

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Télécommunication

Réf:.....

Mémoire de Fin d'Etudes
En vue d'obtention de diplôme:
MASTER

Thème :

**Fusion 2D et 3D pour la
reconnaissance de visage**

Proposée et Dirigée par:

Mm: BELAHCENE. B. Mébarka

Présenté par :

BEN HAMZA Nadjib

Promotion : 2011/2012

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Télécommunication

Mémoire de Fin d'Etudes
En vue d'obtention de diplôme:
MASTER

Thème :

**Fusion 2D et 3D pour la
reconnaissance de visage**

Présenté par :

BEN HAMZA Nadjib

Nom et prénom

BELAHCENE

Mébarka

Avis favorable de l'encadreur

Signature

Avis favorable du président du Jury

Nom et prénom

.....

Signature

.....

Cachet et signature



Résumé du mémoire de fin d'étude

En vue d'obtention de diplôme:

MASTER

Option : Télécommunication

Thème :

Fusion 2D et 3D pour la reconnaissance de visage

Proposé par : **BENATIA BELAHCENE Mébarka**

Dirigé par : **BENATIA BELAHCENE Mébarka**

RESUME

Notre objectif est l'étude et l'évaluation par des expériences sur la fusion de deux systèmes biométriques (des visages 3D et des visages 2D). Ainsi, notre but est la réalisation d'un système multi biométrique de reconnaissance de visages basé sur la fusion par combinaison de scores. Notre approche est fondée sur l'apport couleur et l'apport profondeur du visage 3D. Pour cela l'étape de prétraitement est primordiale particulièrement dans le visage 3D. Nous utilisons les cartes de profondeur en NG pour mesurer la profondeur et nous réalisons la projection de l'image 3D couleur sur l'espace des couleurs. Ces deux modalités sont réduites par le filtre de Gabor puis par la transformation ACP et les scores obtenus sont fusionnés. Une classification est effectuée et une décision par la suite est prise.

Nous avons utilisé plusieurs méthodes de combinaison des scores (moyenne, max et min) utilisés à la suite de l'ACP pour essayer d'exploiter la complémentarité entre ces méthodes.

Les résultats des expériences sur la base de données CASIA 3D, nous ont montré le degré d'amélioration des performances de cette fusion et comment choisir les paramètres pour avoir les meilleurs résultats.

Mots clés : SRV, identification, fusion des scores, ACP, classification, Décision ,visage 2D et 3D .

Remerciements

Je remercie tout d'abord « ALLAH » le Tout Puissant, de nous avoir donné le Courage et la Patience, et de nous guider dans le chemin de la Raison et de la Science.

Si on tente de remercier notre encadreur Mme : BELAHCENE BENATIA MEBARKA.

Rien n'exprimera nos gratitude pour vos conseils et votre aide je ne saurai vous remercier assez..... MERCI infiniment.

Je remercie aussi ALISMAIL Mohamed Raouf pour son aide et sa collaboration.

Je tiens à remercier également tous mes amis pour leurs soutiens.

Je voudrai transmettre aux membres du jury, mes hautes considérations

Dédicaces

*Je dédie ce travail aux gens qui m'ont soutenu et aidé à
préparer ce travail:*

à mes très chers parents qui ont tant prié pour ma réussite,

Ma famille,

Mes amis,

Mes collègues,

Mes enseignants,

et ...

à toute personne qui s'intéresse à la Science.

Nadjib BEN HAMZA

Sommaire

Introduction général	01
Chapitre I : reconnaissance de visage 2D et 3D	
I.1 Introduction.....	04
I.2 Processus de reconnaissance des visages.....	04
I.2.1. Images Fixes:.....	04
I.2.2. Images 3D.....	05
I.3 Reconnaissance e de visage.....	05
I.3.1 Identification.....	05
I.3.2 Vérification.....	06
I.4 Systèmes de reconnaissance de visages.....	07
I.4.1 L'Acquisition de l'image.....	08
I.4.2 Les prétraitements.....	08
I.4.3 L'extraction de paramètres.....	09
I.4.4 La classification (Modélisation).....	09
I.4.5 L'apprentissage.....	09
I.4.6 La décision.....	09
I.5 La reconnaissance automatique de visages.....	10
I.6 Reconnaissance 3D de visage.....	11
I.6.1 Multimodal 2D +3 D reconnaissance de visage.....	14
I.7 Techniques de reconnaissance de visage.....	15
I.7.1 Techniques 2D de reconnaissance de visage.....	16
I.7.1.1 Les Approche globales.....	18

I.7.1.2 Approche locale (Géométrie).....	18
I.7.1.3 Les approches hybrides.....	19
I.7.2 Techniques 3D de reconnaissance de visages.....	19
I.7.2.1 Systèmes d'acquisition 3D.....	19
I.7.2.2 Approches modèle.....	21
I.7.2.3 Approches 3D.....	21
1. Approches surface.....	21
2. Approches holistiques 3D.....	25
I.7.2. 4 Approche géométrique ou locale 3D.....	24
I.7.2.5 Approche 2D+3D.....	25
Conclusion	25
 Chapitre II :fusion des faces 2D et 3D	
II.1 introduction.....	28
II.2 C'est quoi la fusion ?.....	28
II.3 Les type de fusion.....	28
II.4 Les différents niveaux de fusion.....	30
II.4.1 Fusion avant le matching.....	30
II.4.1.1 Niveau capteur (Sensor level).....	31
II.4.1.2 Niveau caractéristiques (Feature level).....	32
II.4.2 Fusion après le matching.....	32
II.4.2.1 Niveau décision (Decision level).....	32
II.4.2.2 Niveau rang (Rank level).....	33
II.4.2.3 Niveau score (Score level).....	33

II.5 Etat de l'art de la fusion.....	33
Conclusion.....	40

Chapitre III : Etude du modèle de fusion 2D et 3D pour SRV.

III.1 Introduction.....	42
III.2 Présentation du modèle	42
III.2.1 Prétraitement d'image.....	44
III.2.1.1 Découpage de l'image.....	44
III.2.1.2 Conversion de l'image	44
III.2.1.3 Filtrage.....	44
III.2.2 Réduction d'espace	47
III.2.1 Principe de PCA.....	47
III.2.3 Classification	48
III.2.4.1 Types de classification.....	48
III.2.4 Calcul de métriques.....	50
1) Distances Euclidiennes.....	51
2) Distances dans l'Espace de Mahalanobis.....	51
III.2.5 Normalisation.....	53
III.2.6 Fusion des scores.....	54
III.2.6 Décision.....	55
III. 3 Conclusion.....	56

CHAPITRE IV : Conception Et Résultat

IV.1 Introduction.....	58
IV. 2 La base de données CASIA 3D face VI.....	58
IV.3 Architecture du système de conception.....	60
IV.3.1 Acquisition d'image.....	61
IV.3.2 Prétraitement de visage	61
IV.3.3 Prétraitement d'image	62
IV.4 Réduction d'espace.....	62
IV.4.1 L'analyse en composantes principales(PCA).....	62
IV.5 Mesure de similarité.....	63
IV.6 Normalisation.....	64
IV.7 La fusion.....	64
IV.8 Décision	65
IV.9 Résultats.....	66
IV.9.1 Extraction des paramètres	66
IV.11Conclusion.....	72
Conclusion générale.....	73

ANNEXE

Annexe A

A.1 Classification de machines à vecteur.....	74
A.1.1 Historique	74
A.1.2. Principe de la technique SVM	74

Annexe B

B.1 Réseaux de neurones artificiels.....	75
Bibliographie.....	76

Liste de figure

Chapitre I

Fig. I.1: Schéma d'identification du visage.....	6
Fig. I.2 : Vérification du visage.....	7
Fig.I.3 : Processus d'un système de reconnaissance de visages.....	8
Fig.I.4 : Fonctionnement d'un système de reconnaissance automatique de visages.....	10
Fig.I.5: 3D scanner and original face data process.....	13
Fig.I.6 : Le visage 3D retrimming et le résultat final.....	14
Fig.I.7 : Taxonomie des approches de reconnaissance de visages 2D, 3D.....	16
Fig.I.8 : Les techniques de reconnaissance de visage.....	17
Fig.I.9 : (a) visage normalisé; (b) 2D image; (c) 3D image.....	20

Chapitre II

Fig. II.1 : Sources de différents types de fusion de traits biométriques.....	29
Fig. II.2 : Illustration de la fusion pondérée d'images par l'algorithme PSO.....	31
Fig. II. 3 : Chaîne de traitement complète avec possible fusion des scores.....	34
Fig. II. 4 : Cadre du projet de DMQI-LBP algorithme de reconnaissance de visage.....	36
Fig. II. 5: Système de fusion et démarche pour l'identification	37
Fig. II. 6 : Le premier bloc dans la Fig. est la main et le traqueur visage.....	38
.Fig. II. 7 : Schéma de la reconnaissance biométrique multimodal.....	39

Chapitre III

Fig.III.1 : schéma du modèle proposé	43
Fig.III.2 : Orientation de l'ondelette.....	45
Fig.III.3 : Fréquence centrale de l'ondelette.....	45
Fig.III.4 : Phase de l'ondelette.....	45
Fig.III.5 : Support temporel de l'ondelette.....	46
Fig.III.6 : Forme de l'enveloppe gaussienne.....	46
Fig.III.7 : Parties réelles (a) et imaginaires (b) d'ondelettes de Gabor à 4 niveaux de résolution et selon 4 orientations	47
Fig. III.8 : : Les méthodes de classification	53

Chapitre IV

FigIV.1 : variations d'éclairage de la base de données CASIA visage 3D.....	59
FigIV.2 Pose des variations de la base de données CASIA visage 3D.....	59
FigIV.3 : variations d'expression de la base de données CASIA visage 3D.....	59
Fig. IV.4 : Architecture structurelle du système de conception.....	60
FigIV.5 : Capturer des scènes de la base de données CASIA visage 3D.....	61
Fig. IV. 6 la conversion de base de donné (A)image 3D,(B)image de profondeur,(C)image RGB.....	61
Fig. IV.7 Schéma de prétraitement des images de visage 3D.....	62
Fig. IV.8 Fusion des distances.....	65
Fig. IV 9 . Taux de réussite (TR) en fonction de N(ACP)	68
Fig. IV 10 Taux d'égale erreur (TEE) en fonction de N(ACP).....	68

Fig. IV 11 Taux de réussite (TR) norme L2.....	69
Fig. IV 12 : Taux d'égale erreur(TEE) norme L2.....	70
Fig. IV 13 : Taux de réussite(TR) Cov L3.....	71
Fig. IV 14 : Taux d'égale erreur(TEE) Cov L3.....	71

Liste de tableau

Tableaux II.1 : état de l'art de fusion.....	39
TableauIV.1 : Taux d'erreurs de norme L1.....	67
Tableau IV.2 Taux d'erreurs pour la norme L2.....	69
Tableau IV.3 : Taux d'erreurs et de réussite pour la covariance L3.....	70
Tableau IV.4 : Comparaison de TR pour L1, L2 et COV.....	71

Liste des abréviations

CASIA: Institute of Automation Chinese Academy of Sciences Iris Image Database

RGB : Red Green Blue

SRV : Système de reconnaissance des visages

EFM : Enhanced Fisher linear discriminant Model

TEE : Taux d'égalité Erreur

ACP: Analyse en Composantes Principales

DLA: Dynamic Link Architecture

SVM: Machines à Vecteurs de Support

TFA : Taux de fausse acceptation

TFR : Taux de faux rejet

TR: Taux de reconnaissance

BDD : Base de données.

DLA: Dynamic Link Architecture

Résumé

Notre objectif est l'étude et l'évaluation par des expériences sur la fusion de deux systèmes biométriques (la reconnaissance des visages 3D, et la reconnaissance des visages 2D), Ainsi, notre but est la réalisation d'un système biométrique de reconnaissance de visages 2D et 3D (identification et vérification) basé sur la fusion de scores par ondelettes associées à plusieurs méthodes de combinaison.

Nous avons utilisé plusieurs méthodes de combinaison des scores (moyenne, max et min) utilisés à la suite de l'ACP pour essayer d'exploiter la complémentarité entre ces méthodes.

Les résultats des expériences sur deux bases de données, nous ont montré le degré d'amélioration des performances de cette fusion et comment choisir les paramètres pour avoir les meilleurs résultats.

Mots clés:

Reconnaissance de visage, face 2D, face 3D, Identification, Authentification, Fusion des scores, ACP, Décision.

Abstract

Our purpose was the study and the evaluation with experiments the fusion of two biometric systems (the 3D face recognition and the 2D face recognition), and our goal was the realization of a biometric system of 3D and 2D face recognition (identification and verification) based on score fusion with wavelet.

We used a several methods of score combination (mean, max and min) obtained from the ACP to try to exploit the complementarity between these methods.

The results of the experiences on the two used databases, gave us the degree of the amelioration of the performances of this fusion and how to choose the parameters to get the best results.

Key words:

Face recognition, 2D face, 3D face, Identification, Authentication, Fusion scores, PCA, Decision.

INTRODUCTION GENERALE

Durant les vingt dernières années, la reconnaissance automatique de visages est devenue un enjeu primordial, notamment dans les domaines de la sécurité et de l'indexation de documents multimédia. Malgré la multitude de techniques proposées et les progrès réalisés, le problème reste non résolu, comme le montre la récente évaluation menée par le *National Institute of Standards and Technology* (NIST).

Plusieurs méthodes ont été développées pour la reconnaissance bidimensionnelle des visages. Cependant, elles présentent un certain nombre de difficultés qui proviennent essentiellement de la forte variabilité des données à classer dues à des changements dans l'expression faciale, dans la pose de la tête, dans les conditions d'illumination, à des occultations partielles et dans le fait que les images de visages constituent des données de très grandes dimensions.

Ces dernières années, on parle de plus en plus des techniques de reconnaissance de visages 3D comme une solution alternative pour résoudre les problèmes cités ci-dessus. En effet, la richesse de l'information fournie par les mesures 3D permet de reconstruire la forme tridimensionnelle du visage. Ce type de représentation du visage est invariant aux changements d'illumination et de pose.

Dans ce mémoire nous présentons un système de reconnaissance faciale basé sur la fusion de score extraites à partir des images 3D et 2D, nous préconisons d'implémenter une technique pour l'extraction de caractéristiques pour l'analyse biométrique tridimensionnelle du visage. Pour cela nous adoptons une méthode statistique bien connue de la reconnaissance faciale bidimensionnelle, l'analyse en composante principale ou ACP. Cela réduit grandement les tailles des images, maintient un temps de calcul raisonnable. Nous allons ensuite fusionner les scores extraites sous formes de distances, ces derniers sont calculer avec plusieurs méthodes de mesure de distance.

Notre étude se présente sous forme de quatre parties essentielles :

Dans le premier chapitre, nous avons cité les différentes méthodes de reconnaissance de visage bidimensionnelle et tridimensionnelle et nous avons essayé

de présenter les différentes étapes du processus de reconnaissance ainsi que les difficultés inhérentes à la reconnaissance de visages.

Le deuxième chapitre commence par un aperçu sur la multi modalité dans les systèmes biométriques. Et décrit brièvement les différents niveaux de la fusion biométrique et un peu plus en détail les méthodes de fusion et de normalisation des scores.

Le troisième chapitre contient l'étude théorique de toutes les méthodes appliquées ou utilisées dans la conception du système bimodale basé sur la fusion de score.

Le quatrième chapitre est composé de deux parties, la première partie présente les étapes qui ont permis la réalisation du système qui se base sur les méthodes décrites dans le troisième chapitre. Les tests et les résultats obtenus sur le système final sont présentés dans la deuxième partie de ce chapitre.

Nous terminerons ce mémoire par une conclusion générale et des perspectives.

I.1 Introduction

Dans le spectre de la biométrie, la reconnaissance de visages revêt un caractère particulier du fait du grand nombre de travaux qui lui sont consacrés.

En effet capturer une image d'un visage, en particulier à travers une caméra 2D, est simple et non invasif. C'est donc une modalité biométrique facilement tolérée par les utilisateurs, et ceci d'autant plus que les ordinateurs sont équipés de plus en plus souvent de « webcam ». Mais les performances de la reconnaissance faciale sont toujours bien en deçà de ce que l'on pourrait espérer pour de telles applications. L'augmentation des taux de reconnaissance, la diminution des reconnaissances à tort et l'accélération des temps de réponse lors de la recherche dans les grandes bases de données biométriques sont les défis auxquels doivent faire face les algorithmes actuellement développés.

I.2 Processus de reconnaissance des visages

I.2.1. Images Fixes

La plus part des algorithmes de reconnaissance automatique de visages portent sur la classification d'images fixes 2D notamment en niveau de gris. Les algorithmes prenant en compte la couleur ne sont, pour la plus part, que des généralisations à trois canaux des techniques utilisées en niveaux de gris [M.Visani05].

On dispose d'une base de connaissance contenant un ensemble d'images de visages de personnes connues, c'est-à-dire les personnes que le système est censé reconnaître. Cette base de connaissance peut contenir plusieurs images d'une même personne sous des conditions de prises de vue différentes. Chaque image est étiquetée par son identité associée. Lorsqu'un visage requête (image d'un visage à reconnaître) est présenté au système, celui-ci va chercher à lui assigner une identité parmi celles contenues dans la base de connaissance. Le système de reconnaissance est basé sur un modèle généralement construit à partir d'une base d'apprentissage contenant un ensemble d'images de visages par le biais d'un algorithme d'apprentissage.

Dans certains cas, la base d'apprentissage et la base de connaissance sont confondues. Le modèle est alors spécifiquement conçu pour les visages qu'il vise à reconnaître. Dans d'autres cas, au contraire, les deux bases sont distinctes soit à cause

de l'application à réaliser soit à cause des spécificités de la technique utilisée [M.Visani05].

