonalise à la fois le numérateur et le dénominateur de critère de Fisher donné par l'équation (III. 19) tel que :

$$A S_w A^T = D_w \quad \text{et} \quad A S_b A^T = I$$

La transformation finale est la suivante :

$$W_{opt} = D_w^{-1/2} A \quad (III.43)$$

La projection des images initiales dans ce nouvel espace est le suivant :

$$Q = W_{opt} X \quad (III.44)$$

Chaque colonne de X est une image de l'ensemble d'apprentissage. La matrice Q , quant à elle, contient les coefficients de projection des images de l'ensemble d'apprentissage dans le sous espace engendré par les colonnes de la matrice W_{opt}^T .

Cet algorithme s'avère plus efficace que celui de FisherFaces car la plupart de l'information discriminatoire contenue dans l'espace nul de S_w est sauvegardée ce qui n'est pas le cas dans FisherFaces. En plus, il peut être lui aussi utilisé dans la réduction de dimension en se basant sur le même principe que celui de la PCA.

Pour les problèmes de classification, ce qui est le cas pour la reconnaissance de visage, la réduction de dimension en ne gardant que les vecteurs propres de S_b (correspondant aux

plus grandes valeurs propres) peut engendrer un chevauchement entre les classes. Dans le but d'éviter cela, une procédure de réduction de dimension tout en gardant les classes très bien séparées a été développée pour les données de petite dimension par R.Lotlikar et R.Kothari dans [35] et s'appelle Fractional-Step LDA (F-LDA). Par la suite les chercheurs J.Lu, K.N. Plataniotis, et A.N. Venetsanopoulos ont adapté dans [36] cette nouvelle procédure pour des données de grande dimension et ils l'ont appelé Direct- Fractional step LDA (DF-LDA).

4.5.Direct Fractional step LDA :

Dans cet algorithme on utilise la matrice de dispersion inter-classe pondérée S_{bp} à la place de S_b . Sa formule est la suivante :

$$S_{bp} = \sum_{i=1}^c \frac{n_i}{N} \sum_{j=1}^c \omega(d_{ij}) (m_i - m_j)(m_i - m_j)^T = \sum_{i=1}^c \Phi_i \Phi_i^T \quad (III.45)$$

C : le nombre de classes.

N : le nombre d'images dans l'ensemble d'apprentissage.

m_i : l'image moyenne de la classe i .

d_{ij} : la distance euclidienne entre les images moyennes des classes i et j .

$\omega(\cdot)$: fonction de pondération.

$$\Phi_i = \sqrt{\frac{n_i}{N}} \sum_{j=1}^c \sqrt{\omega(d_{ij})} (m_i - m_j) \quad (III.46)$$

La fonction de pondération $\omega(\cdot)$ est une fonction monotone décroissante parce que les classes qui sont plus près les unes des autres sont susceptibles d'être confuses ce qui demande une grande pondération pour les séparer. Pour cela, la fonction $\omega(\cdot)$ doit décroître plus rapidement que $\|m_i - m_j\|^2$. La fonction qui satisfait cette condition est de la forme d_{ij}^{-3}, d_{ij}^{-4} mais pas d_{ij}^{-1}, d_{ij}^{-2} .

Dans cet algorithme au lieu d'utiliser le critère de Fisher, on utilisera sa variante donnée par la formule suivante:

$$W_{opt} = \arg \max_W \frac{|WS_{bp}W^T|}{|WS_{tot}W^T|} \quad (III.47)$$

$$S_{tot} = S_{bp} + S_w \quad (III.48)$$

Avec S_w donnée par l'équation (III.29).

Cette variante est équivalente au critère de Fisher. Cela peut être démontré par le théorème suivant :

$$\forall x \in R^n, \text{ si } f(x) \geq 0, \quad g(x) > 0 \text{ et } f(x) + g(x) > 0$$

$$\text{Posant: } h_1(x) = \frac{f(x)}{g(x)} \quad \text{et} \quad h_2(x) = \frac{f(x)}{f(x) + g(x)}$$

la fonction $h_1(x)$ possède un maximum au point $x_0 \in R^n$ (on inclue l'infini positif)
SSI $h_2(x)$ possède un maximum au même point x_0 .

La motivation pour utiliser DF-LDA résulte de la considération suivante : Supposons qu'on veut réduire la dimension de k à $k - 1$. Pour faire cela, on doit calculer et ordonner les vecteurs propres de S_{bp} suivant les valeurs propres qui leur sont associées. Les vecteurs propres sont $[V_1 \dots V_k]$. Afin d'obtenir une dimension $k - 1$ les données sont projetées sur le sous-espace engendré par les vecteurs $[V_1 \dots V_{k-1}]$. Quand on a affaire à plusieurs classes de données, il peut y avoir deux classes i et j dont les projections se superposent (ou sont très proches). Donc au lieu de réduire la dimension de k à $k - 1$ directement la DF-LDA élimine le vecteur V_k après r étapes fractionnaires (r fractional step). Dans chaque étape, les vecteurs propres de S_{bp} sont recalculés en se basant sur le changement de $\omega(d_{ij})$ de sorte que le sous-espace de dimension $k - 1$ sera réorienté et le chevauchement entre deux classes dans l'espace de sortie sera évité [35] [36].

Voici l'algorithme DF-LDA :

La première étape de cet algorithme est une D-LDA et la deuxième étape est une F-LDA.

a. La première étape (D-LDA):

Cette étape consiste à maximiser le rapport des dispersions inter-classes et intra-classes en utilisant la variante du critère de Fisher et la matrice de dispersion inter-classes pondérée.

Pour commencer, la matrice $\Phi = [\Phi_1 \ \Phi_2 \ \dots \ \Phi_C]$ doit être calculée. Ensuite, on calcule les vecteurs et valeurs propres de $S_{bp} = \Phi\Phi^T$ en calculant ceux de la matrice de petite taille $\Phi^T\Phi$ comme expliqué dans la section réservée à la PCA.

Posons V la matrice dont chaque colonne est un vecteur propre de S_{bp} . Ces vecteurs propres sont ordonnés d'une manière à ce que les valeurs propres correspondantes sont dans un ordre décroissant. Les k vecteurs propres, dont les valeurs propres associées sont supérieures à zéro, seront gardés et les autres éliminés. On aura donc :

$$V^T S_{bp} V = \Lambda_b \quad (III.49)$$

V est la matrice qui contient les k vecteurs propres. Avec $k \leq C - 1$ car $\text{Rang}(S_{bp}) = \min(n, C - 1)$ où n représente le nombre de pixels dans l'image. Dans notre mémoire on prend $k = C - 1$.

Λ_b la matrice qui contient comme éléments diagonaux les valeurs propres supérieures à zéro qui correspondent aux vecteurs propres contenus dans V .

Posons $U = V\Lambda_b^{-1/2}$. On doit maintenant calculer les vecteurs et valeurs propres de la matrice $U^T S_{tot} U$. On peut donc écrire :

$$Z^T U^T S_{tot} U Z = D_w \quad (III.50)$$

Z est la matrice qui contient les vecteurs propres de $U^T S_{tot} U$.

D_w est une matrice diagonale qui contient comme éléments diagonaux les valeurs propres de $U^T S_{tot} U$.

Afin de maximiser le rapport donné par la variante du critère de Fisher, on peut éliminer les plus grandes valeurs propres de $U^T S_{tot} U$ ainsi que les vecteurs propres correspondants. Supposons que le nombre de vecteurs propres gardés est égal à $l \leq k$ et P et Λ_w sont respectivement les matrices contenant ces vecteurs propres et les valeurs propres correspondantes. Dans notre mémoire on prend $l = k$.

La matrice de transformation W de cette étape permet de maximiser la variante du critère de Fisher et de réduire la dimension de n à l . Cette transformation est donnée par :

$$W = \Lambda_w^{-1/2} P^T U^T \quad (III.51)$$

Soient X les images initiales et E les images projetées.

$$E = WX \quad (III.52)$$

b. La deuxième étape (F-LDA) :

Cette étape consiste à réduire la dimension de l à l' tout en gardant les classes très bien séparées entre elles.

Les images projetées contenues dans E constituent l'entrée de cette étape. Soit la matrice de transformation de cette étape W_f . On initialise $W_f = I$ (I la matrice d'identité de taille $l \times l$). Pour réduire la dimension de l à $l - 1$ on utilise r étapes fractionnaires. La réduction se fait comme suit :

Pour chaque étape fractionnaire u (u allant de 1 jusqu'à r) les opérations suivantes sont répétées :

Les données d'entrée sont projetées en utilisant la transformation W_f . On aura donc :

$$F = W_f^T E \quad (III.53)$$

Une transformation d'échelle est appliquée sur la matrice C dans le but de comprimer la dernière ligne de F . Cela signifie que la distance entre deux classes suivant le vecteur à éliminer soit très petite. Cette transformation est donnée comme suit :

$$T = \varphi(F, \alpha^u) \quad \text{tel que} \quad T^i = \begin{cases} \alpha^u F^i & \text{pour } i = l \\ F^i & \text{pour } i \neq l \end{cases} \quad (III.54)$$

T^i et F^i sont respectivement la $i^{\text{ème}}$ ligne de T et F .

α^u est le facteur de réduction pour la $u^{\text{ème}}$ étape fractionnaire.

Par la suite, on calcule la matrice de dispersion pondérée S_{bp} de T et on détermine ses valeurs propres ordonnées de manière décroissante $\lambda_1, \lambda_2, \dots, \lambda_l$ et les vecteurs propres Vp_1, Vp_2, \dots, Vp_l correspondants. La matrice de transformation de cette étape sera modifiée comme suit :

$$W_f = W_f Vp \quad \text{tel que :} \quad Vp = [Vp_1, Vp_2, \dots, Vp_l] \quad (III.55)$$

Une fois les r étapes fractionnaires effectuées, l'espace de dimension $l - 1$ sera réorienté de façon à ce qu'il n'y ait pas de chevauchement entre les classes. Alors, le vecteur propre Vp_l peut être éliminé et la dimension sera réduite à $l - 1$.

Ce procédé est alors répété jusqu'à ce que la dimension soit égale à l' .

La transformation finale de la DF-LDA est la suivante :

$$W_{tot} = W_f^T W \quad (III.56)$$

Les projections des images de l'ensemble d'apprentissage sont alors données comme suit :

$$\bar{X} = W_{tot} X \quad (III.57)$$

La DF-LDA est une technique très performante dans la reconnaissance de visage. Cependant elle nécessite trop de calculs donc un temps élevé pour la phase d'apprentissage.

5. La phase de reconnaissance :

Une fois l'apprentissage effectué, les projections des images ainsi que les transformations seront sauvegardés en mémoire afin d'être utilisés dans la phase de reconnaissance. Pour la PCA et FisherFaces en plus de ça, on doit aussi stocker l'image moyenne I_{moy} .

La phase de reconnaissance ou test est identique pour tous les algorithmes vus précédemment. Elle consiste à injecter une image de l'ensemble du test afin de l'identifier. Cette phase comprend plusieurs étapes.

Avant de commencer, l'image de test doit être convertie en vecteur de la même façon que dans la phase d'apprentissage.

Pour la PCA et FisherFaces l'image de test I_{test} doit être centrée et pour cela on doit lui soustraire l'image moyenne I_{moy} . Cette étape ne sera pas effectuée pour la D-LDA et DF-LDA.

Donc pour la PCA et FisherFaces :

$$\psi_{test} = I_{test} - I_{moy} \quad (III. 58)$$

Et pour la D-LDA et DF-LDA :

$$\psi_{test} = I_{test} \quad (III. 59)$$

Par la suite, le vecteur ψ_{test} sera projeté par la transformée finale de l'algorithme utilisé, et ses coefficients de projection seront calculés. Ces coefficients vont représenter l'image de test.

$$Y_{test} = W\psi_{test} = \begin{bmatrix} Y_1^{test} \\ Y_2^{test} \\ \vdots \\ Y_M^{test} \end{bmatrix} \quad (III. 60)$$

Pour la classification, différentes métriques peuvent être utilisées pour le calcul de la ressemblance entre l'image de test et celles de la base de données. Prenons par exemple la distance euclidienne comme métrique.

Les distances euclidiennes entre l'image de test et celles de la base de données seront calculées en utilisant la formule suivante :

$$d_i = \|Y_{test} - Y_i\| = \sqrt{\sum_{k=1}^{M'} (Y_k^{test} - Y_k^i)^2} \quad (III.61)$$

Y_i : les coefficients de la $i^{\text{ème}}$ image de l'ensemble d'apprentissage stockés lors de la phase d'apprentissage.

M' : la dimension des images dans l'espace de projection.

$i = 1, \dots, C$ car généralement chaque classe est représentée par un seul vecteur de coefficients calculé à partir des coefficients de projection des images de visage appartenant à cette classe.