I.2.2. Images 3D

Très récemment, les avancées dans le domaine de l'acquisition de données tridimensionnelles (acquisition par scanners laser) ont permis l'émergence de technologies de reconnaissance 3D [M.Visani05]. Deux grandes familles de méthodes sont apparues ces dernières années :

- 1) Celles basées sur la correspondance d'objets 3D directement (on dispose d'images 3D aussi bien dans la base d'apprentissage que pour les visages requêtes) [Chan05]
- 2) Celles se ramenant pour la classification à un problème de reconnaissance 2D (seules les vues 3D de la base d'apprentissage sont requises) [Rom04].

I.3 Identification et vérification de visage

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

Les systèmes de reconnaissance peuvent être divisés en deux types selon leur fonctionnalité :

I.3.1 Identification

Ayant un flux d'image ou une image qui contient un visage, l'identification et la reconnaissance à partir d'une base des données sans autre connaissance a priori de l'identité de la personne. Cette tâche donne lieu à une comparaison de type 1-N : on compare une image du flux avec les images enregistrées dans la base de donnée (fig.I.1). Une application est la reconnaissance de personne « dangereuses » enregistrées dans une base de données. Cette application sujette à des contraintes de

temps de calcul d'où l'enregistrement de caractéristique du visage est conditionné [Wal09].

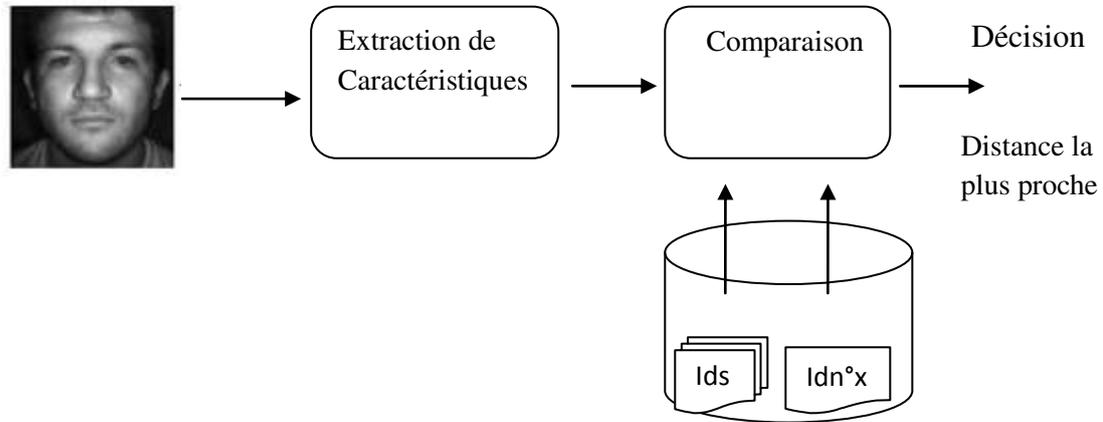


Fig.I.1: Schéma d'identification du visage.

Dans l'identification l'erreur commise par ce type de système est d'attribuer à l'individu présenté une identité autre que la sienne. Les performances de ces systèmes sont mesurées à l'aide du taux d'identification correcte :

$$TIC = \frac{\text{nombre de tests correspondant à une identification correcte}}{\text{nombre total de tests}}$$

I.3.2 Vérification

Ayant un flux d'image ou une image qui contient un visage, la vérification est la reconnaissance de visage à partir d'une image de référence enregistrée au préalable dans une puce ou fichier. C'est une vérification de l'identité revendiquée c'est-à-dire que la personne déclare au préalable son identité (figure I.2). Cette tâche donne lieu de comparaison 1 : 1. Une des applications est le contrôle ou vérification d'identité à l'entrée des Etats Unis par exemple (le nouveau passeport biométrique).

La référence biométrique peut alors être enregistrée dans une base de données mais aussi sur une carte à puce détenue par la personne [Wal09].

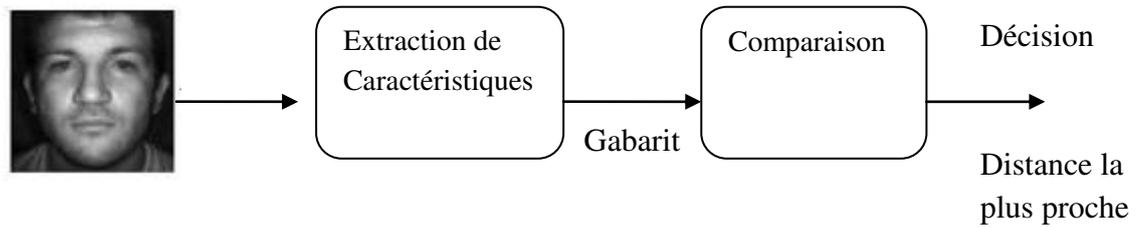


Fig.I.2 : Vérification du visage

Dans la phase de vérification, un utilisateur doit faire une demande d'authentification de son identité. Le système doit déterminer si l'identité proclamée est acceptée ou rejetée. Deux types d'erreurs peuvent être commis par le système :

Faux Rejets **FR** (False-Rejection) : correspond au cas où le système rejette un client légitime.

- Fausse Acceptation **FA** (False-Acceptance) : correspond au cas où le système accepte un imposteur.

La fiabilité d'un système de reconnaissance biométrique est caractérisée par deux valeurs statistiques :

- Le Taux de Faux Rejets (**FRR**) : il exprime le pourcentage d'utilisateurs légitimes rejetés.

$$FRR = \frac{\text{nombre de FR}}{\text{nombre de clients}}$$

- Le Taux de Fausses Acceptations (**FAR**) : il exprime le pourcentage d'imposteurs acceptés.

$$FAR = \frac{\text{nombre de FR}}{\text{nombre d'imposteurs}}$$

I.4 Systèmes de reconnaissance de visages

Dans un système de reconnaissance de visages [TOU 06]

depuis son acquisition, l'image suit un processus bien précis pour arriver à déterminer ou à vérifier l'identité du porteur de visage. Ce processus peut être présenté par le diagramme suivant:

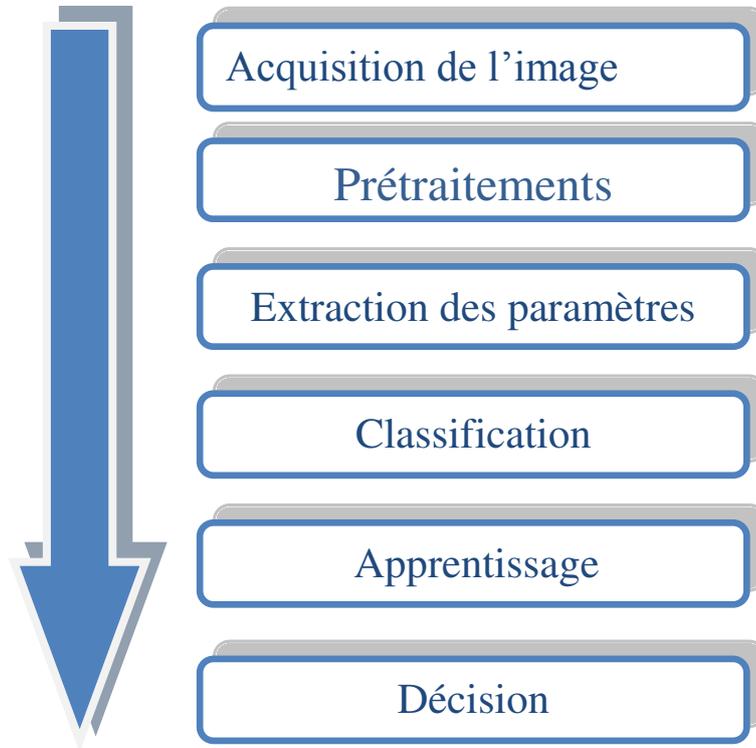


Fig.I.3 : Processus d'un système de reconnaissance de visages.

Comme l'indique la figure ci-dessus, un système de reconnaissance du visage passe par plusieurs étapes qui sont:

I.4.1 L'Acquisition de l'image

Cette étape consiste à extraire l'image de l'utilisateur du monde extérieur dans un état statique à l'aide d'un appareil photo ou dynamique à l'aide d'une caméra.

L'image dans cette étape est dans un état brut ce qui engendre un risque de bruit qui peut dégrader les performances du système.

I.4.2 Les prétraitements

Le rôle de cette étape est d'éliminer les parasites causés par la qualité des dispositifs optiques ou électroniques lors de l'acquisition de l'image en entrée, dans le

but de ne conserver que les informations essentielles et donc préparer l'image à l'étape suivante.

Elle est indispensable car on ne peut jamais avoir une image sans bruit à cause du background et de la lumière qui est généralement inconnue.

Cette étape peut également contenir la détection et la localisation du visage dans une image, surtout là où le décor est très complexe.

I.4.3 L'extraction de paramètres

En plus de la classification, l'étape de l'extraction des paramètres représente le cœur du système de reconnaissance, elle consiste à effectuer le traitement de l'image dans un autre espace de travail plus simple et qui assure une meilleure exploitation de données, et donc permettre l'utilisation, seulement, des informations utiles, discriminantes et non redondantes.

I.4.4 La classification (Modélisation)

Cette étape consiste à modéliser les paramètres extraits d'un visage ou d'un ensemble de visages d'un individu en se basant sur leurs caractéristiques communes.

Un modèle est un ensemble d'informations utiles, discriminantes et non redondantes qui caractérise un ou plusieurs individus ayant des similarités.

I.4.5 L'apprentissage

C'est l'étape où on fait apprendre les individus au système, elle consiste à mémoriser les paramètres, après extraction et classification, dans une base de données bien ordonnées pour faciliter la phase de reconnaissance et la prise d'une décision, elle est en quelque sorte la mémoire du système.

I.4.6 La décision

C'est l'étape qui fait la différence entre un système d'identification d'individus et un autre de vérification.

Dans cette étape, un système d'identification consiste à trouver le modèle qui correspond le mieux au visage pris en entrée à partir de ceux stockés dans la base de données, il est caractérisé par son taux de reconnaissance. Par contre, dans un système

de vérification il s'agit de décider si le visage en entrée est bien celui de l'individu (modèle) proclamé ou il s'agit d'un imposteur, il est caractérisé par son EER (Equal Error Rate).

I.5 La reconnaissance automatique de visages

Un système automatique de reconnaissance de visages doit intégrer une étape d'apprentissage durant laquelle il associe l'allure du visage à l'identité d'une personne. Cette étape est réalisée chez les êtres humains d'une façon spontanée et évolutive. Dans un système artificiel, cette étape permet de construire une base de données des personnes connues, stockant des images étiquetées des identités [AHC09]. Pour ce faire, un système automatique comporte deux modes de fonctionnement : un mode enrôlement et un mode identification. Le premier mode sert à extraire pour chaque personne les éléments caractéristiques et les met sous la forme d'un vecteur caractéristique, appelé par la suite signature. Cette dernière, associée à une étiquette d'identité, sera stockée dans une basée de données dédiée. Le mode d'identification permet de reconnaître une personne à partir de son image faciale, c'est à dire de retrouver l'identité associée à l'image.

Nous présentons dans la (fig.I.4) une illustration générale d'un système de reconnaissance de visages dans ces deux modes de fonctionnement. Les méthodes appliquées dans les modules formant le système dépendent essentiellement de la technique d'acquisition d'images de visage.

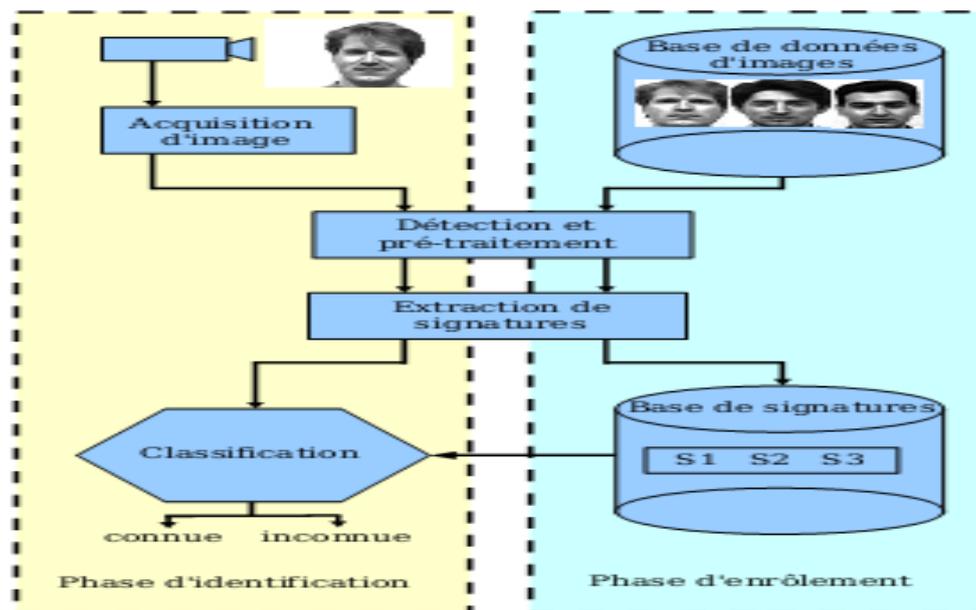


Fig.I.4 : Fonctionnement d'un système de reconnaissance automatique de visages.

En effet, s'il s'agit d'un scanner tridimensionnel ou d'un système d'acquisition stéréoscopique, la détection de visages et l'extraction de signatures reposent sur des techniques de traitement 3D [Sam07]. Cette thématique est actuellement en pleine expansion. Elle met en évidence l'information de profondeur qui enrichit les données utilisées aussi bien dans la phase d'enrôlement que dans la phase d'identification. Toutefois, elle est limitée par le prix élevé de l'appareil d'acquisition (s'il s'agit d'un scanner 3D) et la difficulté d'installation (s'il s'agit d'un système stéréoscopique). Cette limite laisse le traitement 2D des images de visages plus accessible et encore très exploré, tant dans le domaine académique que dans le milieu industriel. Les travaux de recherche menés dans cette thèse portent sur l'utilisation des images 2D de visages. Le système comporte deux modules qui appliquent les mêmes traitements dans les deux modes de fonctionnement. Le premier module permet de détecter le visage dans l'image brute captée par le système d'acquisition, il fournit ainsi au second module une image contenant seulement le visage. À partir de celle-ci, le second module extrait la signature discriminante. Elle sera soit stockée soit utilisée pour la classification. En mode hors-ligne (enrôlement), elle sera étiquetée et stockée dans une base de données dédiée. En mode en ligne (identification), elle servira comme entrée à un troisième module s'occupe de la classification de cette signature et de fournir la décision finale : personne inconnue ou personne connue.

I.6 Reconnaissance 3D de visage

Chaque visage individuel peut générer une grande variété d'images différentes. Cette grande diversité d'images de visages rend l'analyse avec beaucoup de difficultés. Outre les différences générales entre les individus les faces présentent des variations dans l'apparence d'images d'un seul visage. On peut les quatre sources suivantes de variabilité:

- Pose de changements peut entraîner des variations dramatiques dans les images. En raison de différentes occlusions de parties l'objet devient visible ou invisible. En outre, leurs configurations spatiales sont modifiées [Vol00].
- Changements d'éclairage ont un effet sur l'apparition d'un visage, même si la pose de la face est fixée. Les positions de distribution des sources de lumière autour d'un visage ont pour effet de modifier la répartition de la

luminosité dans les images, les emplacements des ombres joints et les réflexions spéculaires. En outre, des ombres peuvent générer des contours importants dans les images faciales.

- Les expressions faciales un outil important dans la communication humaine sont une autre source des variations dans les images. Seuls quelques points de repère du visage qui sont directement couplés avec la structure osseuse du crâne, comme la distance ou la position interoculaire générale de l'oreille sont constants dans un visage. La plupart des autres caractéristiques peuvent changer leur configuration spatiale ou position en raison de l'articulation de la mâchoire ou à l'action des muscles, comme les sourcils mobiles, les lèvres ou les joues.

Avec le facteur temps un visage change en raison du vieillissement, changement de coiffure, selon le maquillage ou d'autres accessoires. L'isolement et la description explicite de toutes ces différentes sources de variations doivent être le but ultime d'un système d'analyse du visage performant et stable. Il est souhaitable que les paramètres caractéristiques d'une personne soient invariables et ne doivent pas être perturbés par ces modifications. Le principal défi pour la conception de tels systèmes est de trouver ou de choisir une description de ces paramètres qui permet à la fois, la modélisation appropriée d'images d'un côté et description avec précision d'une image sur l'autre. Quelques unes des sources de variations, comme l'éclairage et la pose, obéissent à l'intégrité des lois physiques de la nature. Ces lois reflètent des contraintes dérivées de la géométrie tridimensionnelle des visages et l'interaction de leurs surfaces avec la lumière. Ils sont de façon optimale imposée par la représentation en 3D qui a donc été choisie pour le modèle morphable. D'un autre côté, il ya des régularités supplémentaires entre les faces qui ne sont pas formulées que par les lois physiques, mais peut être obtenues en exploitant les statistiques générales de visages. Ces méthodes sont également représentées comme l'apprentissage à partir d'exemples. On s'attend à ce que l'apprentissage conforme à l'incorporation des contraintes physiques sont plus efficaces dans les tâches comme la généralisation à partir d'une seule image de visage dans des conditions d'éclairage différents [Vol00].