La distance minimale est retenue et comparée à un seuil avant de prendre une décision. Le seuil est déterminé de manière expérimentale afin que le taux de fausses acceptations soit identique au taux de faux rejets (FAR=FRR). Si la distance minimale est inférieure au seuil, alors la personne à identifier appartient à la classe qui a fourni cette distance minimale. Sinon (la distance minimale supérieure au seuil) la personne à identifier n'appartient pas à la base de données et sera classée comme inconnue.

Plusieurs autres méthodes hormis la distance euclidienne permettent de calculer la ressemblance entre l'image de test et celles de la base de données. On citera :

- La norme L1 donnée par la formule :

$$d_{L1}(i) = \sum_{k=1}^{M'} |Y_k^{test} - Y_k^i| \quad (III.62)$$

- La distance de Mahalanobis donnée par la formule :

$$d_{Mah}(i) = \sqrt{\sum_{k=1}^{M'} \frac{(Y_k^{test} - Y_k^i)^2}{\mu_k}} \quad (III.63)$$

μ_k : la $k^{\text{ème}}$ valeur propre.

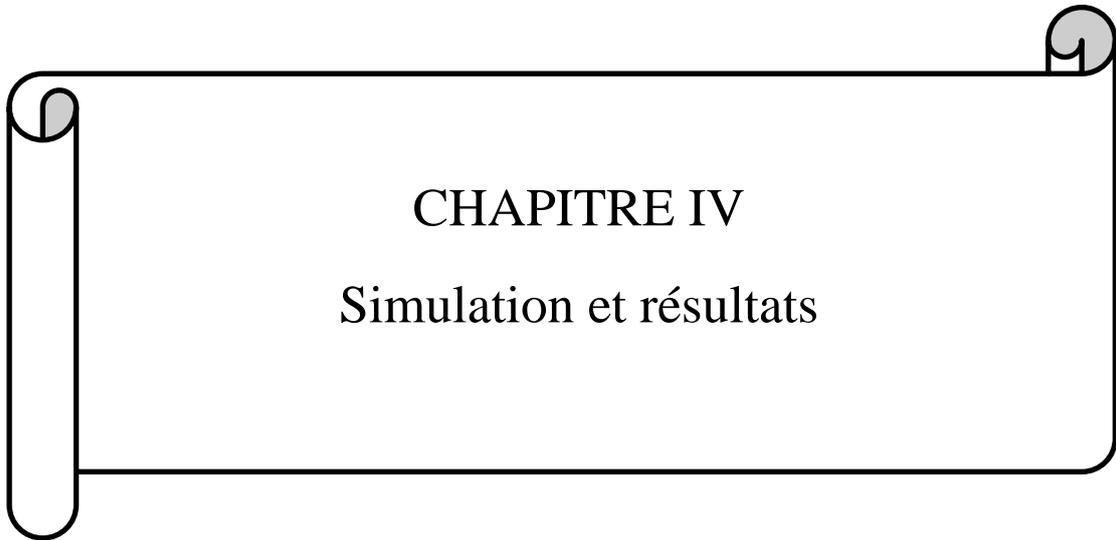
- L'angle entre deux vecteurs (images) ou la distance Cosinus donnée par la formule :

$$d_{cos}(i) = - \frac{W_{test}^T W_i}{\sqrt{\sum_{k=1}^{M'} (w_k^{test})^2} \cdot \sqrt{\sum_{k=1}^{M'} (w_k^j)^2}} \quad (III.61)$$

6. Conclusion :

Les algorithmes PCA et LDA utilisés pour la reconnaissance de visages sont des méthodes globales qui utilisent les pixels de l'image comme entrées. La PCA dé-corrèle des variables initialement corrélées et traite les données dans leur totalité sans prendre en compte la structure des classes. Par contre, la LDA (dont on a étudié et implémenté trois algorithmes : la FisherFaces, la D-LDA et DF-LDA) assure une véritable séparation des classes. La FisherFaces est une combinaison séquentielle de la PCA et de la LDA, l'étape PCA assure une réduction de dimension dans le but d'appliquer la LDA par la suite afin d'obtenir une bonne séparation des classes. La D-LDA quant à elle procède à une diagonalisation simultanée des deux matrices de dispersion inter et intra-classes afin d'éviter une perte de l'information discriminatoire, ce qui peut être le cas dans la FisherFaces à cause de la réduction de dimension par la PCA. Dans le but d'avoir une bonne séparation de données avec un petit nombre de vecteurs formant la base de projection, la DF-LDA rentre en jeu et réduit la dimension tout en réorientant l'espace de projection afin d'éviter tout chevauchement entre les classes.

La facilité d'implémentation de ces algorithmes les rend très utilisés et indispensables pour un certain nombre d'applications, surtout celles qui nécessitent une réduction de dimension.



CHAPITRE IV
Simulation et résultats

1. Introduction :

Les différents algorithmes vus dans le chapitre précédent (PCA, FisherFaces, D-LDA et DF-LDA) ont été entièrement programmés sur MATLAB dans le but de les simuler et de les comparer.

Ce chapitre est consacré à la présentation des résultats fournis par les deux techniques (PCA et LDA) en utilisant trois bases de données différentes. Pour cela, on commencera ce chapitre par une présentation des bases de données utilisées, puis on exposera les résultats de nos simulations et on finira par une comparaison de ces deux techniques.

2. Présentation des bases de données utilisées :

Les trois bases de données utilisées dans notre simulation sont : ORL (Olivetti Research Ltd) Faces94 et UMIST. Nous allons présenter brièvement chacune d'elles.

2.1. La base de données ORL :

Cette base de données [37] contient 40 personnes (36 hommes et 4 femmes) distinctes avec 10 images de visages par personne. L'âge des personnes de cette base de données varie entre 18 ans et 81 ans, avec un âge de 25 ans à 30 ans pour la majorité des personnes. Cette base de données a été développée par Ferdinando Silvestro Samaria à l'université de Cambridge en Angleterre.

Ces images ont été prises à des intervalles de temps différents en faisant varier les conditions d'éclairage, les expressions faciales et les détails faciaux (exemple : le port des lunettes). Toutes les personnes sont en position droite et frontale, avec une tolérance pour un certain mouvement latéral. Les images sont en niveaux de gris avec une résolution de 112×92 ce qui fournit une dimension $n = 10304$ après concaténation.



FIG IV.1 : Quelques images de la base de données ORL.

2.2.La base de données Faces94 :

Cette base de données [38] contient en tout 153 personnes distinctes (133 hommes et 20 femmes) avec 20 images de visages par personne. Elle a été développée par le Dr Libor Spacek à l'Université d'Essex en Angleterre.

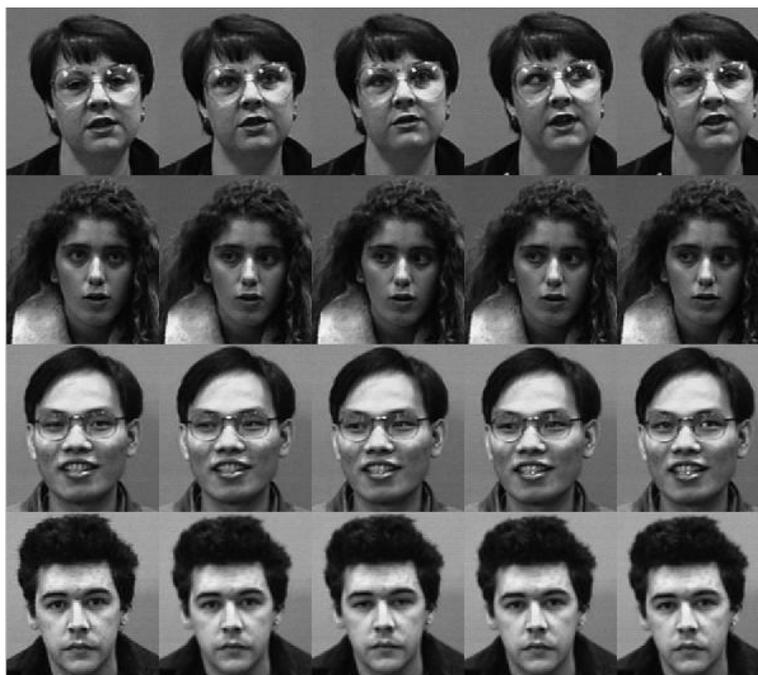


FIG IV.2 : Quelques images de la base de données Faces94.

Les images de visages de cette base de données ont été prises à l'aide d'une caméra dans des conditions d'éclairage fixe. La personne est assise à une distance fixe de la caméra, une séquence d'images de visage de la personne sera prise à des intervalles de temps très proches. Durant cette séquence la personne doit parler afin d'introduire des variations dans ses expressions faciales. Les images sont en couleur avec un arrière-plan vert et de résolution de 200×180 ce qui fournit une dimension $n = 36000$ après la concaténation.

2.3.La base de données UMIST :

Cette base de données [39] contient 20 personnes distinctes mais le nombre d'images par personne varie de 19 à 48 selon la personne. En tout, elle contient 575 images de visages. Développé par Daniel B. Graham et Nigel M. Allinson à l'université de L'université de Sheffield, Angleterre



FIG IV.3 : Quelques images de la base de données UMIST.

Les images de cette base de données couvrent une large gamme de poses, de la vue de profil à la vue frontale. Les images sont en niveaux de gris avec une résolution de 112×92 ce qui fournit une dimension $n = 10304$ après concaténation.

3. Performances des différents algorithmes :

Pour déterminer les performances du système de reconnaissance de visages se basant sur l'un des algorithmes vus précédemment, plusieurs paramètres doivent être déterminés, comme :

- Taux d'identifications correctes (le taux de reconnaissance): le pourcentage des reconnaissances justes. Sa formule est donnée par l'équation (I. 3).
- FAR : le taux de fausses acceptations. Sa formule est donnée par l'équation (I. 2).
- FRR : le taux de faux rejets. Sa formule est donnée par l'équation (I. 1).
- ERR : le taux d'égale erreur qui est le point d'intersection entre le FAR et le FRR.

On va commencer cette section par une présentation des résultats obtenus pour les qualités des différents algorithmes en terme de séparation de classes. Par la suite on exposera les taux d'identifications correctes obtenus avec les différents algorithmes pour les différentes bases de données puis on tracera les taux de fausses acceptations (FAR) et de faux rejets (FRR) en fonction du seuil de décision afin de déterminer le ERR.

3.1. Qualités des algorithmes en terme de séparation de classes :

Afin de voir les performances de chaque algorithme pour la séparation de classes, on utilisera la base de données ORL et on prendra 10 classes (personnes), avec 10 images de visages par classe. Appliquera, ensuite, à ces images les différents algorithmes tel que les transformations de projection ne contiennent que les deux vecteurs propres correspondant aux plus grandes valeurs propres. Cela dans le but de pouvoir représenter ces classes dans le plan. Les résultats obtenus sont données par la figure IV.4.

On voit bien que les classes sont très mal séparées avec la PCA et la FisherFaces qui gardent seulement 10 visages propres à la fin de l'étape PCA. En revanche la D-LDA fournit une meilleure séparation mais il existe toujours un chevauchement entre les classes. La DF-LDA, quant à elle, assure une très bonne séparation et on ne voit aucun chevauchement. Cependant l'utilisation de la FisherFaces en gardant 100 visages propres lors de la réduction

de dimension par la PCA a permis de rapprocher les images de visages de la même classe tout en éloignant les classes entre elles, cela a donné une meilleure séparation.