En conséquence, le modèle 3D morphable utilise des lois physiques modèle pose et l'éclairage et des méthodes statistiques à l'identité du modèle et d'expression. Comme on le voit dans les deux prochaines sections, ces méthodes statistiques exigent que les

faces à mettre en correspondance les données de chaque visage en 3D sont composés de plus de 200 000 sommets et 400 000 correctifs triangulaires avec une collection de bords connectés et les sommets comme dans la figure 1

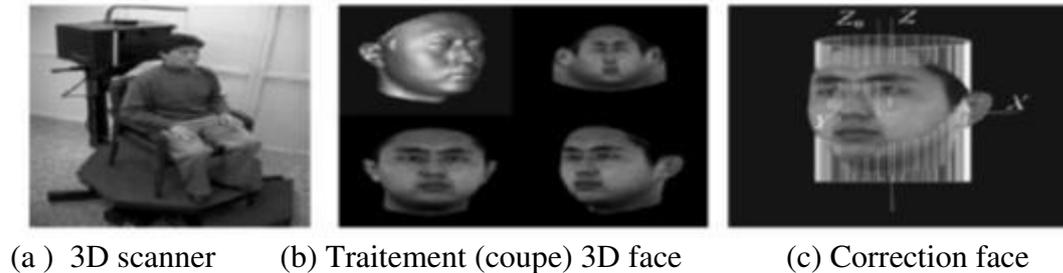


Fig.I.5: 3D scanner et traitement original de données visage.

C'est parce que les échantillons acquis visage 3D sont bruyants et épineux, un préprocesseur est nécessaire pour enlever ces pointes, à remplir les trous qui résultent de la suppression des pointes par interpolation et pour lisser la surface de tout le visage. Puis, le visage est coupé à partir de l'ensemble des données numérisées par la coupe limite et à supprimer les données 3D se trouvant sur les cheveux et l'épaule (voir Fig. 1 (b)).

Les données visage enfin, tout garni 3D sont corrigées à un système de coordonnées homogène. Ici, les discrets sommets visage 3D / un cylindre. L'axe central du cylindre est défini comme l'axe Z avec la direction positive vers le haut, et la direction à travers la pointe du nez de l'avant et perpendiculaire à l'axe Z est l'axe Y, et l'axe X est obtenu par produit vectoriel de l'axe Z et l'axe Y (voir Fig. 1 (c)). Comme chaque échantillon visage 3D ne doit pas avoir les mêmes quantités de données nombre de sommets en reconnaissance de visage, nous prêtons attention à la fonctionnalité de visage des domaines tels que les yeux, front, nez, bouche, et ainsi de suite. Dans l'ordre de réduire les informations inutiles, les échantillons face alignées sont réalisée comme suit. Tout d'abord, deux points caractéristiques clés (les bout du nez et un autre hors-bord sommet sur le front à gauche) sont situé (voir Fig. 2 (b)). Parce que l'axe des Y représente les informations sur la profondeur du visage, le bout du nez sommet peut être recherché facilement que le sommet avec le maximum de la valeur y de 3D des données de face. La texture extérieure sommet sur le sourcil gauche peut être déterminée par l'algorithme de détection de visage. Correspondant de

coordonnées peut être acquise par le rapport de correspondance entre la texture et l'information de forme. Il doit être mentionné que tous les échantillons visage 3D utilisés dans le présent document conservent l'alignement sur la base des caractéristiques du visage. Le procédé de chercher les deux points caractéristiques doit être mené une seule fois, et les numéros de séquence de la clé à deux points sur la surface de face peuvent être enregistrés. Étant donné une nouvelle 3D échantillon visage, les deux points clés peuvent être repérés par leurs numéros de séquence directement et triangulaire patches. Il est nécessaire d'aligner les faces 3D avec correspondance pixel avec les propriétés du visage de l'homme par le procédé de maillage [Vol00].

En outre, le résultat des échantillons visage 3D garde l'alignement sur la base de la face caractéristiques, contiennent les mêmes numéros des sommets et triangulaires patches. Cela signifie que le sommet set triangulaires des tâches de chaque échantillon peut être enregistré par la même règle, et les sommets ayant le même nombre séquentiel mais de différents échantillons sont dans le même endroit du visage.

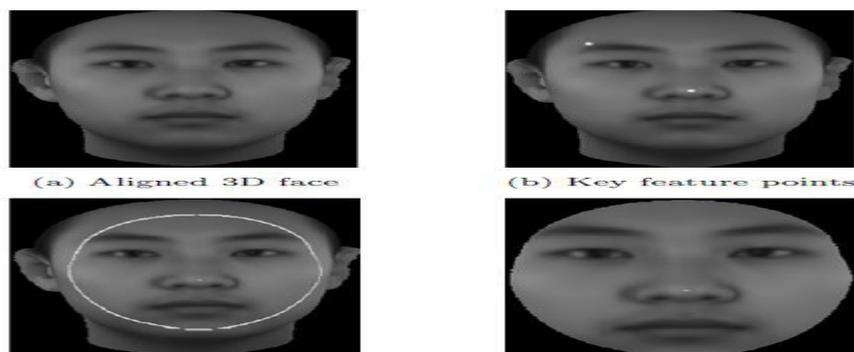


Fig.I.6 : Le visage 3D retrimming et le résultat final.

I. 6.1 Multimodalité 2D +3 D reconnaissance de visage

Notre reconnaissance de visage optimise l'avantage d'être utiliser à la fois en données 2D et 3D du visage grâce à l'utilisation de l'information aux niveaux description de l'image et le score 2D et 3D du visage. Au cours de la description, les acquis en 2D et 3D des données faciales subissent une série de prétraitement afin d'obtenir images 2D et 3D normalisées tel que décrit dans la section précédente. Au niveau de l'image, les images 2D et 3D sont recombinaison dans un ensemble de données plus indépendantes et discriminables par la recombinaison d'image basée sur l'APC.

Les données recombinaées peuvent être utilisés individuellement ou en parallèle en tant qu'entrée au classificateur visage. Dans notre visage système de reconnaissance, deux procédés correspondants sont effectués indépendamment sur les données recombinaées # 1 et # 2. Les données sont recombinaées transformé dans leurs espaces respectifs métrages utilisant Fisher faces.

Les deux fonctions sont utilisées pour déterminer les scores de dissemblance entre visages de la sonde et de la galerie. Les processus d'appariement donnent lieu à deux notes distinctes. Enfin, au niveau score, le score final est obtenu en tant que somme.

I. 7 Techniques de reconnaissance de visage

Nous classons toutes ces approches en deux sous catégories à savoir :

- a) Les méthodes de sous espace (subspace methods)
- b) Les approches à base de caractéristiques géométrique (geometric feature based methods).

Dans la suite, nous détaillons ces deux sous catégories et les approches qui en découlent. Il est à noter que quelques unes de ces approches ont été appliquées sur des images de profondeurs, profitant ainsi du développement mathématiques considérable que les approches 2D ont gagnées, ces quelques dernières années.[BBA06]

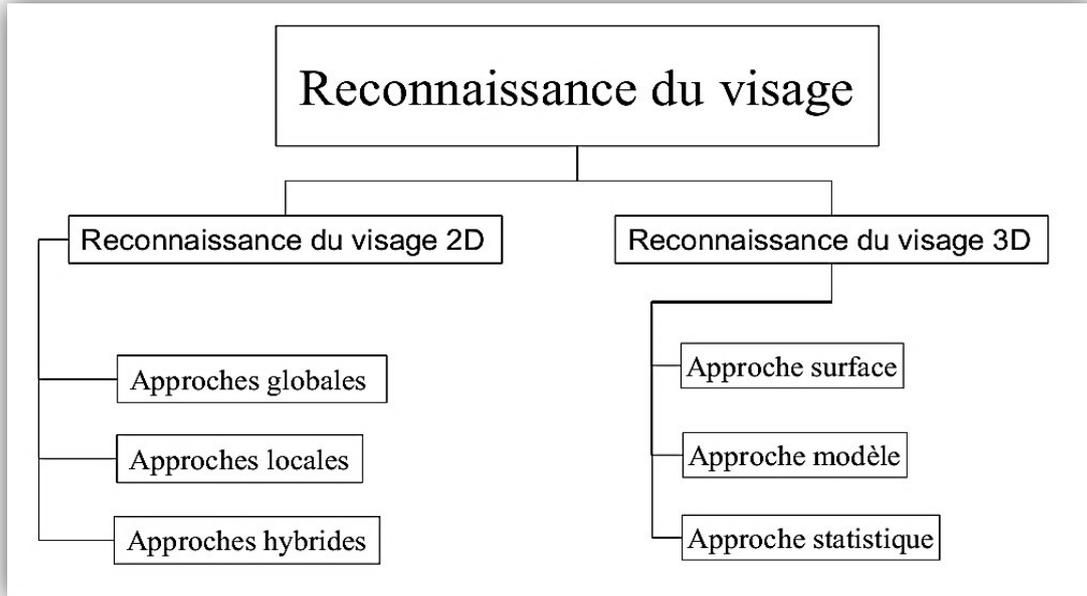


Fig.I.7 : Taxonomie des approches de reconnaissance de visages 2D, et 3D.

I.7.1 Techniques 2D de reconnaissance de visage

Plusieurs méthodes d'identification de visages ont été proposées durant les vingt dernières années. L'identification de visage est un axe de recherche ouvert attirant des chercheurs venants de disciplines différentes : psychologie, reconnaissance de formes, réseaux neuraux, vision artificielle et infographie. [SOU08]

Avant de détailler les différentes techniques liées à la reconnaissance de visage 2D, nous allons d'abord présenter un aperçu des études faites par les chercheurs en cognition et en reconnaissance faciale du visage. En effet, la connaissance des résultats de ces études est importante car elle permet le développement de nouvelles approches. Le but ultime de la reconnaissance faciale est de rivaliser, voir même dépasser, les capacités humaines de reconnaissance. Les résultats fondamentaux de ces études est comme suit :

- Les humains peuvent reconnaître des visages familiers dans des images de faible résolution
- La capacité de tolérer les dégradations des images augmente avec la familiarité
- Les informations haute fréquence seules, soit les contours, sont insuffisantes pour obtenir une reconnaissance faciale performante.

- Les caractéristiques faciales sont traitées de manière holistique
- Parmi les différentes caractéristiques faciales, les sourcils sont les plus importants pour la reconnaissance
- La forme du visage est généralement codée de manière caricaturale
- La pigmentation du visage est aussi importante que sa forme.
- La couleur joue un rôle important spécialement lorsque la forme est dégradée
- Les changements d'illumination influencent la capacité de généralisation
- Le mouvement des visages semble faciliter la reconnaissance de manière conséquente
- Le système visuel progresse d'une stratégie locale vers une stratégie holistique au cours des premières années de la vie
- Identité faciale et expressions sont traitées par des systèmes séparés. [SOU 08]

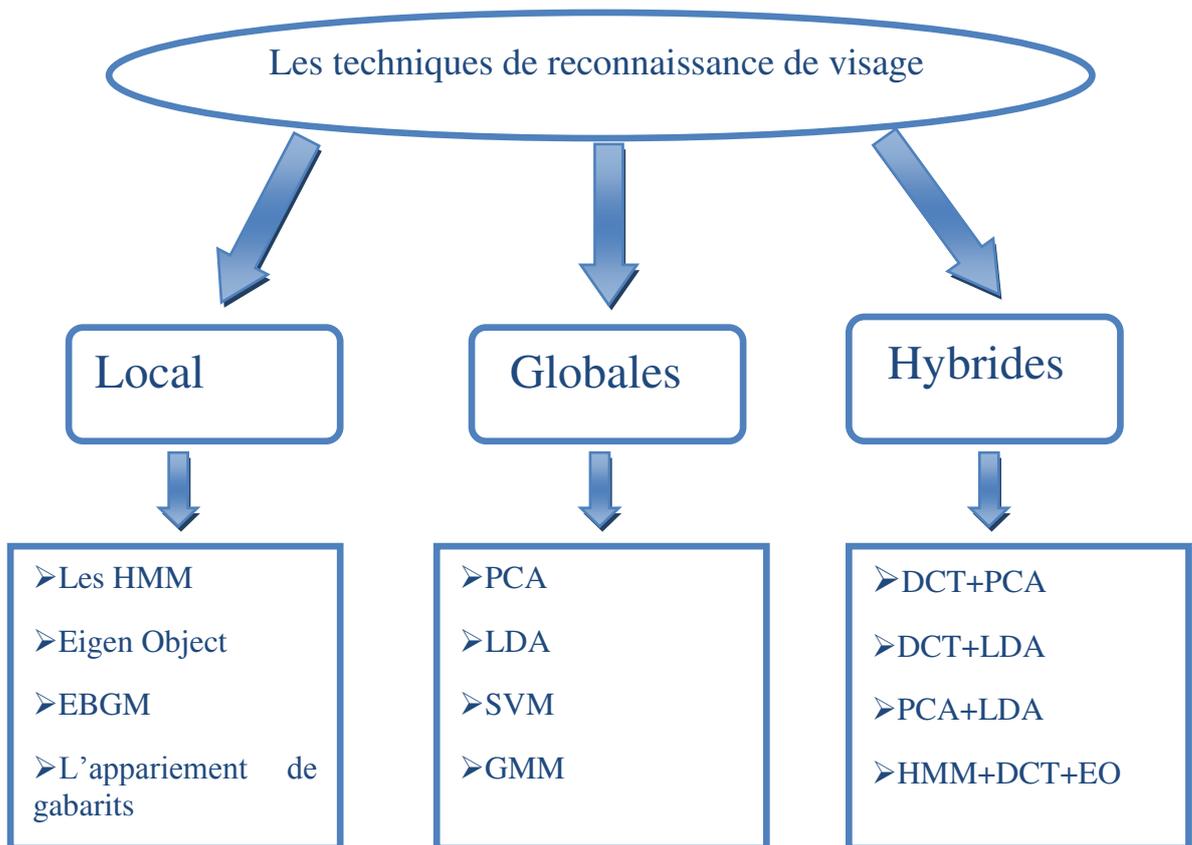


Fig.I.8 : Les techniques de reconnaissance de visage.

I.7.1.1 Les Approche globales

Le principe de ces approches est d'utiliser toute la surface du visage comme source d'information sans tenir compte des caractéristiques locales comme les yeux, la bouche,...etc.

Les algorithmes globaux s'appuient sur des propriétés statistiques bien connues et utilisent l'algèbre linéaire. Ils sont relativement rapides à mettre en œuvre, mais sont sensibles aux variations d'illumination, de pose et d'expression faciale.

Parmi les approches les plus importantes réunies au sein de cette classe on trouve:

- L'analyse en Composantes Principales (PCA ou Eigen Faces)
- L'analyse Discriminante Linéaire (LDA)
- Machine à Vecteurs de Support (SVM)
- Les réseaux de neurones (RNA)
- Mélange de gaussiennes (GMM)
- Modèle surfacique du visage (3D)
- L'approche statistique et l'approche probabiliste. [ARO 11]

I.7.1.2 Approche locale (Géométrique)

On les appelle aussi les méthodes à traits, géométriques, à caractéristiques locales, ou analytiques.

Ce type consiste à appliquer des transformations en des endroits spécifiques de l'image, le plus souvent autour des points caractéristiques (coins des yeux, de la bouche, le nez, ...), l'énergie sera accordée aux petits détails locaux évitant le bruit engendré par les cheveux, les lunettes, les chapeaux, la barbe, etc. . [ARO 11]

Mais leur difficulté se présente lorsqu'il s'agit de prendre en considération plusieurs vues du visage ainsi que le manque de précision dans la phase "extraction" des points constituent leur inconvénient majeur. Parmi ces approches on peut citer :

- Modèles de Markov Cachés (Hidden Markov Models)
- L'algorithme Elastic Bunch Graph Matching (EBGM)
- Eigen objectes (EO)
- L'appariement de gabarits [ARO 11]

I.7.1.3 Les approches hybrides

Comme on a vu précédemment plusieurs approches ont été proposées pour la reconnaissance de visages, sauf qu'aucune d'elle n'est capable de s'adapter aux changements d'environnements tels que la pose, expression du visage, éclairage, la balance...etc.

La robustesse d'un système de reconnaissance peut être augmentée par la fusion de plusieurs méthodes. Il est par ailleurs possible d'utiliser une combinaison de classificateurs basés sur des techniques variées dans le but d'unir les forces de chacun et ainsi pallier à leurs faiblesses. [ARO 11]

I.7.2 Techniques 3D de reconnaissance de visages

Nous avons passé en revue les méthodes de reconnaissance 2D de visages. Malgré les avancées réalisées ces dernières années, les techniques de reconnaissance 2D de visages robustes aux différents facteurs de variabilité (éclairage, pose, occultation) sont loin d'être développées. La reconnaissance 3D de visages constitue une alternative prometteuse pour surmonter ces problèmes, surtout depuis l'apparition de dispositifs d'acquisition 3D performant. L'avantage principal des approches basées modèle 3D réside dans le fait que le modèle 3D conserve toutes les informations sur la géométrie de visage, ce qui permet d'avoir une représentation réelle de ce dernier. Dans cette section, après avoir rapidement évoqué les systèmes d'acquisition 3D, nous présenterons les travaux récents sur la reconnaissance 3D de visages.

I.7.2.1 Systèmes d'acquisition 3D

Généralement le modèle du visage est représenté par des images 2.5D et 3D. L'image 2.5D (image de profondeur) correspond à une représentation bidimensionnelle d'un ensemble de points 3D (x,y,z) où chaque pixel dans le plan X-Y stocke la valeur de sa profondeur z . On peut assimiler une image 2.5D à une image en niveau de gris où les pixels noirs correspondent au fond tandis que les pixels blancs représentent les points de surface les plus proches de la caméra. Par ailleurs, la méthode la plus simple pour représenter un visage 3D est le maillage polygonal 3D, ce dernier correspond à une liste de points connectés par des arêtes (polygones). Il existe plusieurs techniques pour construire un

maillage 3D, les plus utilisées combinent des images 2.5D ou bien exploitent des systèmes d'acquisition 3D tel que le scanner 3D.