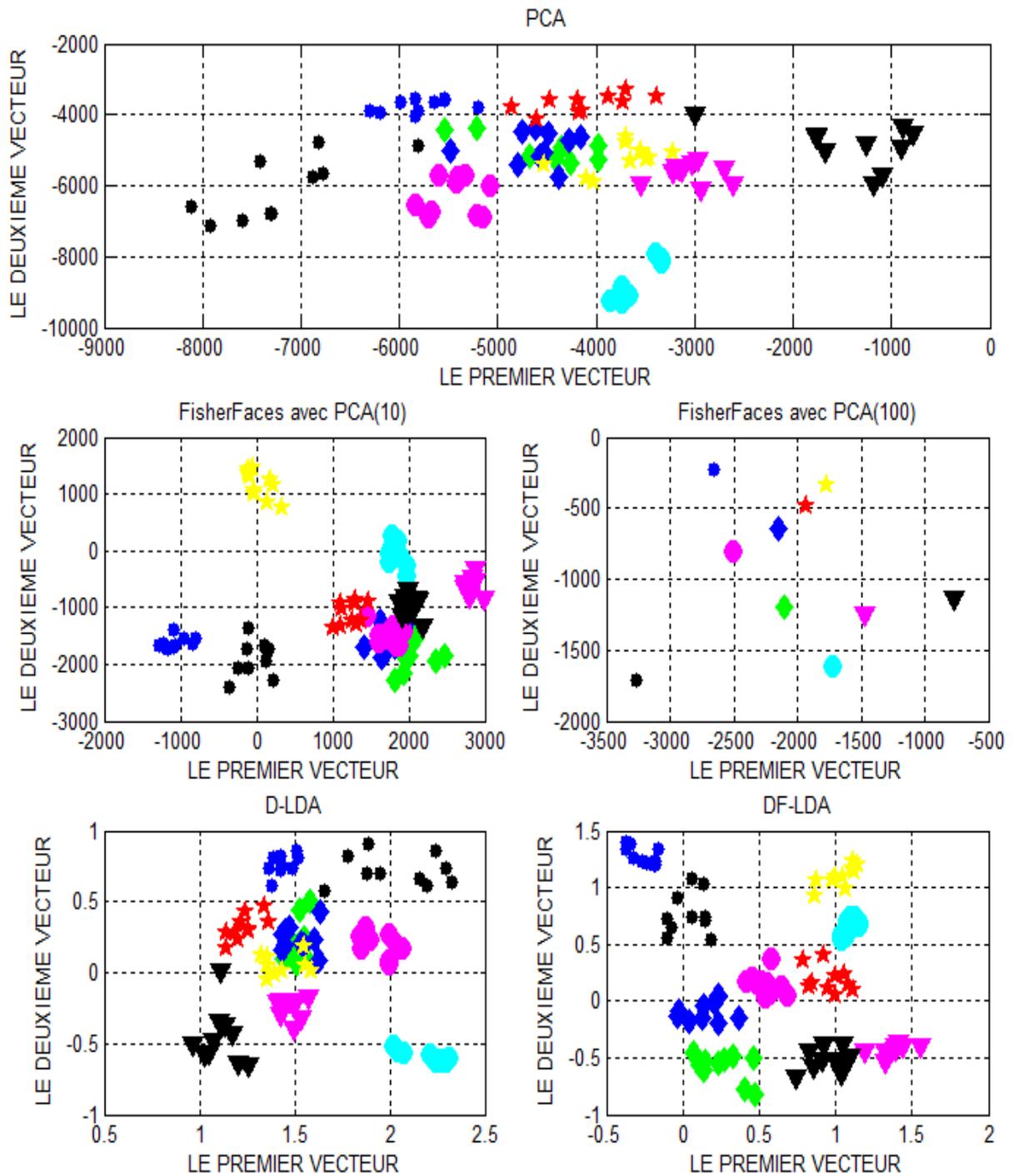


FIG IV.4 : Qualité des différents algorithmes pour la séparation de classes.

Le nombre d'étapes fractionnaires de la DF-LDA utilisée dans cet exemple est $r = 50$ et la fonction de pondération est $\omega(d_{ij}) = d_{ij}^{-8}$.

3.2. Le taux d'identifications correctes:

Pour déterminer le taux d'identifications correctes, les images de tests utilisées doivent appartenir aux personnes présentes dans l'ensemble d'apprentissage, autrement dit, on ne va pas utiliser pour le test des images de visages de personnes inconnues.

Avant de commencer, il faut diviser chacune des bases de données en un ensemble d'apprentissage et un ensemble de test. Les images d'apprentissage pour une personne donnée sont prises de manière aléatoire. Dans le but de réduire les variations, chaque calcul est répété 10 fois pour l'ORL et l'UMIST et 3 fois pour Faces94. Le taux d'identifications correctes est, alors, égal à la moyenne des taux d'identifications correctes obtenus lors de chaque calcul. Deux expériences sont réalisées.

a. La première expérience :

Le but de cette expérience est de déterminer le taux d'identifications correctes en fonction du nombre de vecteurs constituant la base de projection. On utilise, pour cela, les trois bases de données. Les résultats sont présentés sous forme de graphes.

Il est à noter que pour l'algorithme FisherFaces, deux résultats sont exposés. Dans le premier, le nombre de vecteurs gardés lors de la réduction de dimension par la PCA est égal au nombre de vecteurs formant la base de projection finale (représenté par FisherFaces sur le graphe). Cependant, dans le deuxième résultat, ce nombre de vecteurs est fixe et est égal à 70 (représenté par FisherFaces70 sur le graphe). Pour la DF-LDA on a pris le nombre des étapes fractionnaires égal à 30 et la fonction de pondération $\omega(d_{ij}) = d_{ij}^{-12}$. Durant la phase de test on a utilisé la distance euclidienne pour calculer la ressemblance entre les images de test et celles de la base de données.

La base de données ORL :

On forme un ensemble d'apprentissage contenant 200 images de visages des 40 personnes (on prend 5 images par personne). Les 200 images restantes formeront l'ensemble de test.

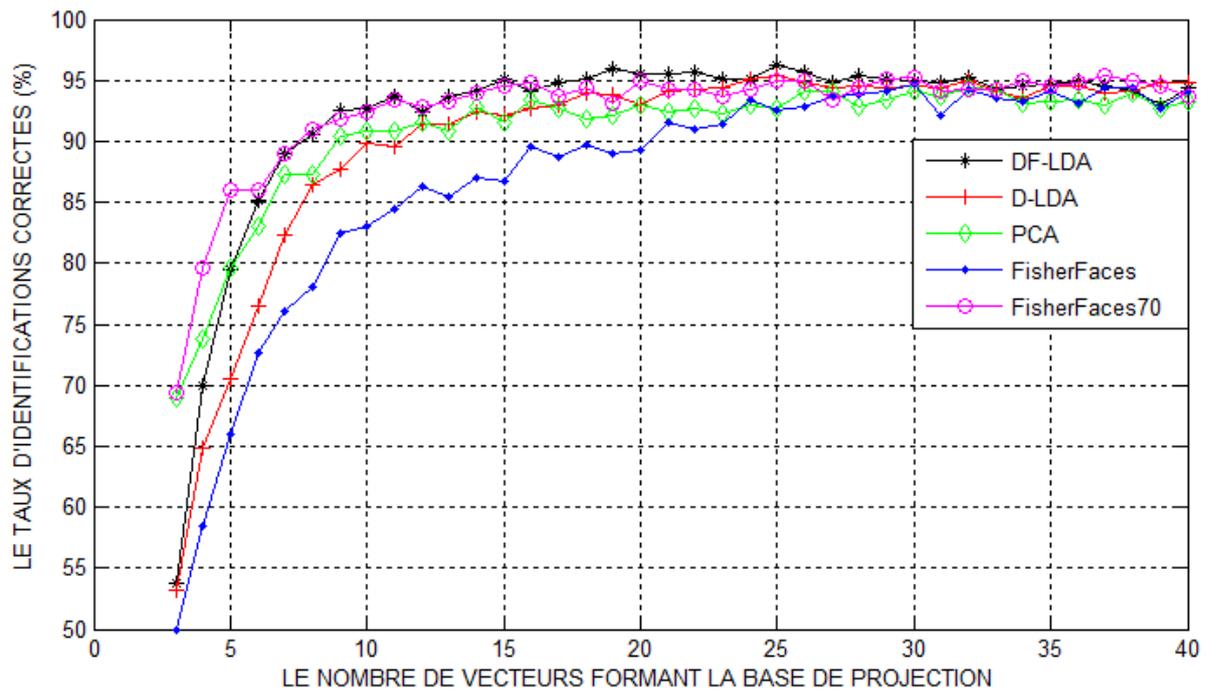


FIG IV.5 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données ORL.

La base de données Faces94 :

L'ensemble d'apprentissage contient 3 images par personne pour 133 personnes utilisées. Le reste des images de visages des mêmes personnes serviront pour le test et formeront alors l'ensemble de test.

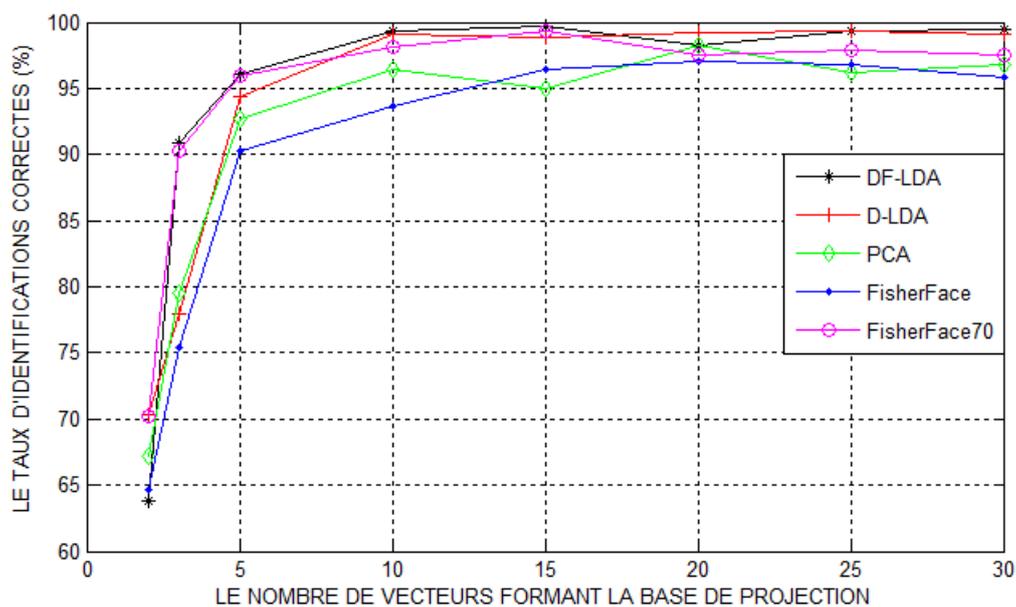


FIG IV.6 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données Faces94.

La base de données UMIST :

L'ensemble d'apprentissage est constitué de 8 images par personne pour les 20 personnes présentes dans UMIST, ce qui donne un nombre total de 160 images. Les 415 images restantes formeront l'ensemble de test. Le nombre d'étapes fractionnaires pour la DF-LDA est égal à 30 et la fonction de pondération est $\omega(d_{ij}) = d_{ij}^{-8}$.

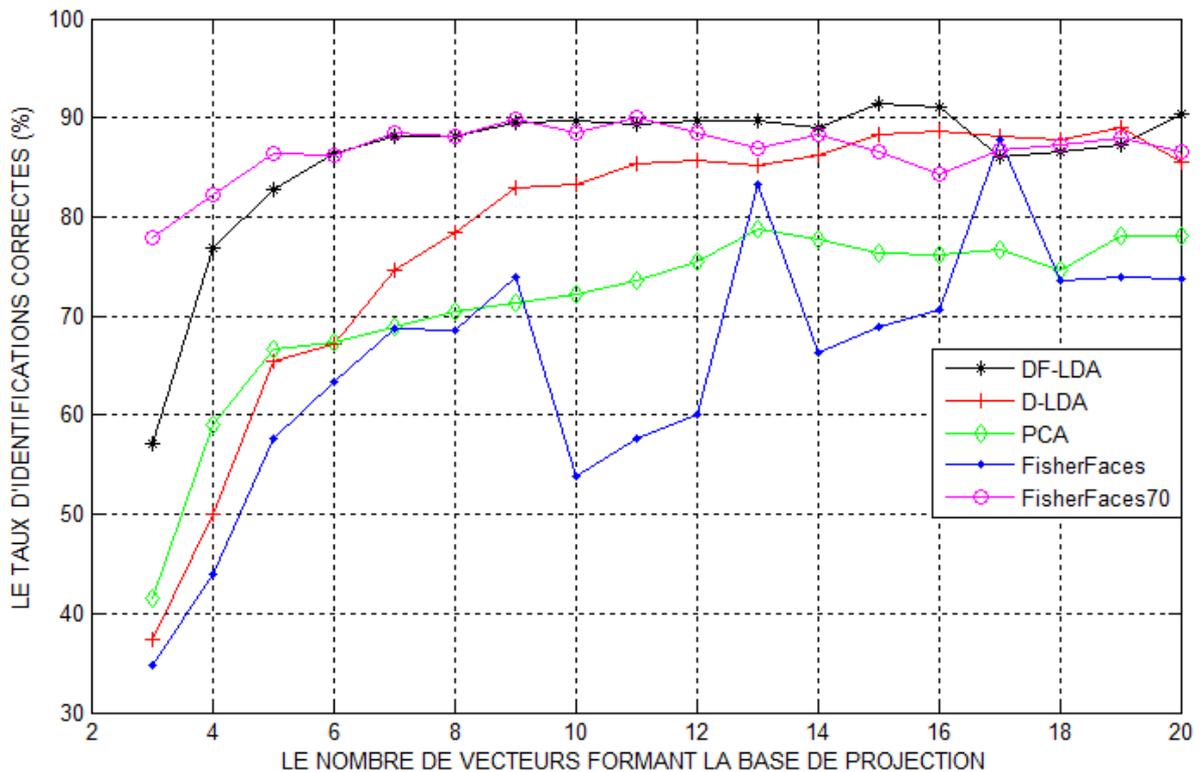


FIG IV.7 : Taux d'identifications correctes en fonction du nombre de vecteurs pour la base de données UMIST.

Interprétation des résultats :

Le taux d'identifications correctes est proportionnel au nombre de vecteurs formant la base de projection. Plus le nombre de ces vecteurs augmente, meilleure sera la précision du système. Cela est évident car en augmentant le nombre de vecteurs, on ajoute une information permettant la discrimination des personnes. En revanche, la rapidité du système diminue au fur et à mesure que le nombre de ces vecteurs augmente, ce qui donne des temps de reconnaissance élevés. Pour cela il faut toujours chercher le meilleur compromis permettant d'assurer une très bonne précision du système tout en maintenant une vitesse de traitement acceptable.

On constate d'après les figures précédentes, qu'à partir d'un certain nombre de vecteurs (pour l'ORL environ 30 pour la PCA et FisherFaces, 25 pour la D-LDA et 15 pour la DF-LDA) le taux d'identifications correctes reste constant quel que soit l'algorithme ou la base de données utilisée.

La PCA fournit de meilleurs résultats que la FisherFaces, ce qui peut être expliqué par l'élimination de l'information discriminatoire, à cause de la réduction de dimension effectuée par la PCA. En revanche, quand le nombre de vecteurs gardés après la PCA est suffisant (FisherFaces70), c'est-à-dire qu'une grande quantité de l'information discriminatoire est sauvegardée, l'algorithme FisherFaces donne un meilleur taux d'identifications correctes. Le temps de reconnaissance, quant à lui, est identique pour la PCA, FisherFaces et FisherFaces70, la seule différence est dans la durée de la phase d'apprentissage.

Les algorithmes PCA et FisherFaces ne donnent pas un bon taux d'identifications correctes quand les images de la base de données présentent une grande variation de poses et d'éclairage (le cas de l'UMIST), contrairement aux algorithmes DF-LDA, D-LDA et FisherFaces70 qui sont performants pour ces mêmes conditions. Il est alors plus intéressant de voir les variations du taux d'identifications correctes de la DF-LDA en fonction de ses paramètres (sa fonction de pondération $\omega(d_{ij}) = d_{ij}^{-2p}$ et le nombre d'étapes fractionnaires r). On a effectué cette expérience sur la base de données ORL en gardant les mêmes ensembles d'apprentissage et de test vus précédemment.

Premièrement on fixe le nombre d'étapes fractionnaires $r = 30$ puis on relève les variations du taux d'identifications correctes en fonction du paramètre p et de nombre de vecteurs propres gardés. On obtient le graphe de la figure IV.8, après simulation.

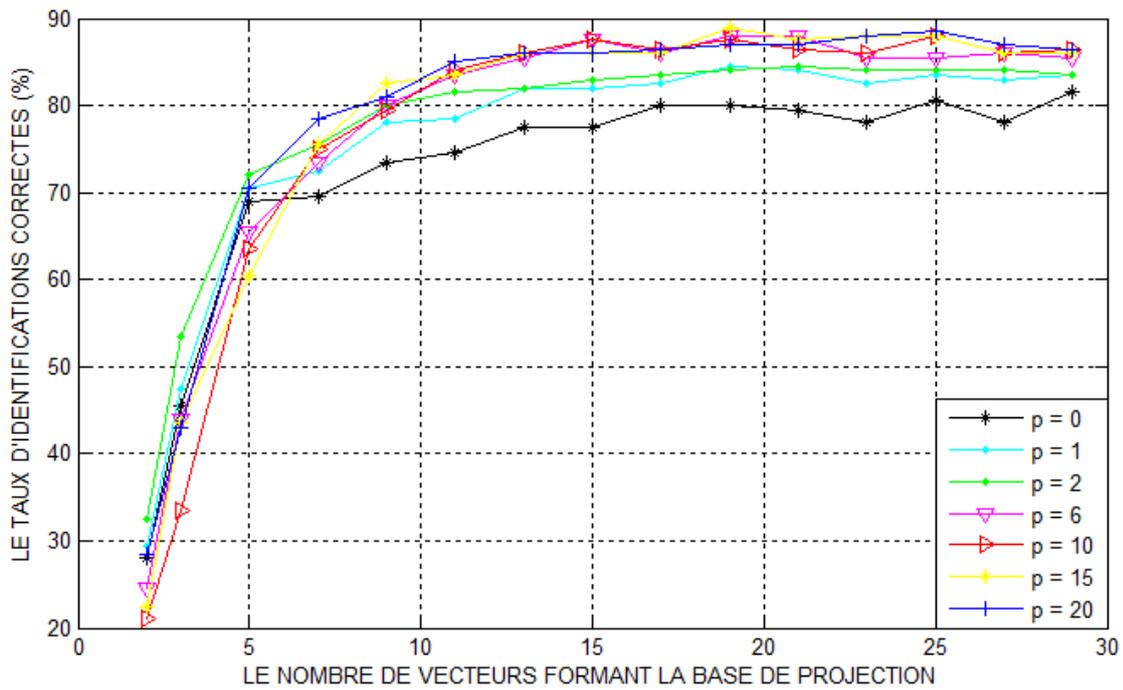


FIG IV.8 : Taux d’identifications correctes en fonction du paramètre p et du nombre de vecteurs pour la DF-LDA sur la base données ORL.

Maintenant, l’étude du taux d’identifications correctes sera en fonction de paramètre r de la DF-LDA et le nombre de vecteurs gardés. Dans la simulation, le paramètre p est fixe et est égal à 4. Les résultats obtenus sont représentés en figure IV.9.

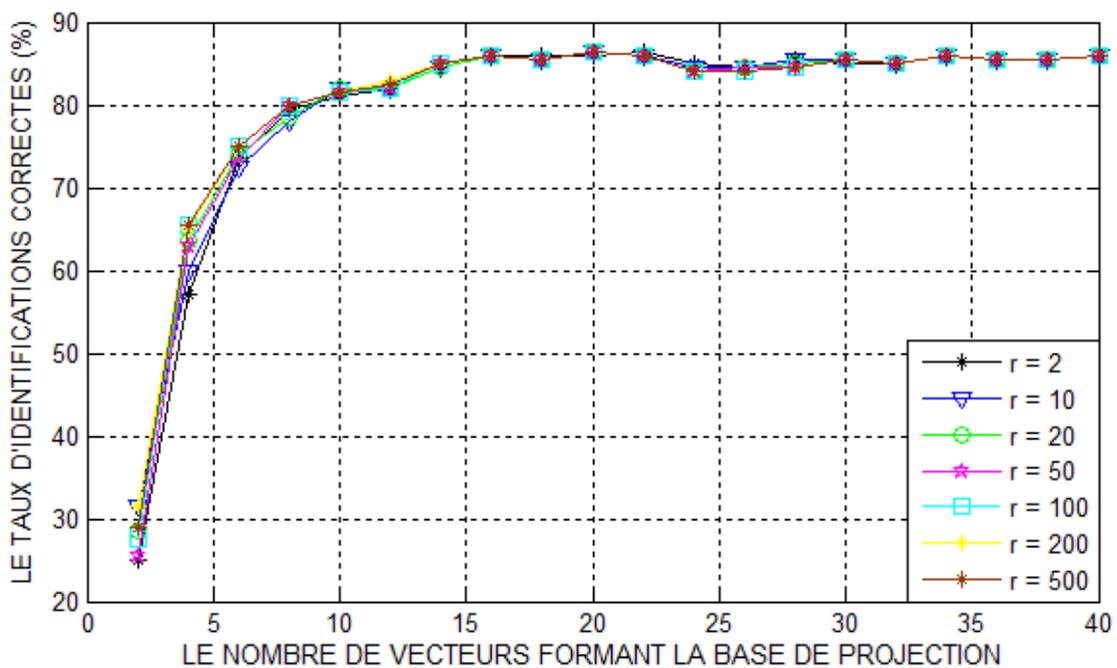


FIG IV.9 : Taux d’identifications correctes en fonction du paramètre r et du nombre de vecteurs pour la DF-LDA sur la base données ORL.

A partir des deux courbes obtenues (IV.8 et IV.9) on constate que le taux d'identification est proportionnel à p et r ; ce qui est évident, car en augmentant le paramètre p la fonction de pondération augmente pour deux classes très proches ce qui permet de les séparer. Le paramètre r , quant à lui, n'a pas une grande influence à partir d'un certain nombre de vecteurs gardés, mais si un petit nombre de vecteurs est choisi alors r influe pour le taux d'indentifications correctes.

b. La deuxième expérience :

On s'intéresse à travers cette deuxième expérience au taux d'identifications correctes en fonction du nombre des images de l'ensemble d'apprentissage. Dans cette partie le nombre de vecteurs constituant la base de données est constant. Cette expérience est effectuée seulement sur les bases de données ORL et UMIST.

La base de données ORL :

Le nombre de vecteurs formant la base de projection est égal à 50 pour la PCA, 35 pour FisherFaces, 25 pour la D-LDA et 20 pour la DF-LDA.

Le tableau IV.1 donne une idée de l'effet de nombre d'images d'apprentissage sur le taux d'identifications correctes.

Nombre d'images d'apprentissage	2	3	4	5	6	7	8	9
PCA (50)	78.9688	87.2500	91.8333	94.6500	95.6250	96.3333	97.6250	98.5000
FisherFaces (35)	69.8125	86.3571	90.3750	93.3500	95.8750	96.8333	97.5000	98.5000
D-LDA (25)	70.3125	86.8571	91.3333	92.6000	95.3125	97.1667	97.6250	98.7500
DF-LDA (20)	38.1250	85.7143	91.6250	95.2500	96.6875	97.3333	97.6250	99.2500

TAB IV.1 : Taux d'identifications correctes en fonction du nombre d'images de l'ensemble d'apprentissage pour la base de données ORL.

La base de données UMIST :

Le nombre de vecteurs formant la base de projection est égal à 20 pour la PCA, 20 pour FisherFaces, 15 pour la D-LDA et 10 pour la DF-LDA.

Nombre d'images d'apprentissage	2	4	6	8	10	12
PCA(20)	57.3458	68.2828	73.0989	74.1928	78.8267	79.0448
FisherFaces(20)	48.9533	85.7576	75.7582	89.9518	73.6267	74.6567
D-LDA(15)	62.5234	80.4444	80.4444	85.7802	86.6747	89.1200
DF-LDA(10)	52.2430	81.1717	84.4835	90.1205	91.0400	91.7612

TAB IV.2 : Taux d'identifications correctes en fonction de nombre d'image de l'ensemble d'apprentissage pour la base de données UMIST.

Interprétation des résultats :

A partir des tableaux IV.2 et IV.2, on constate qu'en augmentant le nombre d'images par personnes dans l'ensemble d'apprentissage le taux d'identifications correctes augmente. En effet, plus ce nombre augmente plus l'information aidant à mieux faire la reconnaissance sera grande.

En résumé, on peut dire que l'algorithme DF-LDA assure une bonne séparation entre les classes et fournit des taux d'identifications correctes supérieurs à ceux fournis par les autres algorithmes. La FisherFaces, quant à elle, donne un bon résultat en ce qui concerne le taux d'identifications correctes quand le nombre de vecteurs gardés après la PCA est suffisant. De plus, on peut augmenter le taux d'identifications correctes en augmentant le nombre de vecteurs constituant la base de données où le nombre d'images par personne utilisées pour l'apprentissage.

3.3.Le taux de fausses acceptations (FAR) et le taux de faux rejets (FRR) :

Pour mesurer le taux de fausses acceptations, on injecte au système de reconnaissance, des images de visage de personnes inconnues. Ce qui mène à enlever certaines personnes de

la base de données lors de la phase de l'apprentissage pour pouvoir les utiliser lors des tests. Ces personnes sont appelées imposteurs. Pour ce qui est du taux de faux rejets, on injecte au système de reconnaissance, des images de visage de personnes connues, c'est-à-dire, les images de tests doivent appartenir aux personnes présentes dans l'ensemble d'apprentissage. Ces personnes sont appelées clients.

La mesure de FAR et FRR se fait en fonction d'un seuil de décision. Le seuil de décision permet au système d'accepter ou de rejeter une personne. Soit S ce seuil de décision; lors de la reconnaissance quand la distance minimale d entre l'image de test et celles de la base de données est déterminée, elle sera comparée à ce seuil S . Si d est inférieure à S , la personne sera acceptée et dans le cas contraire elle sera rejetée. Dans un système biométrique, on cherche toujours le seuil optimal permettant de donner FAR=FRR, dans le but d'équilibrer les erreurs dues aux fausses acceptations et aux faux rejets. Car, si le FAR est grand, le nombre d'imposteurs acceptés par le système sera élevé. Dans le cas où c'est le FRR qui est grand on aura un nombre élevé de clients rejetés par le système. On a procédé à deux expériences pour cela.

a. La première expérience :

Dans ce qui suit, on cherche à comparer les différents algorithmes entre eux afin de déterminer lequel d'eux donne le plus faible ERR. Pour cela, il est intéressant de prendre un même nombre de vecteurs formant la base de projection pour les différents algorithmes. Il est à noter que pour la Fisherfaces le nombre de vecteurs propres gardés après la PCA est égal à 70. Concernant la DF-LDA, le nombre d'étapes fractionnaires est égal à 30 et sa fonction de pondération est d_{ij}^{-12} .