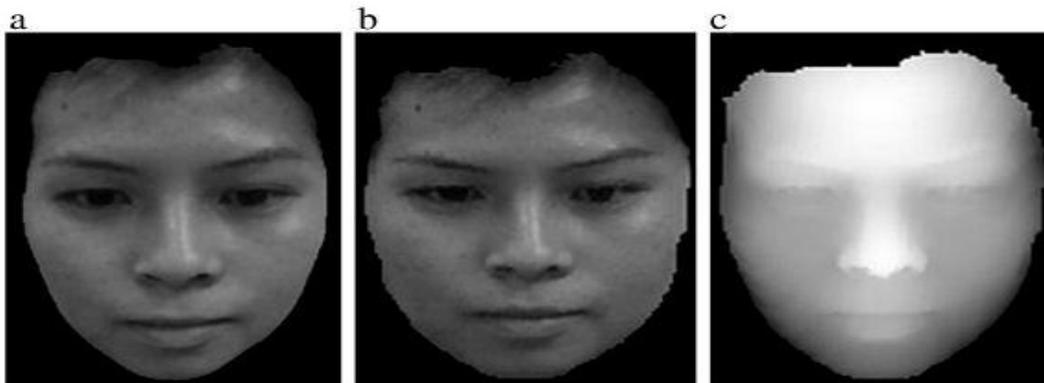


Fig.I.9 : (a) Visage normalisé; (b) 2D image; (c) 3D image.

Différents systèmes d'acquisition sont utilisés pour la reconstruction 3D de visages. Ces derniers peuvent être décomposés, en fonction des capteurs de vision qu'ils utilisent, en deux grandes classes: la *vision dite active* et la *vision dite passive*.

- Nous appelons *vision active* tout système de vision utilisant un capteur actif qui permet d'illuminer les objets de la scène d'une manière particulière en projetant sur leur surface des motifs structurés (un spot, un plan lumineux, une grille, etc.), ceci permet d'en extraire des informations 3D indépendamment des conditions propres d'éclairage de la scène. Ainsi, l'objectif de la vision active est de réduire la complexité des traitements à effectuer pour reconstruire la scène. En effet, les motifs projetés sur la scène sont facilement identifiables [Hu89] à l'inverse des primitives « naturelles » tel que les contours. Le scanner 3D, par exemple, est un système de vision active en lumière structurée qui est capable de scanner les formes 3D via un laser et les textures (couleur de la peau) via une caméra simultanément.
- En *vision passive*, la vision binoculaire ou vision stéréoscopique (utilisation de deux caméras fixes).

Les techniques de reconnaissance 3D de visages peuvent être regroupées en trois catégories principales : approches basées modèle, approches 3D, et approches multimodales 2D + 3D.

I.7.2.2 Approches modèle

Ces approches construisent, à partir des points 3D, des modèles de visages qu'elles utilisent par la suite pour la reconnaissance. L'ensemble des visages est représenté par un espace vectoriel.

Cette approche a été appliquée à une base de données qui contient 100 visages d'hommes et 100 visages de femmes acquis avec un *scanner laser CyberwareTM3030PS*. Les points 3D des modèles de visages générés sont représentés par leurs coordonnées cylindriques définies par rapport à un axe vertical. Pour chaque visage de référence, les coordonnées et les valeurs de texture de tous les sommets ($n=75\ 972$) sont regroupées pour former deux vecteurs : un vecteur de forme et un vecteur de texture. Une fois le modèle générique créé, l'étape suivante consiste à l'ajuster sur l'image 2D à partir des paramètres de forme et de texture. La synthèse d'image permet de rendre les nouvelles positions projetées des sommets du modèle 3D, à l'aide de l'illumination et de la couleur extraites. Enfin, l'étape de reconnaissance est réalisée en mesurant la distance de Mahalanobis entre la forme et les paramètres de texture des modèles dans la galerie et le modèle d'ajustement. L'identification a été évaluée sur deux bases de données d'images à accès libre : CMU-PIE et FERET. Un taux de reconnaissance de 95 % sur l'ensemble de données CMU-PIE et 95.9 % sur l'ensemble de données FERET a été obtenu.

I.7.2.3 Approches 3D

Elles sont subdivisées en deux catégories : les approches basées surface qui utilisent la géométrie de la surface du visage et les approches holistiques 3D.

1) Approches surface

Dans ce cas, le problème de la reconnaissance 3D de visages est celui de l'alignement de deux surfaces 3D qui modélisent les deux visages à appairer. L'algorithme généralement utilisé est l'algorithme du plus proche voisin itéré, ou ICP (Iterative Closest Point), qui a été introduit par [Bes92].

Il consiste en une optimisation alternée d'appariements et de transformations. Ainsi, à partir d'une transformation initiale, les deux étapes suivantes sont réitérées :

- mise en correspondance (plus proche voisin) : on apparie chaque primitive du modèle transformé avec la primitive la plus proche dans la scène.
- recalage : la transformation (translation + rotation) est généralement calculée aux sens des moindres carrés, surtout si l'on travaille avec des points. Si l'on possède une information d'incertitude, on peut l'utiliser dans les étapes terminales pour affiner la solution.

Une approche basée sur l'algorithme ICP est donnée par Cook et al [Coo04]. Ils utilisent l'algorithme ICP pour établir la correspondance entre les surfaces 3D afin de compenser des erreurs dues à la nature non rigide des visages. En effet, les visages sont comparés en employant un modèle statistique, en l'occurrence un Mélange de Modèles Gaussiens (MMG). La distribution des erreurs est alors paramétrée. Les auteurs ont réalisé des expériences sur la base de données 3D RMA [Beu00], une base créée au sein du Signal et Image Center du Royal Military Academy (Bruxelles - Belgique). Un dispositif de vision active (structured light-based) est utilisé pour l'acquisition. Cette base contient les images de 120 individus, capturées en deux sessions, sous formes de points 3D non maillés. Un taux d'identification de 97.33% a été obtenu. Irfanoglu et al. [Irf04] proposent une autre approche qui utilise l'algorithme ICP pour mettre en correspondance les points caractéristiques de visages. Ils ont développé une méthode pour obtenir un appariement dense de points du maillage contenant les points présents sur tous les visages, de sorte que l'alignement de visage soit trivialement obtenu. Une fois que l'appariement dense est établi, la distance PSD "Point Set Distance", qui représente une approximation discrète du volume entre les surfaces faciales, est employée pour calculer la distance entre deux nuages différents de points. Dans leur expérimentation, les auteurs ont testé l'algorithme sur la base de données 3D RMA, ils ont relevé un taux d'identification de 96,66%.

Quoique l'ICP soit un outil puissant pour évaluer la similitude entre deux visages, les méthodes basées ICP ne sont pas capables de traiter le problème de changement d'expression, car elles considèrent le visage 3D comme un objet rigide, ce qui n'est pas le cas dans la réalité. Pour résoudre ce problème, plusieurs approches ont été développées.

Moreno et al. [Mor03] proposent de segmenter le visage en plusieurs régions en utilisant la courbure gaussienne. Ainsi, chaque modèle facial 3D est subdivisé en 86 régions différentes grâce à un algorithme de segmentation qui exploite les signes de la médiane et des courbures gaussiennes pour isoler les régions ayant des courbures significatives. Cet espace de caractéristiques est ensuite réduit pour augmenter l'efficacité de la méthode.

Enfin, un vecteur de caractéristiques est créé pour chaque sujet. Les expérimentations ont été menées sur un ensemble de 420 modèles 3D de visages appartenant à 60 sujets, comprenant des images avec des variations de lumière, de rotation et d'expressions de visages. Les auteurs ont obtenu un taux de reconnaissance de 78 %.

Le processus de segmentation peut aussi être utilisé pour traiter la reconnaissance de visage comme un problème de reconnaissance d'objets non rigides, et ainsi améliorer la robustesse de la reconnaissance face aux variations d'expressions. Chua et al. [Chu00] ont remarqué que des régions du visage, comme le nez, l'orbite de l'œil et le front subissent beaucoup moins les déformations en cas de changements d'expressions. Ils identifient les régions faciales "rigides" en utilisant une comparaison "deux par deux" entre les signatures de points [Chu97] pour différentes expressions de visage de la même personne. Ainsi, seulement les parties rigides sont stockées dans une base de données indexée. Les modèles de la base sont classés en fonction de leur similitude.

Leur expérimentation a montré un taux de reconnaissance de 100 % sur une base de données de 60 sujets et 4 variations d'expressions de visage. Enfin, [Ben06] propose d'introduire une nouvelle métrique orientée région dans l'algorithme de l'ICP. Il s'agit de segmenter un visage en plusieurs régions ayant des influences plus ou moins importantes sur la déformation de la forme 3D du visage. Le calcul de la distance (similarité) est pondéré en fonction de ces régions en donnant plus de poids aux régions statiques qu'aux régions mimiques du visage. Un taux de reconnaissance de 97.56 % a été obtenu.

2) Approches holistiques 3D

Les techniques holistiques comme l'ACP ont été largement utilisées dans la reconnaissance faciale 2D. Plus récemment, ces techniques ont été aussi étendues aux données 3D de visage. Ainsi, Heshner et al. [Hes03] ont

appliqué l'ACP sur une base de données contenant 222 images de profondeur (37 personnes), avec six expressions faciales différentes. Ils ont obtenu un taux de reconnaissance de 83%. Heseltine et al. [Hes04a] [Hes04b] ont développé deux approches qui appliquent l'ACP sur des représentations tridimensionnelles du visage. La première méthode [Hes04a] utilise l'approche basée ACP « eigensurface » sur des modèles 3D maillés de visage de la base de données de l'Université de York (The 3D Face Database, 2003). Elle arrive à un taux de reconnaissance de 87%. La deuxième approche est une extension de la méthode Fischerface 2D de Belhumeur [Bel97] aux données surfaciques de visage 3D. Les tests ont été effectués sur une base de données contenant 1770 modèles 3D maillés de 280 sujets ayant différentes poses et expressions faciales. Le taux de reconnaissance obtenu est de 88%. Les techniques basées ACP ont également été combinées avec d'autres méthodes de classification, comme le modèle caché de Markov (EHMM) puis appliquées à la reconnaissance 3D de visages [Tsa04]. Enfin, d'autres approches basées sur l'Analyse Discriminante Linéaire [Gok05] ou l'Analyse des Composantes Indépendantes [Kim03] ont aussi été développées pour l'analyse des données 3D de visages.

I.7.2.4 Approche géométrique ou locale 3D

Par rapport aux approches « holistiques », les techniques d'identification 3D du visage basées sur les caractéristiques faciales locales de type géométriques restent relativement peu développées, alors que dans le même temps des approches locales d'identification 2D, telles que l'« Elastic Bunch Graph Matching (EBGM) [Wis97] » et l'analyse locale de caractéristiques, connaissent un grand succès auprès des fournisseurs de données (Recognition Vendors Test 2002 [Phi03]).

I.7.2.5 Approche 2D+3D

Les combinaisons des techniques de reconnaissance 2D et 3D sont de plus en plus étudiées. Ainsi la combinaison des approches EBGM 2D et EBGM 3D (qui correspond à une extension de la méthode EBGM 2D aux images de profondeur de visage) a donné lieu à une technique d'identification du visage 2D+3D très

efficace évaluée sur la base FRGC [Hus05]. Cependant, l'évaluation de cette approche hybride a démontré que les performances du EBGM 2D surpassent de manière significative celles du EBGM 3D. Elle a aussi mis en avant le besoin de comprendre et d'utiliser les propriétés discriminantes des caractéristiques locales du visage afin de concevoir des techniques efficaces de reconnaissance 3D de visage.

Dans [Lee05] les auteurs proposent un système de reconnaissance de visage 3D utilisant deux dispositifs différents d'acquisition de données 3D : un scanner laser 3D et une lumière structurée. A partir des mesures 3D, huit points caractéristiques du visage (géométriquement invariables) sont extraits puis utilisés pour calculer un vecteur caractéristique comprenant les distances et les angles entre ces points. Pour la phase de reconnaissance, les auteurs ont implémenté deux algorithmes différents : le premier basé profondeur (depth-based DP) et le second utilise les SVM (feature-based SVM). Les résultats expérimentaux (sur une base de 20 personnes) ont donné un taux de reconnaissance de 95% pour la DP et de 96% pour la SVM.

Pondérée des deux partitions. Les pondérations sont déterminées empiriquement par l'analyse de Fisher discriminante linéaire. Dans ce qui suit sections, nous décrivons la recombinaison image proposée, Fisher faces et le score de niveau de fusion.

I.8 Conclusion

Durant ces trois dernières décennies, la reconnaissance automatique des visages est devenue l'une des applications les plus réussies en traitement d'images et en reconnaissance de formes. Néanmoins, et malgré la multitude des techniques proposées et les progrès réalisés, la reconnaissance robuste de visages reste très difficile.

Dans ce chapitre, nous avons défini la reconnaissance de visages 2D et 3D, puis nous avons présenté un état de l'art des systèmes de reconnaissance de visages dans lequel nous avons cité les méthodes les plus utilisées et les plus prometteuses pour évaluer ce type de système. Nous pouvons dire que la méthode hybride 2D et 3D pourrait être une bonne candidate. Car dans ce cas la couleur est assurée par l'image

2D et la profondeur par l'image 3D. La fusion de ces deux entités pourrait être encourageante.

Dans le chapitre suivant nous présenterons l'état de l'art de la fusion ainsi que les techniques utilisées et la fusion de visage 2D et 3D.

II.1 introduction

Dans ce chapitre, nous traitons la question de la fusion et ses différents niveaux. Particulièrement la fusion des scores qui fait objet de notre travail. Les principales méthodes de normalisation des scores et de combinaisons sont mises en revue. Un état de l'art des travaux les plus récents dans le domaine de la fusion pour la reconnaissance de visage est présenté en fin de chapitre.

II.2 C'est quoi la fusion ?

Ajouter une modalité à un système biométrique, c'est ajouter une nouvelle source d'information. C'est pourquoi les systèmes multimodaux permettent d'obtenir de meilleurs résultats que les systèmes unis modaux correspondants. Ajouter des modalités permet également d'augmenter l'universalité du système. Car si l'utilisateur est réfractaire à une caractéristique biométrique donnée, l'information récoltée sur les autres modalités peut compenser la faible qualité de l'acquisition.

Le processus de reconnaissance biométrique se compose de quatre étapes principales : la capture des données, l'extraction de paramètres, la comparaison à une base d'exemples, et la prise de décision. La fusion biométrique est possible à chacun de ces niveaux : aux deux premiers niveaux, elle consiste souvent en une normalisation des données suivie d'une concaténation; au niveau de la comparaison, elle se traduit par une combinaison des scores ; enfin, au niveau de la décision, elle consiste en une confrontation des résultats de classification qui aboutit à une décision finale.

II.3 Les type de fusion

La fusion d'éléments biométriques peut se référer à de nombreux scénarios différents (Figure II.1).

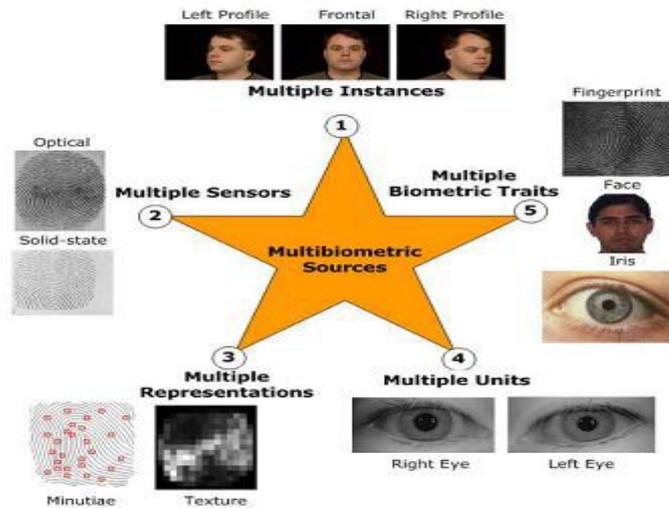


Fig. II.1 : Sources de différents types de fusion de traits biométriques [NAN05].

- **Systèmes multi algorithmes:** C'est le type de système le plus classique implicitement utilisé par de nombreuses approches. Les caractéristiques sont extraites via différents algorithmes puis fusionnées. La fusion de caractéristiques extraites via un algorithme analysant les textures et un autre la forme d'un caractère biométrique entre dans ce cadre.
- **Systèmes multi échantillons:** Un capteur unique peut capturer plusieurs instances du même caractère biométrique dans le but de rendre plus robuste l'extraction des caractéristiques ou d'enrichir le modèle biométrique d'une personne. C'est le cas, par exemple, de plusieurs captures de visage d'une personne sous différents angles. L'utilisation de vidéos entre également dans ce cadre.
- **Systèmes multi capteurs:** Plusieurs capteurs permettent de capturer le même caractère biométrique sous différents « angles ». Ainsi la capture d'un visage à l'aide d'une caméra classique et d'une caméra infrarouge entre dans ce scénario. Ce type de système permet notamment la fusion au niveau capteur, ce que ne permettent pas d'autres systèmes comme les systèmes multi caractères.
- **Systèmes multi instances:** Ce type de système permet de capturer plusieurs instances du même caractère biométrique. L'acquisition de

plusieurs empreintes digitales via le même capteur est l'exemple typique de ce type de système. Ces systèmes n'entraînent pas de surcoût de capteurs, ni le développement de nouveaux algorithmes. À ne pas confondre avec les systèmes multi échantillons.

- **Systèmes multi caractères:** Ce type de système combine différents traits biométriques d'un individu. Les fusions visage iris, ou visage empreinte digitale font partie de ce type d'approche. Ces systèmes nécessitent différents capteurs ainsi que des algorithmes dédiés à chaque caractère biométrique. Ce type de système a comme principale caractéristique que les caractères biométriques considérés peuvent être plus décorrélés que pour les systèmes multi capteurs.