La base de données ORL :

L'ensemble d'apprentissage contient 30 personnes avec 5 images par personne. L'ensemble de test, quant à lui, contient 40 personnes (30 clients et 10 imposteurs). Le nombre d'images utilisées par personne dans l'ensemble de test est le suivant : 5 images par client et 10 images par imposteur. Donc on aura un ensemble d'apprentissage qui contient 150 images et un ensemble de test qui contient 450 images.

En traçant le FAR et le FRR en fonction du seuil de décision pour les différents algorithmes, on obtient les graphes suivants :

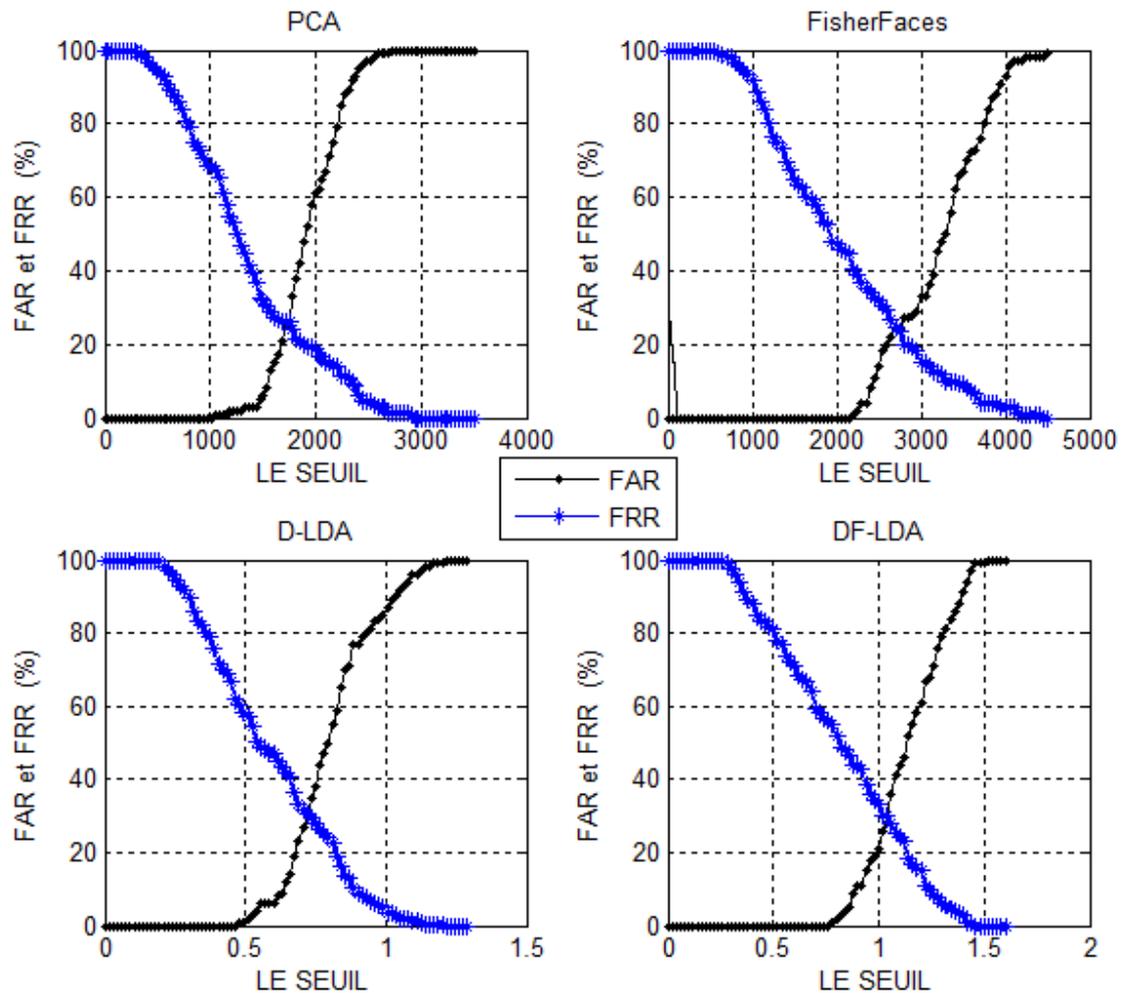


FIG IV.8 : FAR et FRR en fonction du seuil pour 15 vecteurs gardés pour l'ORL.

A partir des graphes de la figure IV.7, le seuil de décision permettant d'avoir le $FAR=FRR=ERR$ ainsi que les valeurs d'ERR pour les différents algorithmes sont données dans le tableau suivant :

Algorithme	Seuil	ERR
PCA	1750	26
FisherFaces	2725	24
D-LDA	0.7213	30.42
DF-LDA	1.044	29.6

TAB IV.3 : Valeurs d'ERR pour le seuil de décision optimal.

Les courbes ROC des différents algorithmes sont présentées dans la figure suivante :

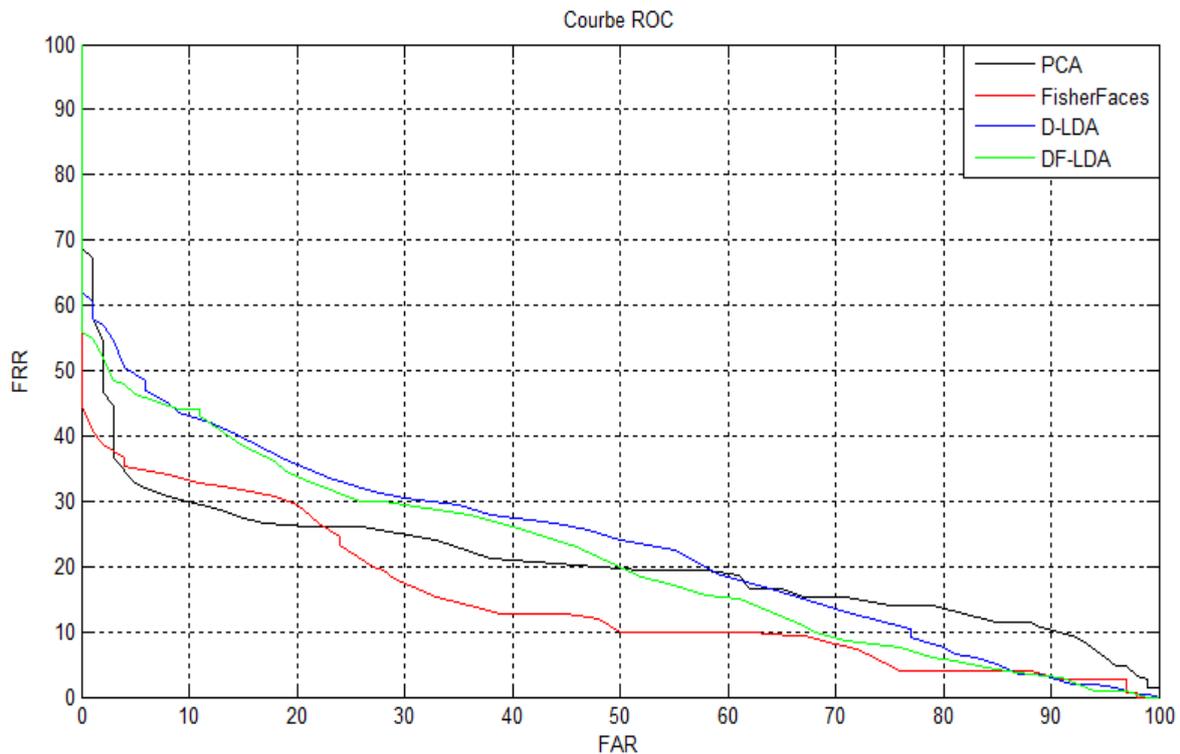


FIG IV.9 : Courbe ROC des différents algorithmes pour 15 vecteurs gardés pour l'ORL.

La base de données Faces94 :

On va utiliser seulement les images de visages des 80 premières personnes. A partir de ces images un ensemble d'apprentissage est formé par les 50 premières personnes, chacune d'elle possédant 5 images. L'ensemble de test quant à lui est formé des 80 personnes (50 clients possédant chacun 15 images et 30 imposteurs possédant chacun 20 images). Les images des clients serviront pour la mesure de FRR et les images des imposteurs serviront pour la mesure de FAR. Après simulation, on obtient les graphes de la figure IV.10.

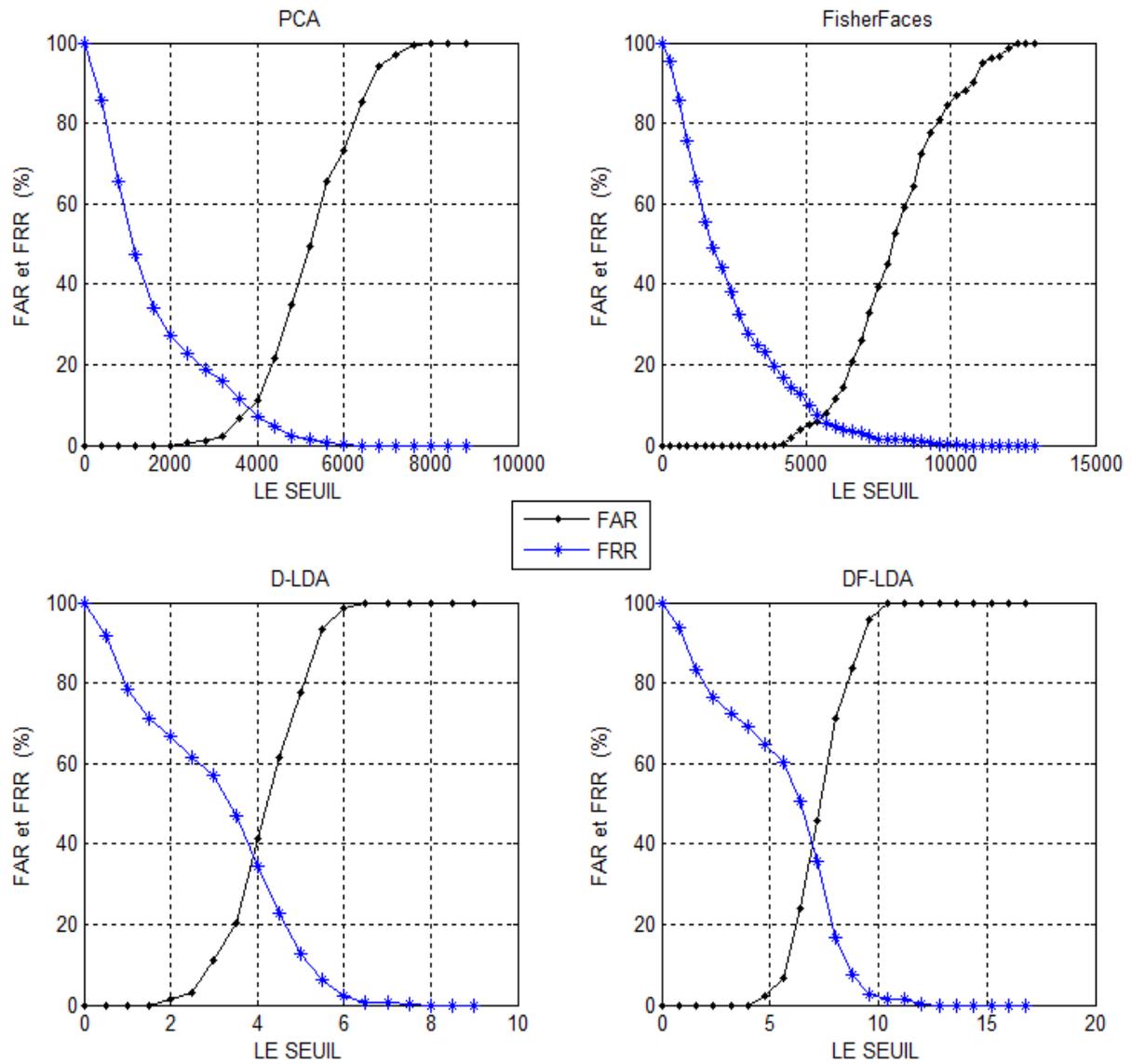


FIG IV.10 : FAR et FRR en fonction du seuil pour 15 vecteurs gardés pour Faces94.

A partir des graphes de la figure IV.10 on obtient le tableau suivant :

Algorithme	Seuil	EER(%)
PCA	3817	9.092
FisherFaces	5516	6.64
D-LDA	3.864	37.04
DF-LDA	6.981	39.71

TAB IV.4 : Valeurs d'ERR pour le seuil de décision optimal pour Faces94.

La courbe ROC suivante est alors obtenue :

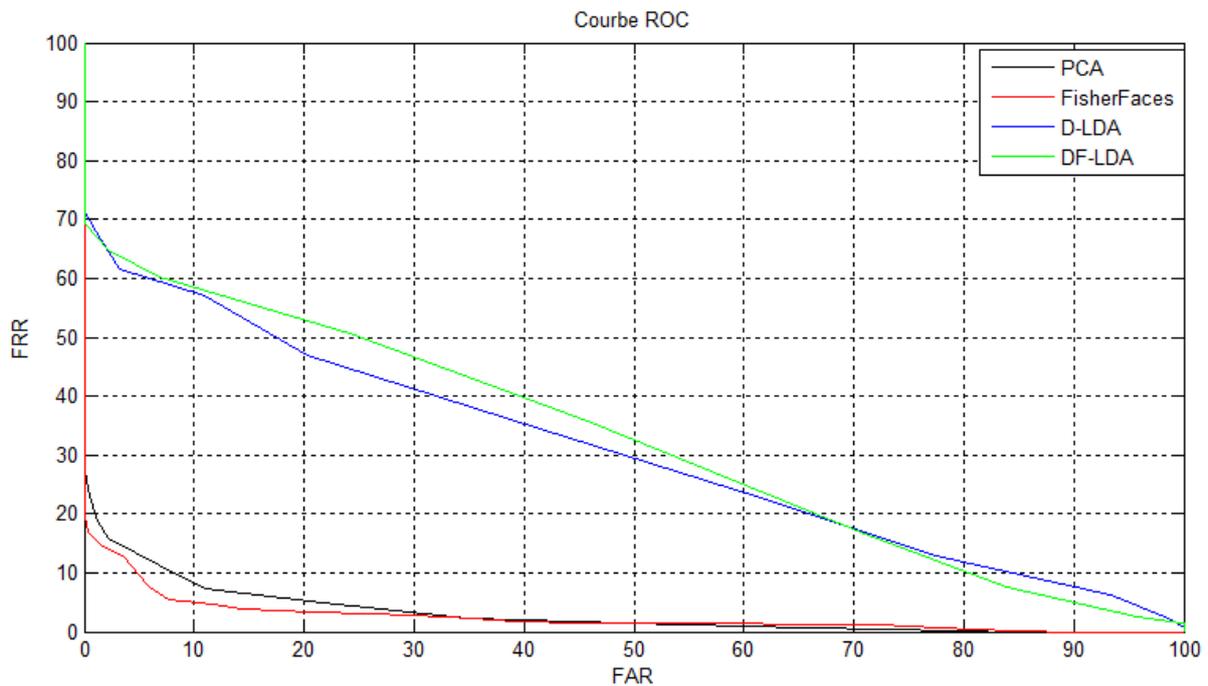


FIG IV.11 : Courbe ROC des différents algorithmes pour 15 vecteurs gardés pour Faces94.

La base de données UMIST :

L'ensemble d'apprentissage contient 15 personnes chacune d'elles possédant 8 images. L'ensemble de test est formé par les images restantes des personnes de l'ensemble d'apprentissage ainsi que les images des personnes non utilisées pour l'apprentissage. Après simulation, on obtient les graphes suivants :

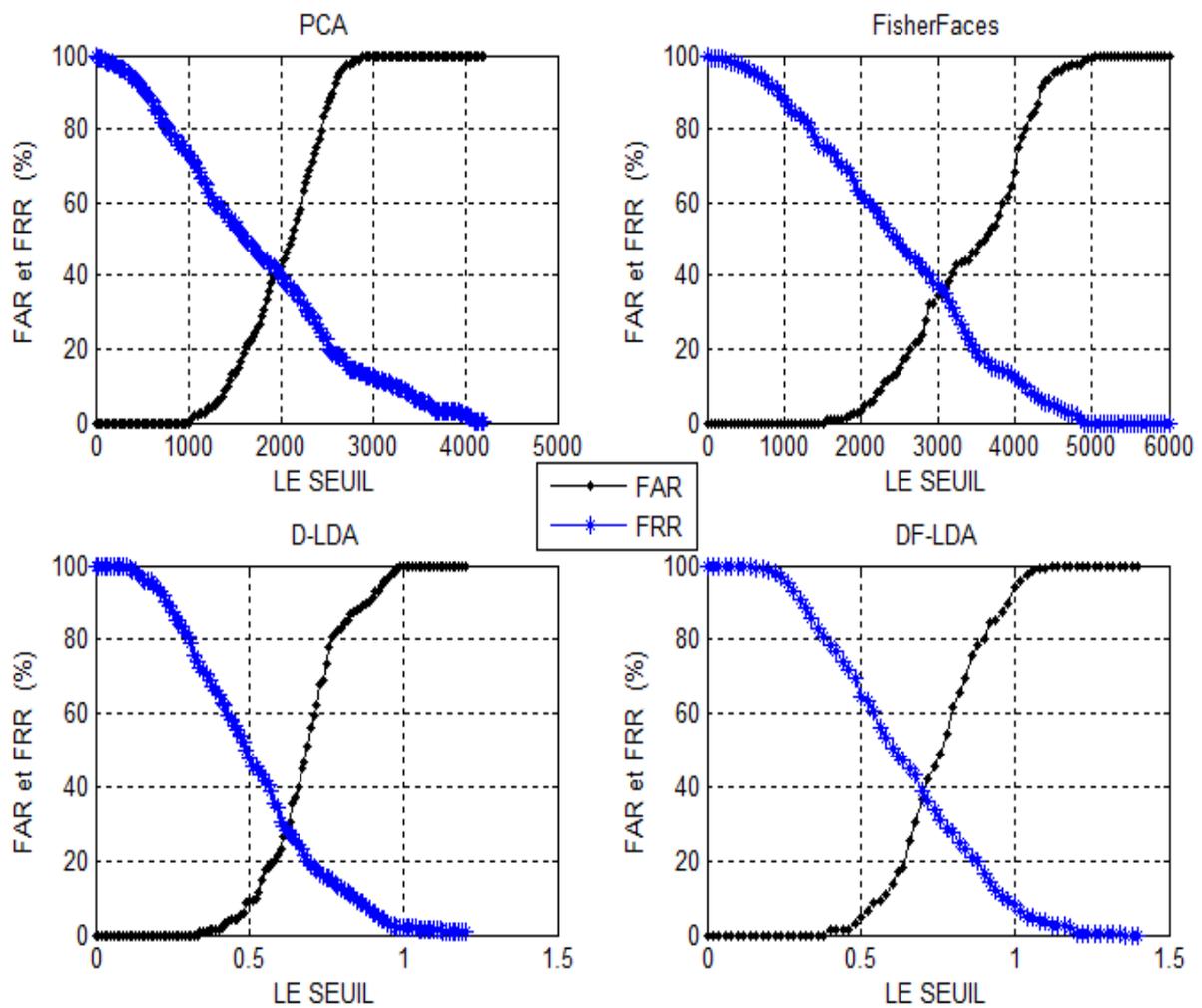


FIG IV.12 : FAR et FRR en fonction du seuil pour 12 vecteurs gardés pour l'UMIST.

A partir des graphes de la figure IV.12 on obtient le tableau suivant

Algorithme	Seuil	EER (%)
PCA	1963	40.72
FisherFaces	3074	35.6
D-LDA	0.5393	28.04
DF-LDA	0.7055	38.02

TAB IV.5 : Valeurs de ERR pour le seuil de décision optimal.

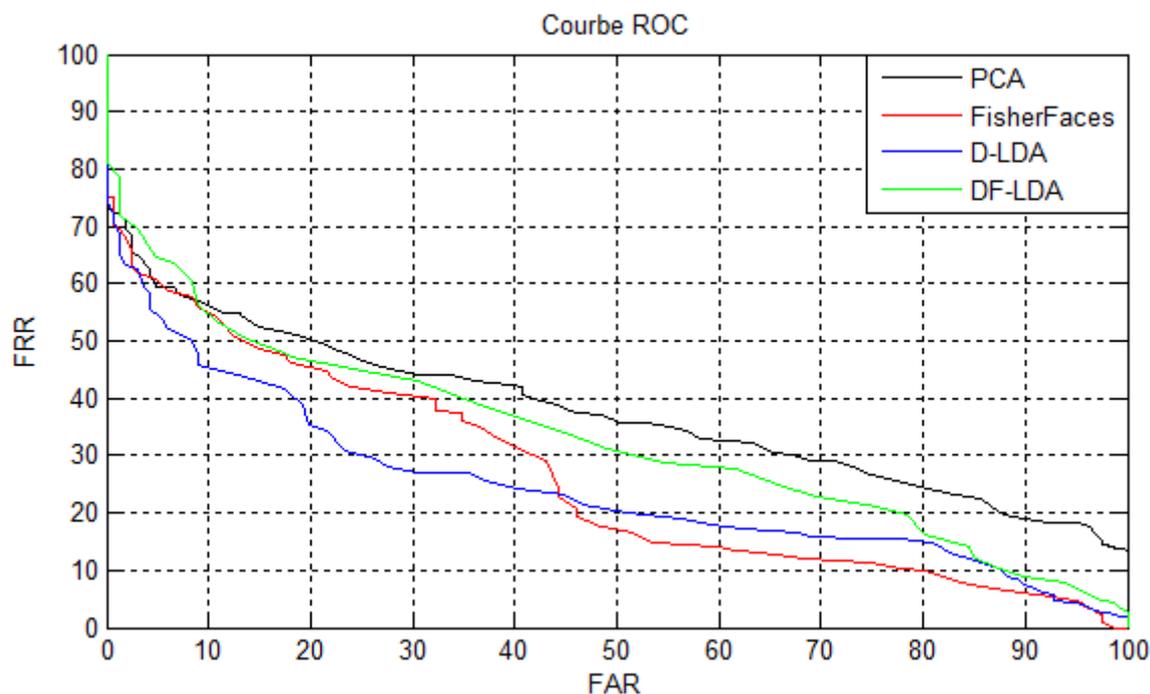


FIG IV.13 : Courbe ROC des différents algorithmes pour 12 vecteurs gardés pour l'UMIST.

Interprétation des résultats :

A partir des graphes et des tableaux obtenus, on constate que le ERR obtenu par les deux algorithmes PCA et FisherFaces est très inférieur à celui de la D-LDA et de la DF-LDA sur tout dans la base de données faces94. De plus, le ERR obtenu pour la base de données UMIST est grand. Cela est dû au fait que les variations de poses et d'éclaircissements sont considérables pour cette base de données.

b. Deuxième expérience :

Il est intéressant, aussi, de voir les variations d'ERR et de la courbes ROC en fonction du nombre des vecteurs formant la base de projection. Pour effectuer cela, on va utiliser les deux algorithmes PCA et FisherFaces sur la base de données ORL et Faces94 avec les mêmes ensembles d'apprentissage et de test vus précédemment.

La base de données ORL :

En appliquant la PCA on obtient le tableau suivant :

Nombre de vecteurs	2	3	4	5	7	10	15	20	25	30
Seuil	191.9	426.3	660.4	815.3	1112	1398	1658	1831	1953	2060
ERR	44.16	34.49	35.83	31.18	30.14	28.31	26.49	26	25.63	25.12

TAB IV.6 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la PCA.

La courbe ROC est la suivante :

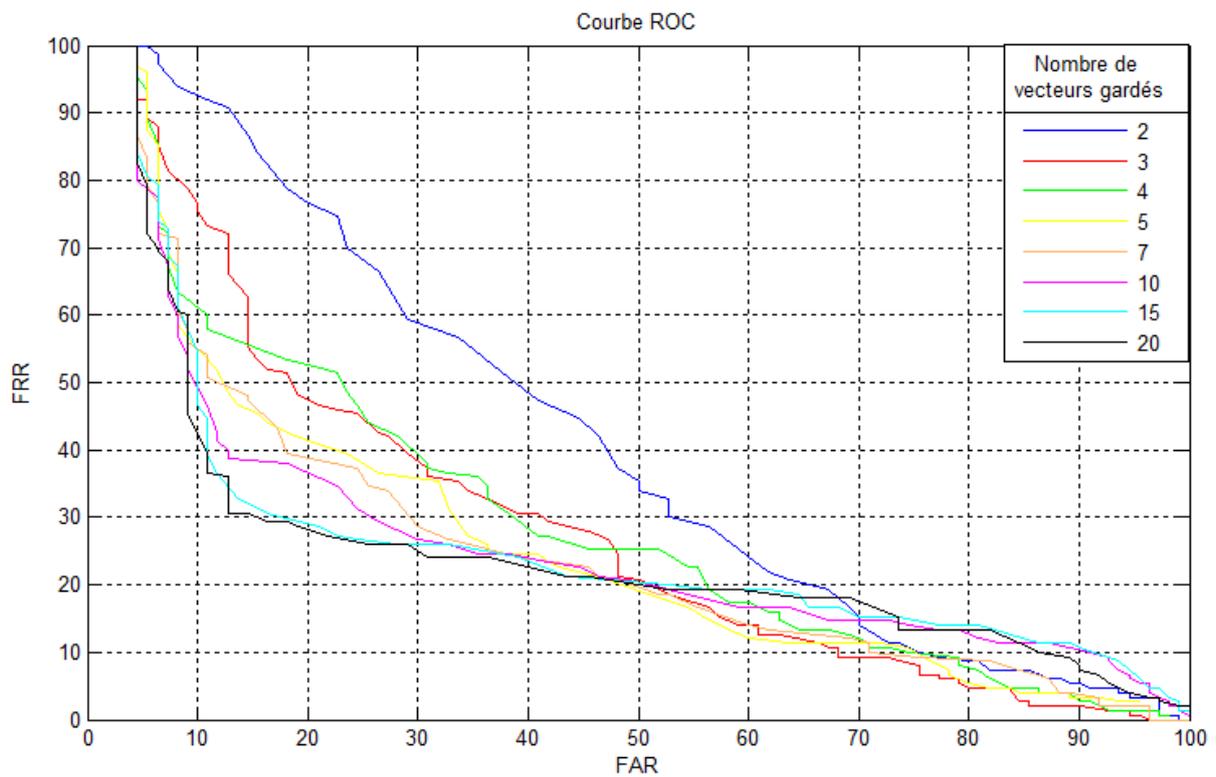


FIG IV.14 : Courbe ROC de la PCA en fonction de nombre de vecteurs gardés pour ORL.