La fusion de données issues de visages capturés via une caméra en lumière visible et une autre en lumière infrarouge entre dans le cadre des systèmes multi capteurs, où il est considéré que les deux captures sont issues de modalités différentes. Même si les deux captures sont sensiblement décorrélées (la chaleur émise par un visage n'est pas visible en lumière visible), la fermeture des yeux d'un individu est visible sur les deux modalités.

À noter la présence de systèmes hybrides combinant plusieurs scénarios. Une revue de nombreux systèmes biométriques multimodaux développés peut être trouvée dans [ROS09].

II.4 Les différents niveaux de fusion

La reconnaissance d'un individu via une certaine modalité suit une chaîne de traitement, de la capture jusqu'à la décision finale. L'introduction de la multi modalité implique une fusion des différentes modalités, cette fusion pouvant intervenir à différents niveaux de la chaîne de traitement.

Deux familles de fusion peuvent être considérées selon qu'elles interviennent avant ou après l'étape de matching (étape qui compare deux empreintes biométriques) [SAN09].

II.4.1 Fusion avant le matching

Avant le matching, la fusion d'informations peut avoir lieu au niveau capteur ou au niveau caractéristique.

II.4.1.1 Niveau capteur (Sensor level)

La fusion de données brutes (raw data) peut se faire uniquement si les données capturées proviennent de la même caractéristique biométrique. Les données capturées doivent en effet être compatibles pour être fusionnées (il est par exemple impossible à ce niveau de réaliser une fusion visage voix). La fusion au niveau capteur permet d'obtenir de nouvelles données par fusion des données acquises. Ce sont ces nouvelles données qui vont ensuite être utilisées pour réaliser la reconnaissance. La création d'une image 3D à partir d'images 2D est un exemple de fusion au niveau capteur. D'autres méthodes utilisent des règles simples comme la somme ou le produit réalisé pixels par pixels. Un autre exemple est la réalisation d'une mosaïque à partir de plusieurs images d'empreintes digitales [JAI02] [MOO04]. D'autres méthodes permettant la fusion au niveau capteur ont été proposées dans la littérature. Dans [BEB06], les auteurs proposent l'utilisation d'algorithmes génétiques pour calculer les poids de la fusion. Dans [RIC08] et [SIN08], les coefficients d'ondelettes des deux images sont fusionnés via de multiples Machines à de Vecteurs de Support (SVM).

Dans [RAG11], l'utilisation de la méthode d'optimisation par essaim de particules (Particle Swarm Optimization),(PSO) permet de pondérer les coefficients issus de décompositions en ondelettes afin de créer une nouvelle image par transformée en ondelettes inverse (Figure II.2).

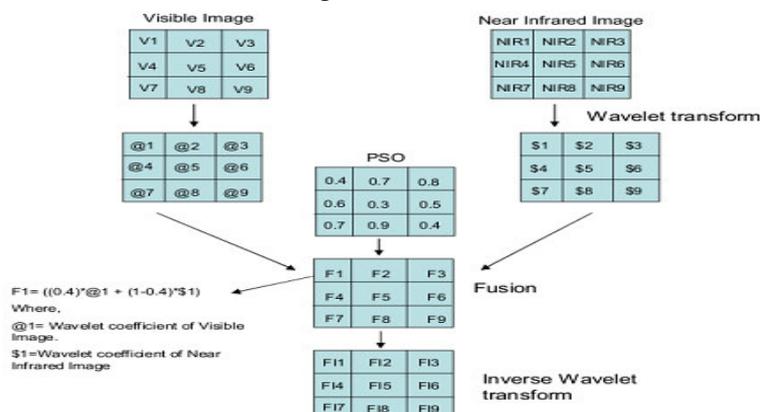


Fig. II.2 : Illustration de la fusion pondérée d'images par l'algorithme PSO [RAG11].

La fusion au niveau caractéristique consiste à combiner différents vecteurs de caractéristiques (feature vectors) issus de différentes modalités ou instances d'une même personne. Une somme pondérée dans le cadre de différentes instances d'un même trait biométrique (caractéristiques homogènes) peut ainsi être une façon de calculer un nouveau vecteur de caractéristiques. Dans le cas de caractéristiques hétérogènes, une concaténation des vecteurs caractéristiques est souvent le moyen utilisé pour la création du nouveau vecteur. Les vecteurs doivent cependant être compatibles pour une telle fusion, ainsi une concaténation des minuties d'une empreinte digitale (représentée par un graphe) avec des coefficients issus d'une ACP sur des visages n'a pas beaucoup de sens. D'autres méthodes ont été proposées comme l'utilisation de la méthode d'optimisation par essaim de particules (Particle Swarm Optimization (PSO)) dans [RAG09] et [RAG10] pour la sélection des caractéristiques discriminantes. Ce niveau de fusion est difficile à réaliser en pratique étant donné que :

- Les relations entre les différents espaces de caractéristiques doivent être connus à l'avance afin d'éliminer les caractéristiques redondantes. Cela nécessite ainsi l'utilisation d'algorithmes de sélection de caractéristiques.
- La concaténation de deux vecteurs de caractéristiques peut engendrer un vecteur de grande taille et conduire au problème de la malédiction de la dimension [DUD01]. La résolution de ce type de problème peut nécessiter l'utilisation de (trop) nombreux échantillons d'apprentissage, ce qui peut être coûteux dans le cadre biométrique.

II.4.2 Fusion après le vote (matching)

Les systèmes intégrant les informations de différentes sources pour les combiner après le vote (matching) peuvent être classés en trois classes principales : ceux réalisant la fusion au niveau décision, au niveau rang et au niveau score.

II.4.2.1 Niveau décision (Decision level)

C'est le niveau de fusion le plus abstrait. Chaque modalité est soumise à sa chaîne de traitement propre et chacun des votants renvoie la décision (accepté/rejeté) pour la modalité associée. Ce sont ces décisions qui sont ensuite fusionnées via des méthodes comme la majorité de vote (majority voting) [LAM95], l'espace de

reconnaissance de voisinage (behavior knowledge space) [LAD93], le vote pondéré (weighted voting) [KRZ92] ou encore les règles ET et OU [HAO06].

II.4.2.2 Niveau rang (Rank level)

Lorsque la sortie de chaque biométrie est une liste de résultats triée dans un ordre décroissant de confiance, la fusion peut se faire au niveau rang. Différentes méthodes existent pour combiner les rangs assignés par différents votants [Hul94]. Parmi celles-ci, citons la méthode de rang élevé (highest rank method) qui sélectionne le meilleur (minimum) rang de chaque votant, la méthode Borda count utilisant la somme des rangs calculés par chaque votant afin d'obtenir les rangs combinés, ou encore la méthode plus générale reposant sur le modèle Borda count. Dans cette dernière méthode, les rangs sont dans un premier temps pondérés à l'aide de poids trouvés via une régression logistique, avant d'être dans un second temps additionnés.

II.4.2.3 Niveau score (Score level)

La fusion des scores intervient au niveau des scores produits par chaque votant. Il s'agit de l'approche la plus courante étant donnée sa simplicité d'implémentation et sa plus grande flexibilité. Les données retournées par les votants possèdent en effet une grande richesse d'information (que ce soit une distance à un modèle ou une mesure de dissimilarité). Ce niveau de fusion fait l'objet de la section suivante.

II.5 Etat de l'art de la fusion

Audrey Poinsot a présenté une chaîne de reconnaissance biométrique basée sur la paume et le visage [AUD 06]. Les traitements utilisés sont adaptatifs, peu coûteux et facilement implantables sur un système embarqué pour une utilisation grand public. Ils respectent de plus des contraintes de faible connaissance à priori. Le système proposé est totalement sans contact, apportant confort et hygiène à l'utilisateur. L'extraction de paramètres par filtrage de Gabor est étudiée dans le cas de la paume et du visage. Les possibilités de fusion sont aussi discutées et testées sur une base multimodale de 40 personnes. De bonnes performances taux de reconnaissance ont été obtenues grâce à la paume seule (97, 92%), mais aussi grâce à la

fusion (99, 7%): l'apport de la multi modalité est confirmé. Cette étude préliminaire valide la faisabilité d'un système biométrique embarqué robuste et fiable.

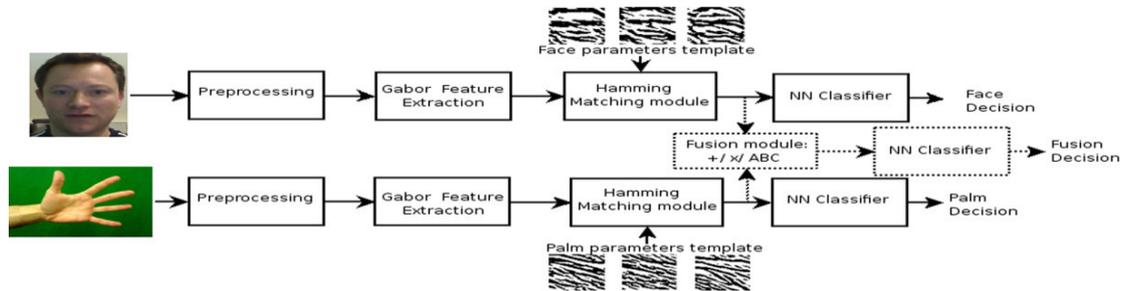


Fig. II. 3 : Chaîne de traitement complète avec possible fusion des scores [AUD 06].

Cette étude nous permet de valider la faisabilité d'un système biométrique multimodal embarqué grand public. Les traitements développés respectent les contraintes de faible coût calculatoire puisque les paumes et les visages sont traités sous Matlab en seulement 120 ms dans le cas de deux exemples par personne.

Une grande base de données (environ 140 sujets) est en cours de construction, afin d'affiner les résultats. Ce travail projette d'implémenter ces traitements sur des SoC (System On Chip). On veut réaliser la chaîne de reconnaissance en parallélisant les opérations sur des FPGA contenant des cœurs de processeur.

Dans [CHE 05], il a été prouvé expérimentalement que la PCA et la LDA peuvent être appliquées sur un nombre réduits de coefficients DCT pour réaliser une meilleure reconnaissance avec un gain en temps de calcul et en espace mémoire. Afin d'améliorer le taux de reconnaissance de la PCA et LDA, leur fusion est proposée par G. L.Marcialis et F. Roli en 2002 [MAR02], ils ont remarqué que la LDA et la PCA ne sont pas corrélées car la LDA génère un espace propre significativement différent de la PCA. Les expériences effectuées dans [MAR02] montrent que la fusion de la LDA et de la PCA ont donné de bons résultats. Dans [MES06] la fusion en score de la DCT-PCA et la DCT- RLDA est réalisée. Les méthodes DCT-PCA et DCT-RLDA sont utilisés au lieu de la PCA et la RLDA, respectivement, afin de réduire le temps d'exécution et d'avoir une meilleure performance. Des expériences ont été effectuées sur la base de données ORL indique que l'utilisation des techniques de normalisation tel que Min-Max et somme suivie d'une stratégie simple de fusion (somme simple, append, la somme pondérée) confirme les avantages de l'approche proposée en termes de taux d'identification et temps de traitement.

Dans [MAS11] on s'intéresse à la fusion radio et vidéo pour la localisation dans le domaine surveillance. On introduit un système de suivi sur les gens hybride basé sur l'utilisation combinée de la RFID La technologie UWB et des techniques de vision par ordinateur. Le système proposé tire parti des caractéristiques différentes des sous-systèmes de vision et sans fil pour atteindre une meilleure précision et de fiabilité pour le suivi des personnes. En outre des données acquises à partir des sous-systèmes peuvent être utilisés pour un système plus complexe de capture de contexte et peut être considérée comme un catalyseur d'un numéro de la demande de la vidéosurveillance. Différents scénarios ont été testés afin d'évaluer la faisabilité et performance du système. Les résultats expérimentaux montrent des avantages en le suivi des personnes tâches encourageant de nouvelles recherches.

La fusion visages Image Dynamique Quotient Morphologique avec le modèle binaire locale est étudiée dans [HOG11]. Dans cet article, on propose un éclairage nouveau normalisé local Motif binaire(LBP) basé sur l'algorithme de reconnaissance de visage sous un éclairage variable. Le projet de DMQI-LBP algorithme fusionne la normalisation éclairage, en utilisant l'image Quotient dynamique morphologique (DMQI), en le courant LBP-système basé sur la reconnaissance du visage. Ainsi, il permet d'utiliser pleinement les avantages de la compensation d'éclairage offert par l'image du quotient, les estimations avec un fonctionnement dynamique fermé morphologique, ainsi que la discrimination puissante la capacité fournie par le descripteur de LBP. Les résultats d'évaluation sur à la base de données Yale face indiquent que le projet basé sur l'algorithme DMQI-LBP améliore les performances de reconnaissance de manière significative (de 5% pour le premier rang) de la première d'origine LBP-système basé sur la reconnaissance du visage avec des variations d'éclairage difficiles. Par ailleurs, l'algorithme est simple et efficace pour mettre en œuvre, ce qui le rend très approprié pour la reconnaissance faciale en temps réel.

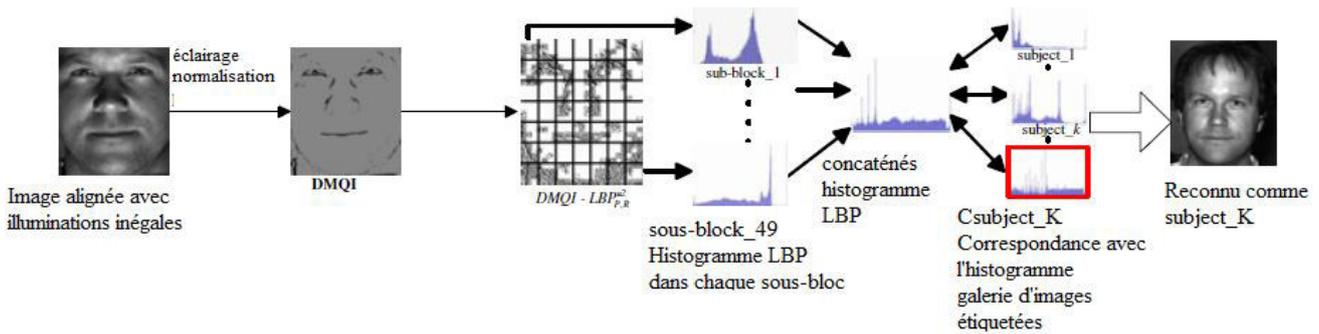


Fig.II.4 : Cadre du projet de DMQI-LBP algorithme de reconnaissance de visage [HOG11]

Des travaux récents proposent une biométrie multimodale utilisant la stratégie de fusion fonction de niveau [San 11]. Il s'agit d'une fusion de la marche et de traits du visage en utilisant l'APC. Les données pour le visage et la démarche sont capturées, puis la combinaison de ces images est effectuée. Les vecteurs caractéristiques issus de la fonction PCA sont combinés. L'analyse en composantes principales est un algorithme utilisé pour réduire la dimension de ces vecteurs caractéristiques. A la phase de test on donne le vecteur d'entrée analogue qui est la combinaison de la fois le visage et la démarche. Les expériences montrent que lorsqu'on utilise la fonction marche seulement de l'individu alors son taux de reconnaissance est de 67%. Mais quand on a combiné la démarche avec le visage, les performances ont pu être améliorées jusqu'à 90%. Ce système peut être utilisé dans les communautés où l'on a besoin de méthode automatisée pour déterminer l'identité de l'individu.

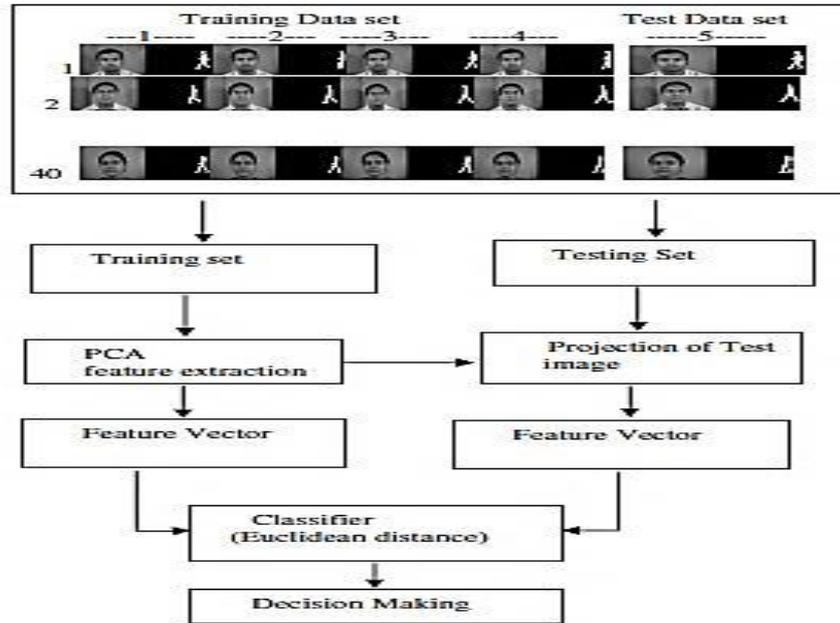


Fig. II. 5 : Système de fusion et démarche pour l'identification [SAN 11]

Des travaux d'actualité présentent une approche intégrée pour le suivi des mains, des visages et des traits du visage dans l'image séquences, destinées à soutenir interaction avec naturelle autonome robots qui naviguent dans les espaces publics et, plus spécifiquement, à fournir des données pour l'analyse des gestes de la main et les expressions faciales que les humains utilisent dans l'exercice de divers états de la conversation avec le robot [HAR12]. Pour la main et suivi du visage, un tracker blob qui est spécialement formé pour poursuivre la peau de couleur régions. Fonction de détection du visage et du suivi a été réalisée par l'intermédiaire de l'emploi l'état de l'art apparence basé sur la détection couplé avec la fonction basée sur le suivi, en utilisant un ensemble de données anthropométriques contraintes. Une approche qui combine le détecteur en cascade stimulé de Viola et Jones avec un tracker basé sur les caractéristiques. Le détecteur de combiné résultant et traqueur s'étend de travaux antérieurs des mêmes auteurs sur la localisation caractéristique du visage [Pat 09]. Un schéma des composants qui composent l'approche proposée est représenté dans la figure.

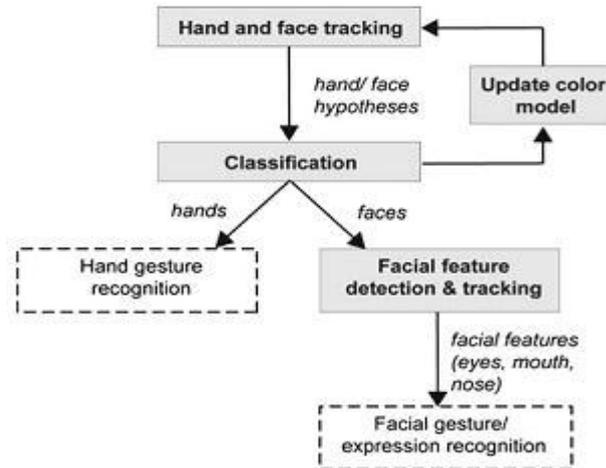


Fig. II. 6 : Le premier bloc dans la Fig. est la main et le traqueur visage [HAR12].

Les résultats expérimentaux ont confirmé l'efficacité de l'approche proposée conduisant à des implémentations qui combinent la précision, l'efficacité et la robustesse. Le but de l'approche de suivi proposé pour faciliter tâche d'interaction homme-robot, mais la méthodologie caractéristique présentée possède ici d'autres tâches. Outre son utilisation pour l'analyse des gestes de la main et les expressions faciales, elle peut être utilisée pour plus de tâches générales de reconnaissance de l'activité et les tâches liées à l'apprentissage de robot par démonstration.

Afin d'améliorer la précision de la reconnaissance du système biométrique une approche de reconnaissance biométrique multimodale basée sur le niveau de fusion fonctionnalité et tenseur courbe est proposé. L'approche tenseur courbe est une extension de la méthode d'analyse tensorielle sur la base de coefficients de l'espace cervelets [GUA12]. On utilise deux sortes de biométrie: reconnaissance d'empreinte palmaire et la reconnaissance du visage. Tous les éléments d'une image sont extraites en utilisant le courbe algorithme de tenseur, puis les caractéristiques normalisées sont combinées au niveau de fusion fonctionnalité en utilisant des plusieurs stratégies de fusion. Le voisin le plus proche de k-(KNN) classificateur est utilisé pour déterminer le classement final biométrique. Les résultats expérimentaux démontrent que l'approche proposée surpasse la solution monomodale et le projet de fusion presque gaussienne (NGF) stratégie a une meilleure performance que d'autres règles de fusion.

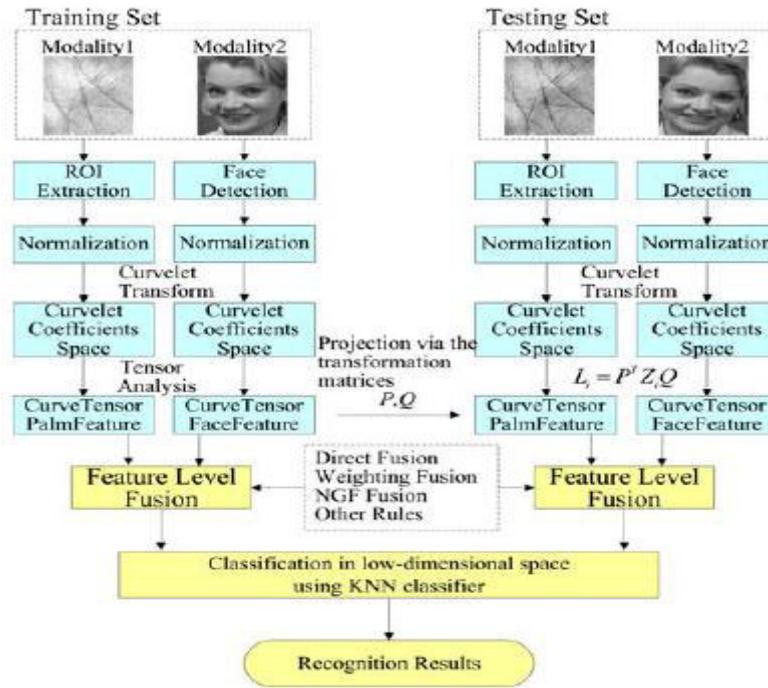


Fig. II. 7 : Schéma de la reconnaissance biométrique multimodal.

Les résultats expérimentaux sur Poly U, CMU-PIE et bases de données ORL démontrent l'efficacité et la robustesse de cette méthode.

II.6 Tableau récapitulatif des travaux récents sur la fusion multimodale.

Auteurs	Année	BDD	Méthode	Résultat
Aud	2006	ACP	Averaged Bayes Classifier(ABC)	99.7%
Mes	2009	ORL	DCT-PCA ,DCT-RLDA	96.75%
jir	2011	GMM-UBM	AC W – C M S	98.98%
San	2011	RFID UWB	PCA	90%
Ami	2011	CASIA V2	PSO	
Ade	2012	TRN	CGD	98.5%
X. XU	2012	PIEP	NGF	99.71%

Tableaux II.1 état de l'art de fusion.

II.7 Conclusion

Dans ce chapitre, nous avons présenté les types et les différentes méthodes de fusion puis nous avons détaillé les états de l'art de la fusion.

III.1 Introduction

La biométrie consiste à reconnaître une personne grâce à des traits distinctifs de son corps. Il existe différentes technologies biométriques telles que l'analyse des empreintes digitales, de la voix, de l'iris, etc.

Ce projet se base sur la fusion de scores entre les visages à deux et trois dimensions. L'algorithme de réduction de données utilisé pour la reconnaissance dans ce travail est l'analyse en composantes principales (PCA). Nous essayons de faire la fusion des scores calculés d'une façon parallèle de chacune des modalités visage 2D et 3D.

III.2 Présentation du modèle

Après notre étude bibliographique et à travers l'état de l'art des travaux sur la fusion. Notre choix s'est porté sur un modèle de reconnaissance de visages basé sur la fusion des scores par combinaison [OUA11]. Mais avant de faire la fusion nous traitons chaque modalité à part. Et bien sûr nous avons jugé de garder les mêmes techniques de réduction d'espace, calcul de métrique et normalisation pour les visages 2D et 3D. Seule la partie prétraitement est particulière à la nature des images mises à l'épreuve. Notre module conçu se présente sous la forme du schéma de la figure III.1.

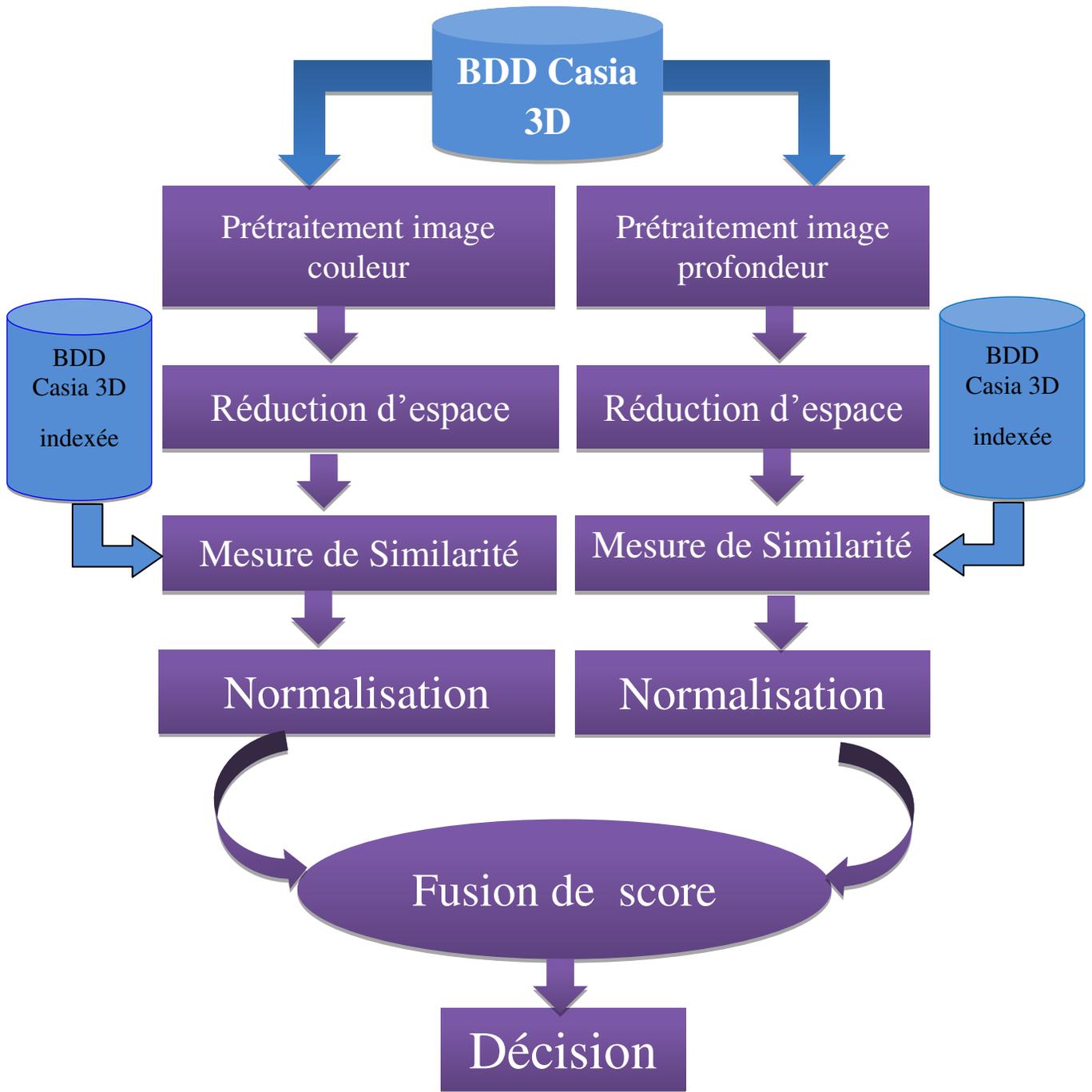


Fig.III.1 : Schéma du modèle proposé

III.2.1 Prétraitement de l'image

III.2.1.1 Découpage de l'image

Le découpage de l'image consiste à conserver les maximums des variations intrinsèques du visage, et de supprimer les autres informations comme l'arrière-plan, les cheveux, les cols de chemise, les oreilles etc.

III.2.1.2 Conversion de l'image

Nous utilisons deux types de prétraitement :

- Extraction de la couleur du visage 3D
- Extraction des cartes de profondeur du visage 3D.

III.2.1.3 Filtrage

1) Ondelettes de Gabor

En 1993, Martin La des a proposé pour la première fois l'utilisation des filtres de Gabor comme outil d'identification de visage. Les ondelettes de Gabor 2D ont la caractéristique d'être modulables par différents paramètres qui contrôlent leurs orientations, fréquences, phases, tailles et aspects. L'équation générale d'une ondelette de Gabor en 2D est complexe:[Hub00]

$$w(x, y, \theta, \varphi, \lambda, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \times e^{i(\frac{x'}{\lambda} + \varphi)} \quad (\text{III.1})$$

Partie réelle :

$$w(x, y, \theta, \varphi, \lambda, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \times \cos^i(\frac{x'}{\lambda} + \varphi) \quad (\text{III.2})$$

Partie imaginaire :

$$w(x, y, \theta, \varphi, \lambda, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \times \sin^i(\frac{x'}{\lambda} + \varphi) \quad (\text{III.3})$$

Où $\mathbf{x}' = x\cos\theta + y\sin\theta$ et $\mathbf{y}' = -x\sin\theta + y\cos\theta$

Il y a donc 5 paramètres qui contrôlent l'analyse par ondelette. Voici une explication des cinq paramètres:

- **L'orientation de l'ondelette (θ)**

Ce paramètre détermine l'angle des contours ou des lignes de l'image aux quelles l'ondelette va être sensible.

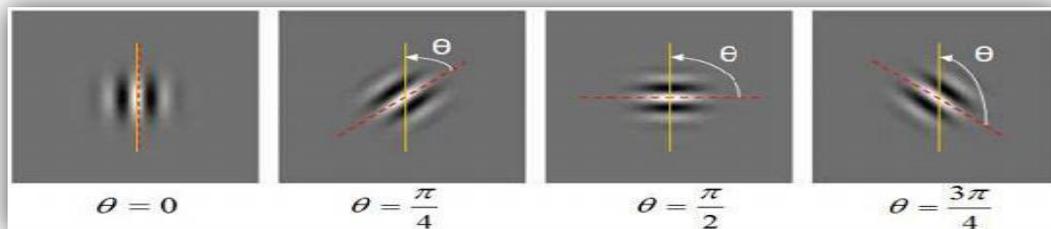


Fig.III.2 : Orientation de l'ondelette.

- **La fréquence centrale de l'ondelette (λ)**

Ce paramètre spécifie la longueur d'onde du cosinus ou inversement la fréquence centrale de l'ondelette (voire la figure). Les ondelettes avec une grande longueur d'onde seront sensibles à des changements progressifs d'intensité dans une image. Les ondelettes avec une petite longueur d'onde seront sensibles à des contours et des lignes abruptes

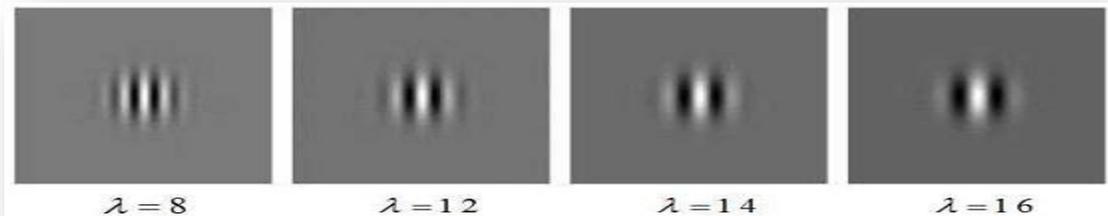


Fig.III.3 : Fréquence centrale de l'ondelette.

- **La phase de la sinusoïde (φ)**

On utilise à la fois la partie réelle et la partie imaginaire de l'ondelette complexe de Gabor.

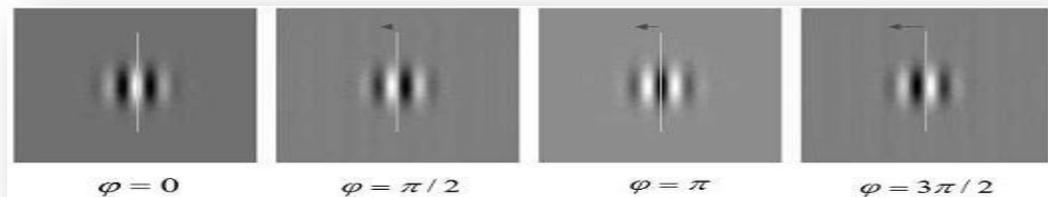


Fig.III.4: Phase de l'ondelette.

- **Le support temporel de l'ondelette (σ)**

Ce paramètre spécifie le rayon de la gaussienne (Fig.III.5). La taille de la gaussienne détermine la quantité de pixels de l'image qui vont être pris en compte dans la convolution.

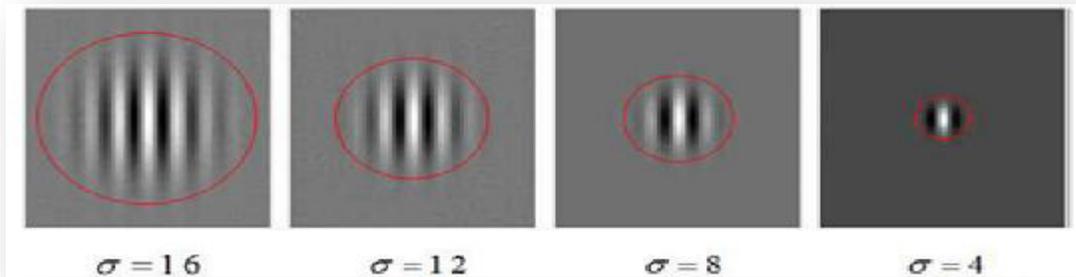


Fig.III.5: Support temporel de l'ondelette.

- **L'enveloppe de la gaussienne (γ)**

Ce paramètre agit sur la forme de l'enveloppe gaussienne, en l'étirant spatialement

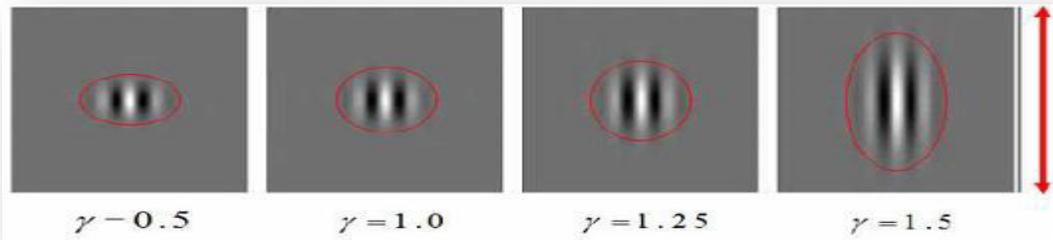


Fig.III.6: Forme de l'enveloppe gaussienne.