En appliquant la FisherFaces on obtient :

Nombre de vecteurs	2	3	4	5	7	10	15	17	20	25	30
seuil	399.6	778.4	925	1213	1514	2121	2625	2843	3254	3805	4078
ERR	45.39	37.64	33.75	31.8	29.86	30.38	28.18	24.43	20.00	19.29	18.47

TAB IV.7 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la Fisherfaces.

La courbe ROC obtenue est donnée par la figure suivante :

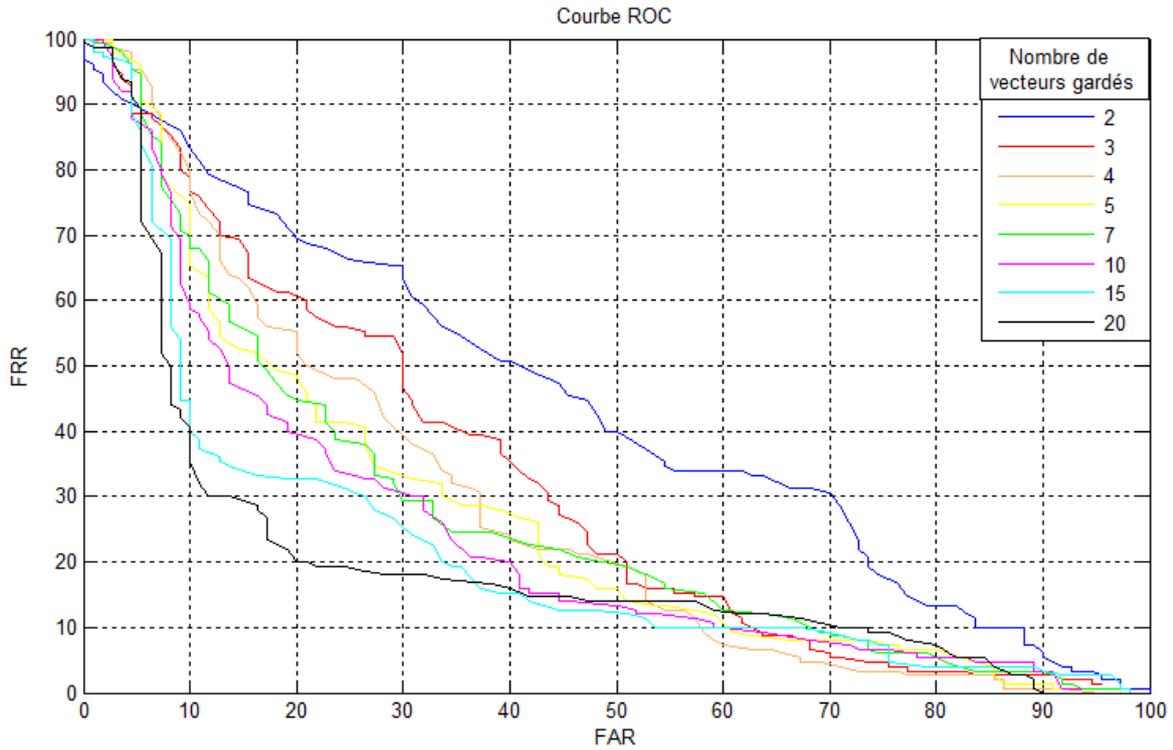


FIG IV.15 : Courbe ROC de la FisherFaces en fonction du nombre de vecteurs gardés pour ORL.

La base de données Faces94 :

En appliquant la PCA, on obtient la courbe ROC suivante:

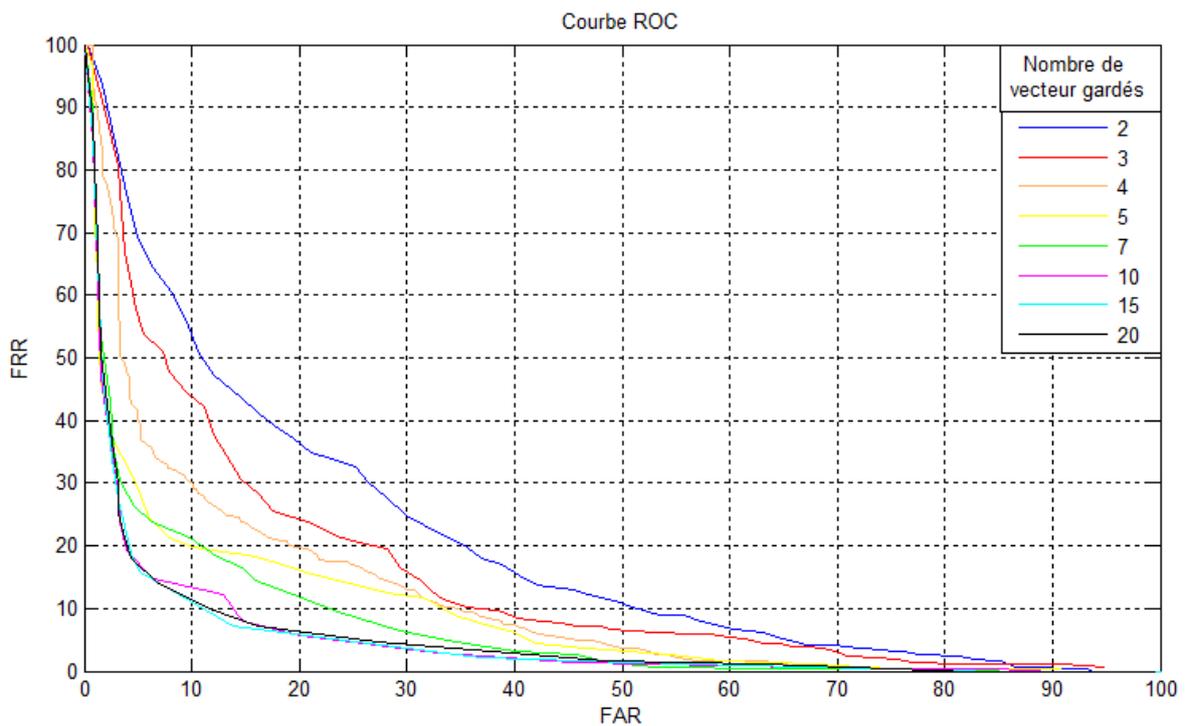


FIG IV.16 : Courbe ROC de la PCA en fonction du nombre de vecteurs gardés pour faces94.

Nombre de vecteurs	2	3	4	5	7	10	15	20	25	30
seuil	537.5	928.2	1481	1982	2505	3120	3673	3921	4147	4318
ERR	27.93	22.47	19.73	17.44	15.35	12.36	10.58	10.69	11.12	10.98

TAB IV.8 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la PCA.

FisherFaces :

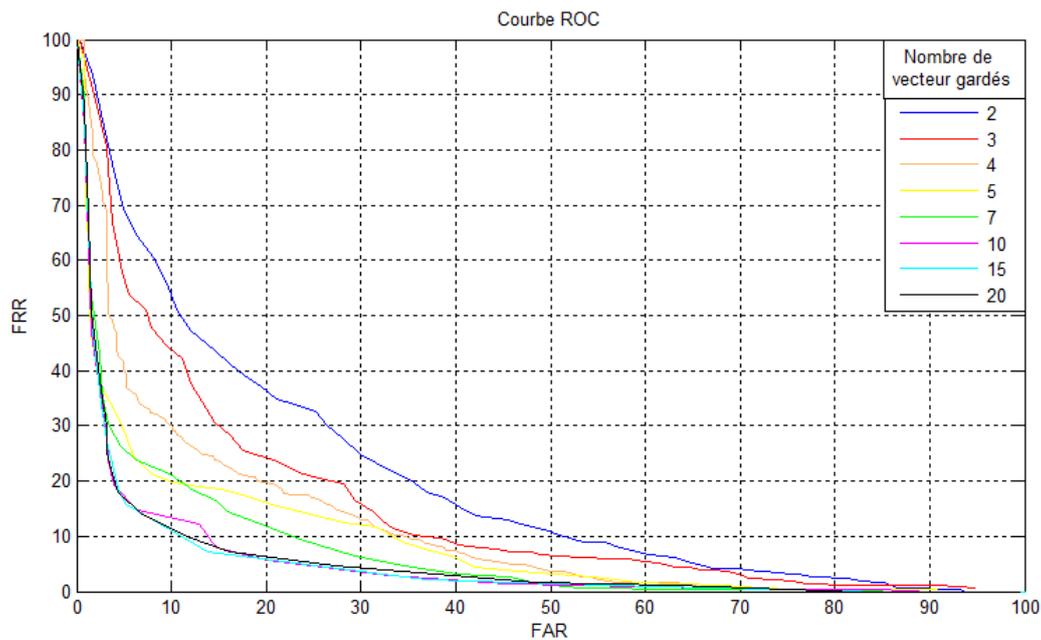


FIG IV.17 : Courbe ROC de la FisherFaces en fonction du nombre de vecteurs gardés pour faces94.

Nombre de vecteurs	2	3	4	5	7	10	15	20	25	30
seuil	668.8	1231	1822	2163	2937	3790	5264	6119	7054	7763
ERR	19.09	14.30	9.737	12.46	8.908	9.771	8.544	10.59	10.55	10.86

TAB IV.9 : Valeurs du seuil et de ERR en fonction du nombre de vecteurs gardés pour la FisherFaces.

Interprétation des résultats :

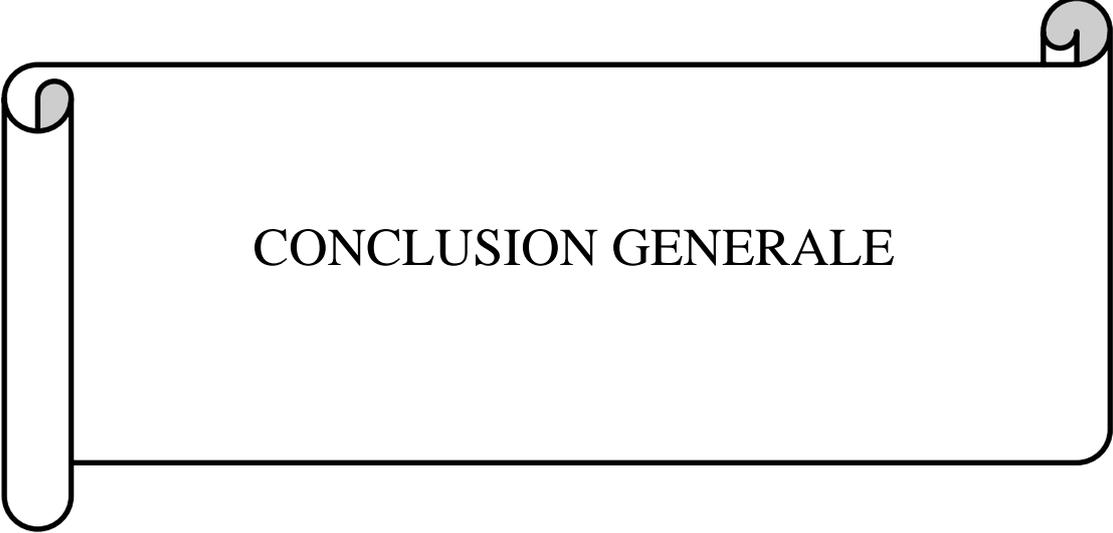
Le ERR est inversement proportionnel au nombre de vecteurs formant la base de projection. En plus, quand le nombre de ces vecteurs augmente, la courbe ROC tend à épouser le repère; ce qui est tout à fait évident car en augmentant le nombre de vecteurs une information supplémentaire sera ajoutée ce qui permet une bonne discrimination des classes. A partir d'un certain nombre de vecteurs, le ERR devient relativement constant. Le seuil de

décision optimal varie lui aussi en fonction de nombre de vecteurs constituant la base de projection.