Le jeu de ces cinq paramètres permet dès lors une analyse complète de la texture d'une région de l'image. Comme on peut remarquer dans l'équation (Eq III.1), que les ondelettes de Gabor ont une forme complexe: partie réelle et partie imaginaire.[Bev03]

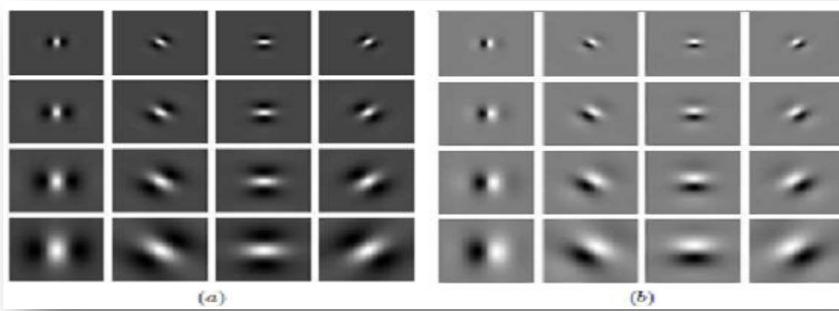


Fig.III.7: Parties réelles (a) et imaginaires (b) d'ondelettes de Gabor

Sur la figure III.7 il s'agit d'ondelettes de Gabor à 4 niveaux de résolution et selon 4 orientations (les colonnes représentent les orientations et les lignes, les résolutions) [Bev03]

III.2.2 Réduction d'espace

Dans notre cas nous réalisons une réduction d'espace par Analyse en composantes Principales (ACP ou PCA). La PCA peut construire un sous-espace vectoriel en ne retenant que les meilleurs vecteurs propres, tout en conservant beaucoup d'information non redondante. C'est pourquoi la PCA est très efficace pour réduire la dimension des données.

III.2.1 Principe de PCA

Pour trouver les principales composantes, on peut réaliser les opérations suivantes:

- Organiser les données en vecteurs colonnes pour obtenir une matrice

$$X \text{ de } (N^1 \times M^2) \quad \Gamma_i \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}; \quad X = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{bmatrix}$$

- Calculer le vecteur moyen de X

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad \text{III.4}$$

Soustraire le vecteur moyen Ψ à toutes les colonnes de X afin d'obtenir des données centrées, la matrice obtenue est A

$$\Phi_i = \Gamma_i - \Psi, \quad A = [\Phi_1, \Phi_2, \dots, \Phi_M]$$

- Calculer la matrice de covariance C de $N \times N$

$$C = A^T A$$

III.5

- Calculer et trier par ordre décroissant les vecteurs propres λ_i de C selon leurs valeurs propres u_i correspondantes dans E de $N \times N$
- Garder les K premiers vecteurs propres λ_i pour définir un espace vectoriel engendré par ces K vecteurs propres, que l'on appelle l'espace \bar{E} de $N \times K$
- Projeter les données de départ sur E . Un vecteur Γ_i est alors transformée en Y dans l'espace réduit par l'opération:

$$Y_i = \bar{E}^T(\Phi_i) \quad \text{III.6}$$

III.2.3 Classification

La Classification est une tâche qui regroupe les instances en groupes ayant une ou plusieurs propriétés communes, dont les groupes sont les « classes » distinguées.

Le classificateur doit estimer une fonction $f(x)$ qui est l'estimation de la fonction qui représente la relation entre l'objet et sa catégorie, cette fonction est appelée fonction de décision : $F(x) = y$

X : L'ensemble des objets à classifié (appelé espace d'entrée).

Y : L'ensemble des catégories (appelé espace de sortie) [Khe06]

III.2.4.1 Types de classification

Plusieurs techniques de classification ont été proposées dans la littérature. Ces méthodes peuvent être subdivisées en deux catégories qui sont la **classification supervisée** et **non supervisée** selon les résultats qu'on veut obtenir et la nature des données.

Le schéma ci-dessous représente les types de classification :

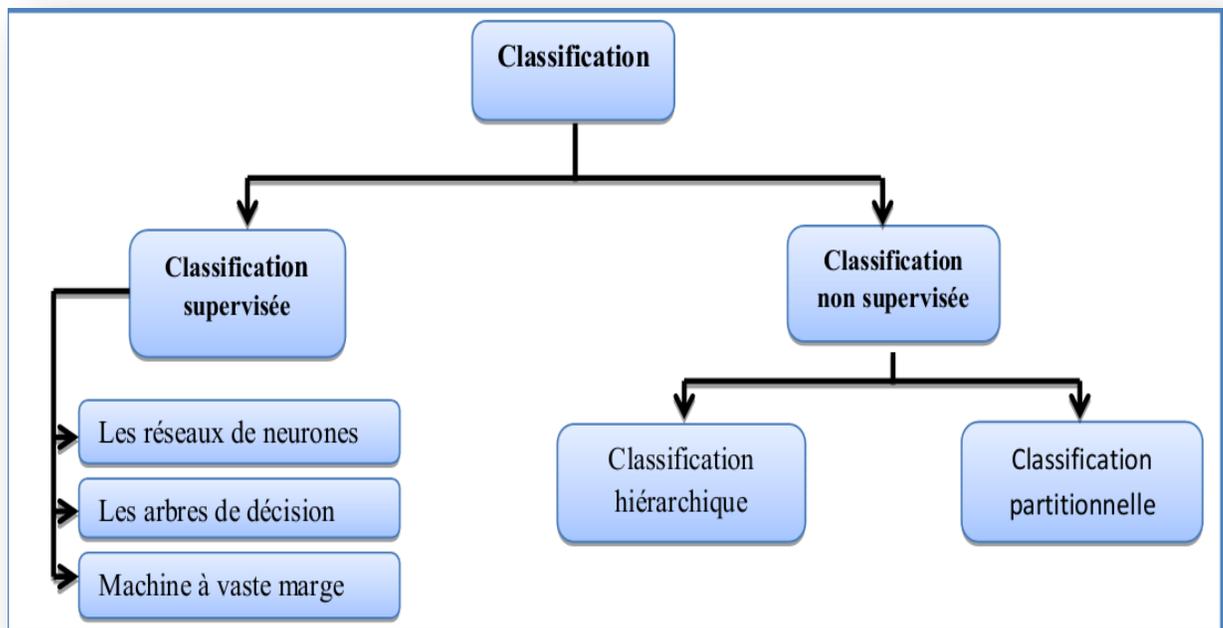


Fig. III.10 : Les méthodes de classification

- **Classification supervisée**

La classification est dite supervisée car on dispose d'exemples déjà étiquetés (étiquetés par une classe). Son but principal est de trouver, à partir des exemples déjà classés (appelé ensemble d'apprentissage), un modèle afin de prédire les classes de nouvelles données [Say10]. Parmi les méthodes de classification supervisée, (SVM)

- **Classification non supervisée**

Contrairement à la classification supervisée, la classification non supervisée ne dispose pas d'un ensemble d'apprentissage. De ce fait, cette dernière cherche à diviser et à classer un ensemble de données de sorte que les classes soient les plus homogènes possibles [Say10].

Il existe deux types de méthode de constitution des groupes (classification non supervisée) : les méthodes dites **hiérarchiques** et les méthodes **partitionnelles**.

Dans le cas de notre application nous devons absolument passer par l'enrôlement du système biométrique. Au cours de cet enrôlement nous procédons à l'apprentissage à notre système de notre base de données clients. Donc, notre classification ne peut être que supervisée. Parmi les classifieurs les plus adaptés à la reconnaissance des personnes en biométrie on trouve le SVM, les réseaux de neurones (RN) qui sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit. Ils constituent des neurones adaptatifs interconnectés par des poids synaptiques [Cla92], la mesure de similarité...

Nous nous limitons au calcul de métriques pour mesurer les similarités dans le visage. Notre choix est justifié par la simplicité de la méthode.

III.2.4 Calcul de métriques

Lorsqu'on souhaite comparer deux vecteurs de caractéristiques issus du module d'extraction de caractéristiques d'un système biométrique, on peut soit effectuer une mesure de similarité (ressemblance), soit une mesure de distance (divergence). La première catégorie de distances est constituée de distances Euclidiennes et sont définies à partir de la distance de Minkowski d'ordre dans un espace euclidien R^N (N déterminant la dimension de l'espace euclidien).

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

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

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

1) Distances Euclidiennes

- **Distance City Block (L1)**

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

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

- **Distance Euclidienne (L2)**

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

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

2) Distances dans l'Espace de Mahalanobis

Avant de pouvoir effectuer des mesures de distance dans l'espace de Mahalanobis, il est essentiel de bien comprendre comment l'on passe de l'espace des images \mathcal{J}_m à l'espace de Mahalanobis \mathcal{E}_{Mah} . En sortie de l'algorithme PCA, nous obtenons des vecteurs propres associés à des valeurs propres (représentant la variance selon chaque dimension). Ces vecteurs propres définissent une rotation vers un espace dont la covariance entre les différentes dimensions est nulle. L'espace de Mahalanobis est un espace où la variance selon chaque dimension est égale à 1. On l'obtient à partir de l'espace des images \mathcal{J}_m en divisant chaque vecteur propre par son écart-type correspondant.

Soient u et v deux vecteurs propres de J_m , issus de l'algorithme PCA, et m et n deux vecteurs de \mathcal{E}_{Mah} . Soit λ_i les valeurs propres associées aux vecteurs u et v , et σ_i l'écart-type, alors on définit $\lambda_i = \sigma_i^2$. Les vecteurs u et v sont reliés aux vecteurs m et n à partir des relations suivantes :

$$m_i = \frac{u_i}{\sigma_i} = \frac{u_i}{\sqrt{\lambda_i}} \text{ et } n_i = \frac{v_i}{\sigma_i} = \frac{v_i}{\sqrt{\lambda_i}}$$

- **Mahalanobis L1 (MahL1)**

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

$$Mah_{L1}(u, v) = \sum |m_i - n_i|$$

- **Mahalanobis L2 (MahL2)**

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

$$Mah_{L2}(u, v) = \sqrt{\sum_{i=1}^N |m_i - n_i|^2}$$

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

- **Cosinus de Mahalanobis (MahCosine)**

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

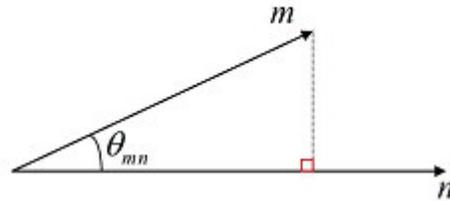


Fig. III.9 – Les deux vecteurs m et n dans l'espace de Mahalanobis

Nous avons donc par définition :

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

De plus, on peut écrire :

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

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

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

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

III.2.5 Normalisation

La normalisation est l'une des transformations les plus utilisées dans le traitement d'image car elle permet d'exploiter toute la gamme de valeurs des niveaux de gris. Elle est donnée par la fonction: $T : [a, b] \rightarrow [0, 255]$

$$x \rightarrow 255 \times \frac{x - a}{x - b}$$

a, b sont respectivement les valeurs minimale et maximale de niveaux de gris que peut prendre les pixels d'une image.

III.2.6 Fusion des scores

Considérons un système de vérification biométrique multimodale qui adopte une approche de fusion par combinaison, au niveau score. Le cadre de travail théorique développé par Kittler et al.[Kitt 98] utilise une approche probabiliste qui peut être appliquée à ce système seulement si la sortie de chaque modalité est de la forme $P\{\text{authentique}|X\}$, c'est-à-dire la probabilité à posteriori qu'un utilisateur soit "authentique" sachant l'échantillon d'entrée biométrique X . En pratique, la plupart des systèmes biométriques donnent en sortie un score de correspondances. Verlinde et al [Ver199] ont proposé que ce score de correspondances soit lié à $P\{\text{authentique}|X\}$ comme suit :

$$s = f(P\{\text{authentique}|X\}) + \eta(X) \quad (\text{III.12})$$

Où f est une fonction mono tonique et η est l'erreur faite par le système biométrique qui dépend de l'échantillon d'entrée biométrique X .

Cette erreur peut être due au bruit introduit par le capteur pendant l'acquisition du signal biométrique et les erreurs faites par l'extraction de caractéristiques et les processus de correspondance.

Si l'on suppose que $\eta = 0$, il est raisonnable d'approximer $P\{\text{authentique}|X\}$ par $P\{\text{authentique}|s\}$. Dans ce cas, le problème revient à calculer $P\{\text{authentique}|s\}$ et cela requiert l'estimation des densités conditionnelles $p(s|\text{authentique})$ et $p(s|\text{imposteur})$.

Les techniques de fusion de Snelick et al.[Sne03] utilisent l'approche probabiliste de Kittler et al. en se servant d'un ensemble de données d'entraînement composé de 100 personnes afin d'estimer les probabilités a posteriori des authentiques $P\{\text{authentique}|s\}$.

Pour ce faire, ils utilisent la moyenne et la variance des scores authentiques et imposteurs provenant de cet ensemble de données d'entraînement en supposant une distribution normale pour leurs densités conditionnelles respectives $p(s|\text{authentique})$ et $p(s|\text{imposteur})$.

Cependant, leur approche possède deux inconvénients majeurs. Tout d'abord, l'hypothèse d'une distribution normale pour les scores n'est pas valide dans certains cas, en particulier pour les scores authentiques. D'autre part, cette approche n'utilise pas les probabilités à priori des utilisateurs authentiques et imposteurs qui pourraient être disponibles au sein du système. En effet, la formule que Snelick et al. Utilisent pour estimer $P\{\text{authentique}|s\}$ [Sne03] n'est vraie que dans le cas où l'on considère les probabilités a priori (c'est-à-dire les fréquences des classes) équiprobables; dans le cas contraire, on doit directement appliquer le Théorème de Bayes avec des densités de probabilités (III.13):

$$P\{\text{authentique}|s\} = \frac{p(s|\text{authentique}) * P\{\text{authentique}\}}{p(s|\text{authentique}) * P\{\text{authentique}\} + p(s|\text{imposteur}) * P\{\text{imposteur}\}} \quad (\text{III.13})$$

Pour ces raisons, une méthode d'estimation de densité par fenêtre de Parzen a été proposée dans [Dud01] ; cette technique permet d'estimer la véritable densité conditionnelle des scores authentiques et imposteurs. Ainsi, bien que la technique d'estimation de densité par fenêtre de Parzen réduise de manière significative l'erreur sur l'estimation de $P\{\text{authentique}|s\}$ (en particulier lorsque les densités conditionnelles ne sont pas gaussiennes), l'estimation de densité possède toujours des inexactitudes dues à la taille finie de l'ensemble des données d'entraînement et aux problèmes liés au choix de la largeur de fenêtre optimale pendant le processus d'estimation de densité. De plus, l'hypothèse selon laquelle la valeur de η dans l'équation (III.12) est nulle n'est pas valide, d'un point de vue pratique, dans la plupart des systèmes biométriques. Puisque η dépend de l'échantillon d'entrée biométrique X , il est possible d'estimer η seulement si le système biométrique donne une mesure de confiance (qui prend en compte la nature de l'entrée X) sur le score de correspondance en plus du score lui-même. En l'absence de cette mesure de confiance, la valeur calculée de $P\{\text{authentique}|s\}$ n'est pas un bon estimateur de $P\{\text{authentique}|X\}$ et cela peut amener à une faible performance de reconnaissance du système multimodal.

III.2.6 Décision

Dans un processus d'authentification, la stratégie de décision nous permet de choisir entre les deux alternatives suivantes : l'identité de l'utilisateur correspond bien à l'identité de la personne proclamée (ou cherchée) ou elle ne correspond pas. En comparant le modèle de l'utilisateur à celui qui existe dans la base de données. Cette phase est basée sur le seuil prédéfini.

III.3 Conclusion

Ce chapitre contient le modèle de fusion 2d et 3d pour SRV avec base de données Casia 3d qui se divise en deux parties la première partie inclut le prétraitement image couleur et la deuxième partie inclut le prétraitement image profondeur et toutes les deux parties fonctionnent de la même manière (réduction de l'espace, mesure de similarité et normalisation) après fusion des scores, finalement la décision, chaque étape qui nous est détaillée dans ce chapitre. En va voir la conception et le résultat dans le chapitre suivant.

IV.1 Introduction

Notre but est d'étudier la fusion des scores entre les images 3D couleur et 3D profondeur. à partir de la base de données CASIA 3D, et de classifier des personnes particulières pour les identifier en temps réel (par exemple dans un système de surveillance (sécurité) et pour permettre certain accès (par exemple accès à une banque, à un ordinateur etc...)).

Dans ce chapitre nous allons présenter les différentes étapes de la conception de notre système. Nous développons en suite la partie d'extraction de paramètre.

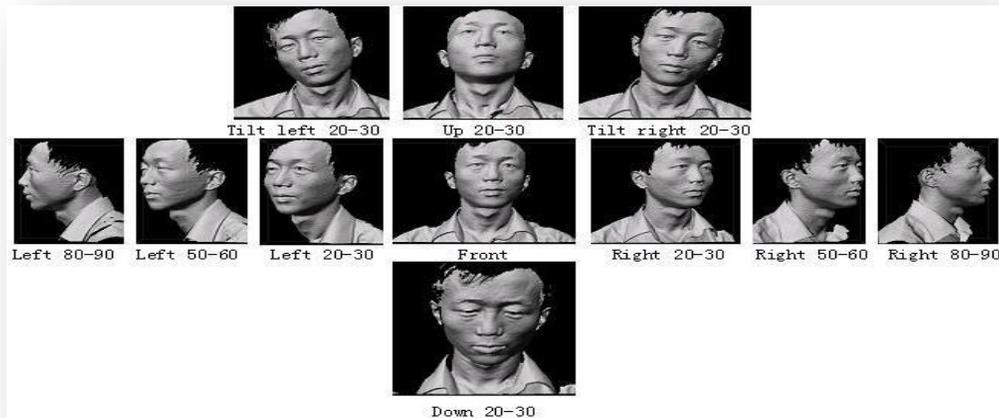
Finalement, on fusionne les scores extraits à partir de ces deux technique de conversion étudié est appliqué sur les images des visages de la base de données CASIA 3D, et les résultats obtenus sont exposés.