4. Conclusion :

Après étude des performances des différents algorithmes, on a constaté que les algorithmes D-LDA et DF-LDA présentent des taux d'identifications correctes élevés. Cependant le ERR obtenu n'est pas faible, ce qui engendre une forte acceptation d'imposteurs et rejection de clients. La PCA, quant à elle, offre des taux d'identifications correctes relativement petit surtout quand la base de données présente une forte variation d'éclairage, de poses et d'expressions faciales; cependant le ERR est assez faible ce qui lui permet de rejeter les imposteurs et d'accepter les clients. Pour ce qui est de la FisherFaces, sa performance dépend entièrement du nombre de vecteurs gardés lors de la réduction de dimension avec la PCA. Quand ce nombre est faible cela signifie qu'une grande quantité d'information utile est éliminée et à cause de ça le taux d'identifications correctes sera faible. Dans le cas où la PCA réduit la dimension tout en gardant un nombre suffisant de vecteurs propres de telle manière que la plupart de l'information utile pour la discrimination soit gardée, la FisherFaces présentera alors un taux d'indentification correcte élevé et un EER faible ce qui lui permet d'être très efficace et performante.

Afin d'améliorer le taux d'identification correcte et de réduire le ERR il suffit d'augmenter le nombre de vecteurs formant la base de projection. Cependant, cette solution est très couteuse en terme de temps de reconnaissance car ce dernier est proportionnel au nombre de vecteurs gardés. Pour cette raison, un bon compromis performance/temps de reconnaissance est toujours recherché dans un système de reconnaissance de visages.



CONCLUSION GENERALE

CONCLUSION GENERALE

La biométrie est une technologie en plein essor. Elle est de plus en plus utilisée dans les applications en relation avec la sécurité, vu les avantages qu'elle offre contrairement aux anciennes méthodes.

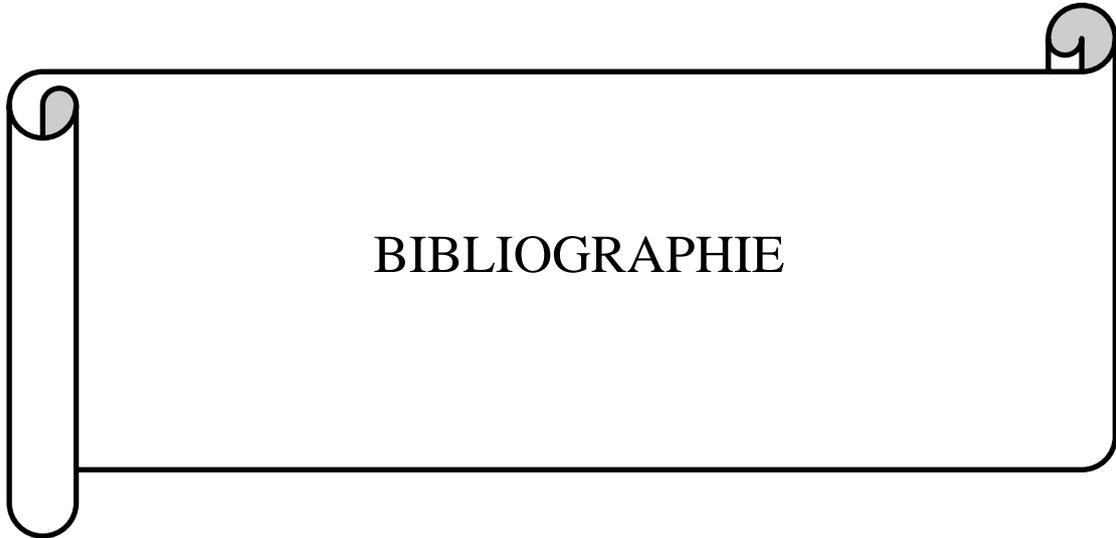
Après avoir introduit les concepts généraux en biométrie et présenté les systèmes biométriques ainsi que les différentes modalités qui existent, on s'est intéressé aux différentes techniques et algorithmes utilisés en reconnaissance des visages en présentant un état de l'art sur cette modalité. Par la suite, on a étudié en détail l'analyse en composantes principales et l'analyse discriminante linéaire et on a exposé les résultats obtenus après simulation des différents algorithmes.

L'objectif principal du travail effectué est la comparaison entre les différents algorithmes (PCA, FisherFaces, D-LDA et DF-LDA) en se basant sur leurs performances dans la reconnaissance de visages. Ces performances sont traduites par le taux d'identifications correctes, le taux d'égale erreur (ERR) et le temps de reconnaissance qui est directement lié au nombre de vecteurs formant la base de projection des différents algorithmes.

Après comparaison, on a constaté que le meilleur algorithme, parmi ceux utilisés, est le FisherFaces dans lequel un nombre suffisant de visages propres est gardé après l'étape de la PCA. Cet algorithme assure un taux d'identifications correctes acceptable même si les images de la base de données présentent une variation d'éclairage, de poses ou d'expressions faciales. Le taux d'égale erreur quant à lui est faible pour les bases de données à faibles variations de paramètres (éclairage, poses,...), cependant il s'avère important dans l'autre cas.

Les deux méthodes globales utilisées dans ce travail font partie des approches les plus courantes en reconnaissance de visages grâce à leur simplicité et leur efficacité. Cependant il serait intéressant de combiner ces deux méthodes avec d'autres techniques de reconnaissance de visages telles que les HMM (modèles cachés de Markov) et les SVM (Machines à vecteurs de support) ou carrément les utiliser dans des systèmes multimodaux afin d'augmenter les performances.

Dans le but d'améliorer le temps de reconnaissance tout en gardant les mêmes performances, il est souhaitable de penser à une architecture matérielle travaillant en parallèle telle que l'utilisation des FPGA ou des DSP.



BIBLIOGRAPHIE

BIBLIOGRAPHIE

- [1] R. JOURANI, Reconnaissance de visages, thèse d'études supérieures approfondies, faculté des sciences de rabat, 2006.
- [2] N. MORIZET, Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris, thèse de doctorat, école nationale supérieure des télécommunications de Paris, 2009.
- [3] A. Ross et A. K. Jain, Biometrics: When Identity Matters, Advances in Biometric Person Authentication, 5th Chinese Conference on Biometric Recognition, Sinobiometrics 2004, LNCS 3338, pp. 1-2. 2004.
- [4] F. PERRONNIN et J. L. DUGELAY, Introduction à la Biométrie Authentification des Individus par Traitement Audio-Vidéo, TS. Traitement du signal, 2002, vol. 19, no4, pp. 253-265.
- [5] E. BOUTELLAA, Système biométrique de vérification de signatures manuscrites en ligne, mémoire de magister, école supérieure d'informatique d'Alger (ex INI).
- [6] N. HAMITOUCHE et Z. SALMI, Système d'identification biométrique de personnes par reconnaissance de l'iris, mémoire de fin d'études pour l'obtention du diplôme d'ingénieur, Ecole nationale Supérieure d'Informatique, 2009.
- [7] journal du net, <http://www.journaldunet.com/solutions/0611/061127-biometrie/6.shtml>
- [8] journal du net, <http://www.journaldunet.com/solutions/0611/061127-biometrie/8.shtml>
- [9] M. A. KHEBRI et M. BIREM, Implémentation en temps réel sur FPGA d'une combinaison séquentielle des algorithmes PCA et WPD pour la reconnaissance de visages, Ecole Nationale Polytechnique d'Alger, 2010.
- [10] M. ZAIED, Etude des réseaux d'ondelettes Bêta : Application à la reconnaissance de visages, thèse de doctorat, Ecole Nationale d'Ingénieurs de Sfax, 2008.
- [11] Stan Z. Li and Anil K. Jain, Handbook of face recognition, December 2004, pp. 1-12.
- [12] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, Face Recognition: A Literature Survey, ACM Computing Surveys, Vol. 35, No. 4, December 2003, pp. 399-458
- [13] M. Turk & A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, Volume 3, Numéro 1 pp. 71-86

BIBLIOGRAPHIE

- [14] A. R. Omondi et J. C. RAJAPAKSE, FPGA implementations of neural networks, Springer. 2006
- [15] M. BOUDJEMA et N. BELHAOUAS, Commande d'une poursuite du point de puissance maximum (MPPT) par les Réseaux de Neurones et implémentation sur FPGA et DSP, mémoire pour l'obtention du diplôme d'ingénieur, Ecole Nationale Polytechnique, 2007.
- [16] http://labo.algo.free.fr/neuro/reseau_de_neurones_artificiel.html
- [17] http://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support
- [18] R. RIOUX, T. SIMON, Redondance de classes pour la reconnaissance de visages par SVM, ELMAR, 2008. 50th International Symposium, Sept. 2008, pp.117 - 120
- [19] P. Comon, Analyse en composantes indépendantes et identification aveugle, Traitement du Signal, pp. 435-450, Vol 7, No 5 spécial Non linéaire non gaussien, 1990.
- [20] C. Havran, L. Hupet, J. Czyz, J. Lee, L. Vandendorpe, M. Verleysen, Independent Component Analysis for face authentication, KES'2002 proceedings - Knowledge-Based Intelligent Information and Engineering Systems, Crema (Italy), 16-18 September 2002, IOS Press, Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies E. Damiani, R.J. Howlett, L.C. Jain, N. Ichalkaranje eds., ISBN 1-58603280-1, pp. 1207-1211.
- [21] M. S. Bartlett, J. R. Movellan et T. J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 13, NO. 6, NOVEMBER 2002, pp. 1450-1464.
- [22] N. MORIZET, T. EA, F. ROSSANT, F. AMIEL, A. AMARA, Revue des Algorithmes PCA, LDA et EBGm utilisés en reconnaissance 2D du visage pour la biométrie, Conférence MajecSTIC'06, 24 Novembre 2006, Lorient, France.
- [23] L. Wiskott, J. M. Fellous, N. Kruger et C. Malsburg, Face Recognition by Elastic Bunch Graph Matching, In Intelligent Biometric Techniques in Fingerprint and Face Recognition, eds. L.C. Jain et al., publ. CRC Press, ISBN 0-8493-2055-0, Chapter 11, pp. 355-396, (1999)
- [24] J. K. Fung, Face recognition by eigenface and elastic bunch graph matching, Premier rapport de projet de recherche de thèse PHD, université de Hong Kong (Chine).

BIBLIOGRAPHIE

- [25] F. S. Samaria, Face Recognition Using Hidden Markov Models, thèse PHD, université de Cambridge (Angleterre), 1994.
- [26] L. Vincent & Y. Ainane Modèles de Markov cachés pour Reconnaissance de la parole, projet multimédia, faculté polytechnique de Mons, 2005.
- [27] A. Pentland, B. Moghaddam et T. Starner. Viewbased and modular eigenspaces for face recognition. In Proceedings of the 1994 Conference on Computer Vision and Pattern Recognition, pages 84–91, Seattle, WA, 1994. IEEE Computer Society.
- [28] A. Mellakh, Reconnaissance des visages en conditions dégradées, Thèse de Doctorat, l'Institut National des Télécommunications de Paris, Avril 2009.
- [29] B. A. Draper, K. Baek, M. S. Bartlett, J. R. Beveridge, Recognizing Faces with PCA and ICA, Computer Vision and Image Understanding - Special issue on Face recognition, Volume 91 Issue 1-2, July 2003, pp. 115 – 137.
- [30] T. J. Chin and D. Suter, A Study of the Eigenface Approach for Face Recognition, Technical Report MECSE-6-2004, Computer Science Publication, Monash University, 2004.
- [31] Jacques Baillargeon L'analyse en composantes principales, janvier 2003.
- [32] K. Liu, Y.Q. Cheng, J.Y. Yang, and X. Liu, An efficient algorithm for foley-sammon optimal set of discriminant vectors by algebraic method, Int. J. Pattern Recog. Artif. Intell., vol. 6, pp. 817–829, 1992.
- [33] H. Yu and J. Yang, A direct LDA algorithm for high-dimensional data with application to face recognition, Pattern Recognition, vol. 34, pp. 2067–2070, 2001.
- [34] L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, A new lda-based face recognition system which can solve the small sample size problem, Pattern Recognition, 33(10):1713-1726, Oct 2000.
- [35] R. Lotlikar and R. Kothari, “Fractional-step dimensionality reduction”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 6, pp. 623–627, 2000
- [36] Juwei Lu, K.N. Plataniotis, and A.N. Venetsanopoulos Face Recognition Using LDA Based Algorithms, IEEE TRANSACTIONS ON NEURAL NETWORKS, ACCEPTED IN MAY 2002.
- [37] ORL face database, website: <http://www.cam-orl.co.uk/facedatabase.html>, AT&T Laboratories Cambridge, Angleterre.

BIBLIOGRAPHIE

[38] Facses94 face database, <http://cswww.essex.ac.uk/mv/allfaces/faces94.html>,
Université de l'Essex, Angleterre.

[39] UMIST face database, <http://www.sheffield.ac.uk/eee/research/iel/research/face>,
L'université de Sheffield, Angleterre.