IV. 2 La base de données CASIA 3D face VI

Entre Août 2004 et Septembre 2004, nous avons recueilli une base de données de visages 3D constitué de 4624 scans de 123 personnes qui utilisent le sans contact numériseur 3D. Au cours de la construction de la base de données, nous considérons non seulement les variations individuelles de poses, les expressions et les illuminations, mais aussi les variations combinées des expressions sous illumination et pose sous des expressions, comme montré sur les figures. Pour les sujets avec des lunettes, nous allons recueillir un scan supplémentaire avec des lunettes. Ainsi, chaque personne dispose de 37 ou 38 balayages. Et à partir de chaque balayage, une image couleur 2D et une surface 3D du visage triangulaire sont également générés. Nous vise à construire une base de données complète du visage en 3D, qui est en outre amenés à être une plateforme publique pour tester les algorithmes de reconnaissance faciale 3D ou autres.



FigIV.1 : Variation d'éclairage de la base de données CASIA visage 3D



FigIV.2 Variation de pose de la base de données CASIA visage 3D



FigIV.3 : Variation d'expression de la base de données CASIA visage 3D

IV.3 Architecture du système de conception

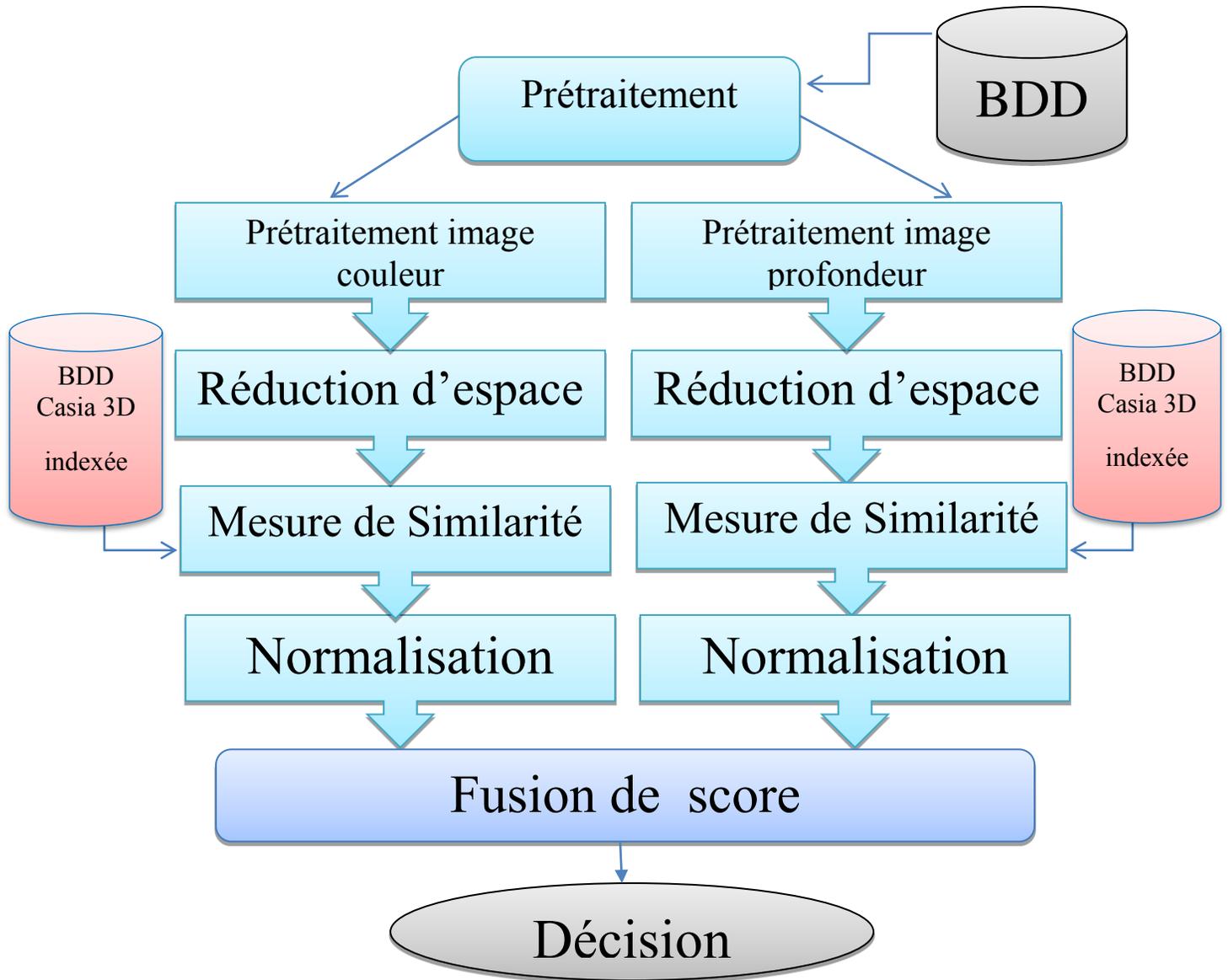


Fig. IV.4 : Architecture structurelle du système de conception.

IV.3.1 Acquisition d'image

Les images de visage



Fig.IV.5 : Capturer des scènes de la base de données CASIA visage 3D

IV.3.2 Prétraitement de visage

En projetant l'image 3D couleur sur un espace plan nous obtenons notre image couleur 2D. Par un programme de prétraitement nous obtenons la carte profondeur appelée aussi l'image 2.5D. Ces deux images après avoir subies des traitements séparés seront fusionnées au niveau des scores.

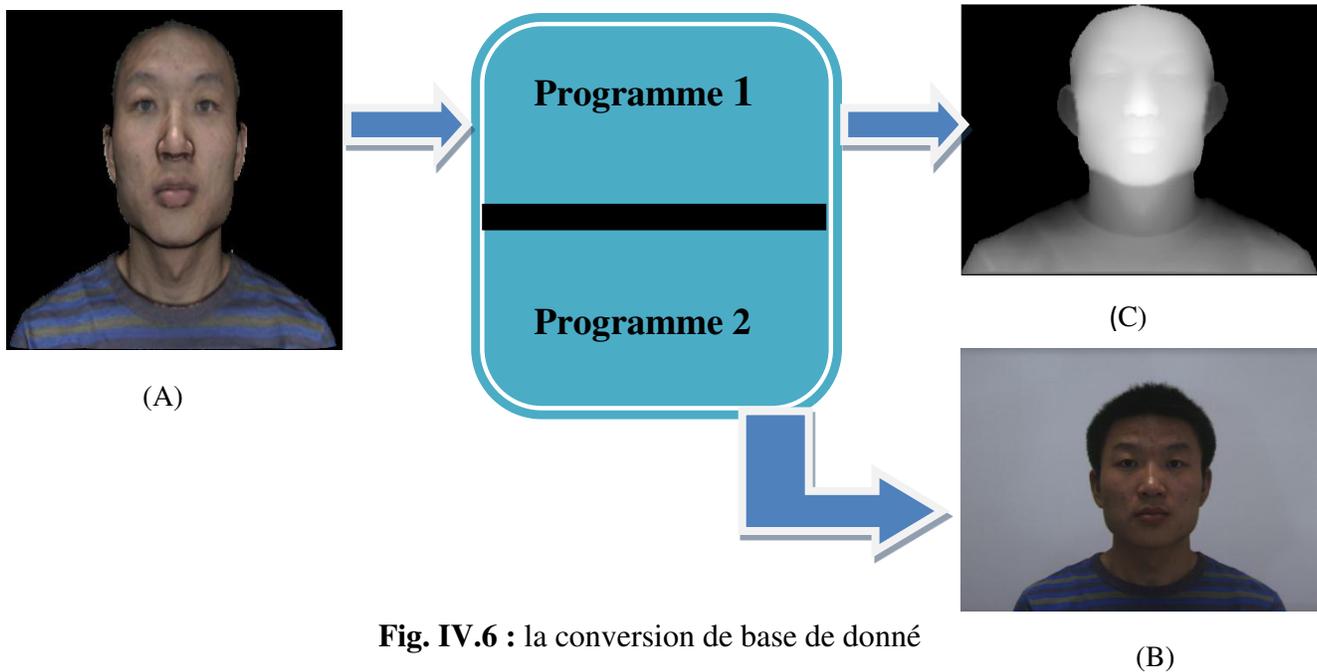


Fig. IV.6 : la conversion de base de donné

(A) image 3D, (B) image de profondeur, (C) image RGB

Dans cette figure nous avons fait deux programmes de conversion

Programme 1 : la conversion d'image de visage 3D à 2D profondeur (2.5).

Programme2 : la conversion d'image de visage 3D à 2D couleur.

IV.3.3 Prétraitement d'image :

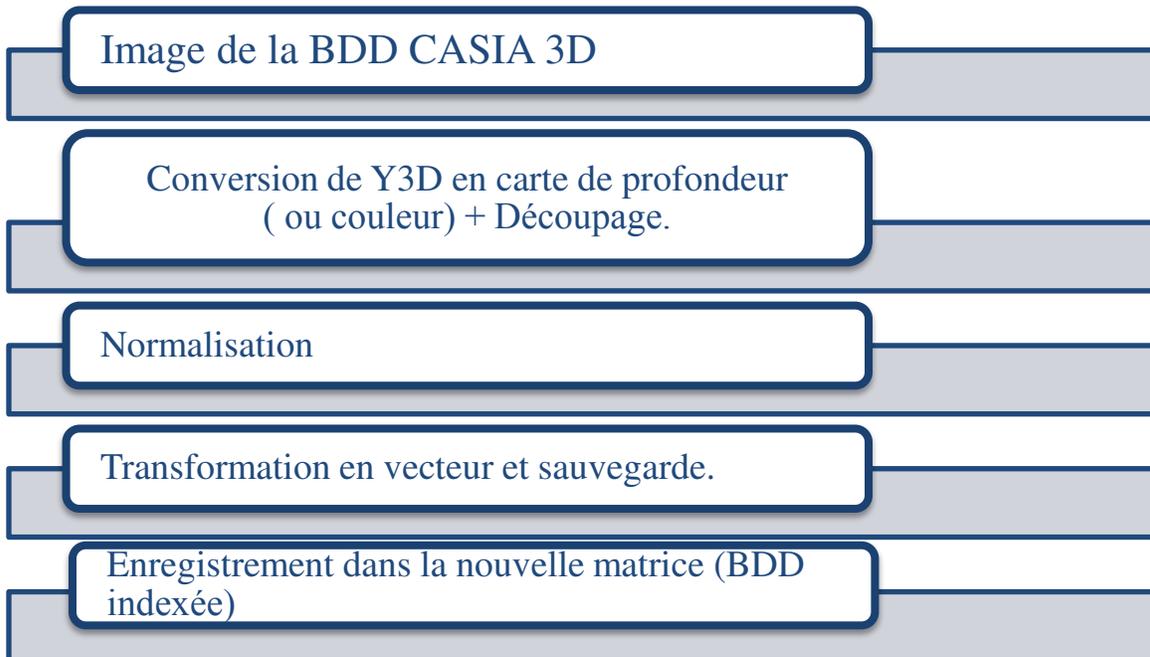


Fig. IV.7 : Schéma de prétraitement des images de visage 3D.

IV.4 Réduction d'espace

Dans la partie de réduction d'espace nous avons proposé l'algorithme de l'ACP pour sa simplicité et apport en information, les étapes de cet algorithme est présenté comme suit:

IV.4.1L'analyse en composantes principales(PCA)

- *Prétraitement.*

- *Création de la matrice des données X (L'ordonnement des vectrices images des clients coté à coté).*
- *Centrer la matrice des données X*
- *Calcul de la matrice de covariance des données XX^T .*
- *Calcul des valeurs et vecteurs propres de la matrice de covariance.*
- *Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes.*
- *La matrice de transformation de la PCA est les m premiers vecteurs propres (UPCA).*

IV.5 Mesure de similarité

Nous avons deux types de distance soit pour les visages :

La distance intra '**distance_intra()**': *mesure la distance entre deux images de (visage 2D/visage 3D) de la même personne.*

- La distance extra '**ditance_extra()**': *mesure la distance entre deux images de (visage 2D/visage 3D) de deux individus différents.*

Les distances que nous avons utilisées dans notre expérience pour la mesure similarité sont de la distance Euclidienne (**norme L1**), distance Euclidienne (**norme L2**), **Covariance**, elles sont définies comme suit :

$$L1 = \sum_{i=1}^N |Ai - Bi|$$

$$L2 = \sum_{i=1}^N (Ai - Bi)^2$$

$$cov(A, B) = \frac{A}{\|A\|} - \frac{B}{\|B\|}$$

IV.6 Normalisation

La photo normalisation a un double effet : d'une part elle supprime pour tout vecteur image x un éventuel décalage par rapport à l'origine et ensuite elle supprime tout effet d'amplification (multiplication par un scalaire). Pour chaque image on effectue l'opération suivante :

$$\text{Photo normalisation}(x) = \frac{x - \text{mean}(x)}{\text{std}(x)}$$

mean(x) : Le moyen de x

std(x): L'écart type de x .

IV.7 La fusion

Dans cette étape nous réalisons la fusion des scores, ces derniers sont représentés par les distances calculées précédemment, la méthode de fusion que nous utilisons est la fusion par combinaisons simples des scores basés sur les calculs de moyenne, max et min. Nous utilisons des instructions disponibles dans Matlab, cette instruction fusionne deux matrices ou deux vecteurs, Nous n'avons utilisé que les méthodes simples qui sont présentées comme suit :

'**max**' : $D = \text{abs}(A) \text{ abs}(B)$; $C = A(D) + B(\sim D)$

'**min**' : $D = \text{abs}(A) \text{ abs}(B)$; $C = A(D) + B(\sim D)$

'**mean**' : $C = (A+B) / 2$; $D = \text{ones}(\text{size}(A))$

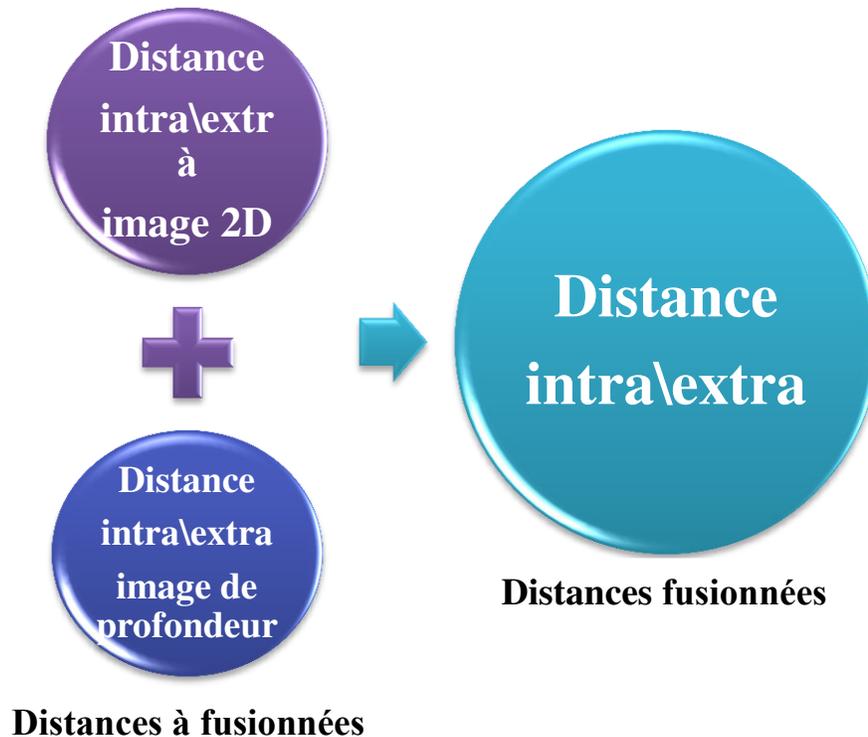


Fig.IV.8 : Fusion des distances

IV.8 Décision

Dans un processus d'authentification, la stratégie de décision nous permet de choisir entre les deux alternatives suivantes: l'identité de l'utilisateur correspond bien à l'identité de la personne proclamée (ou recherchée) ou elle ne correspond pas. En comparant le modèle de l'utilisateur à celui qui existe dans la base de données. Cette phase est basée sur le seuil prédéfini. L'algorithme suivant explique le processus de la phase de décision :

- Algorithme de processus de la phase de décision

Lire V test, P test

Y1 = prétraitement (XV)

Y2 = prétraitement(XP)

Extraction des paramètres (VP)

DIS_P = mesurer distance entre XV, V test

DIS_V = mesurer distance entre XP, P test

DIS = Fusion entre DIS_V, DIS_P

Seuillage

Si DIS < seuil

Personne connue (identité parfaite)

Sinon

Personne inconnue (identité non parfaite)

Fin

Vtest : image de visage test.(Image de l'utilisateur)

XV \ XP : model de la personne proclamée stocké dans la base des données

DIS_V2D couleur : distance entre XV et Vtest. (Projetées dans le même sous espace)

DIS_V2D profondeur : distance entre XP et Vtest. (Projetées dans le même sous espace)

IV.9 Résultats

En effet, nous avons utilisé la conversion en deux niveaux, la conversion en couleur et l'extraction des cartes de profondeur du visage 3D. Nous effectuons nos expériences sur plusieurs types de fusion et nous gardons le type qui convient à notre étude.

Notre programme principal se présente comme suit :

1-Conversion de l'image

- *Extraction de la couleur du visage 3D*
- *Extraction des cartes de profondeur du visage 3D.*

2- Projection des images sur l'espace PCA (visages 2D et visages 2D de profondeur).

3- Mesure des distances (visages 2D et visages 2D de profondeur): Euclidienne norme 1 (L1), Euclidienne Norme 2(L2), covariance.

4- Fusion des scores (distances des visages 2d et distances des visages de profondeur).

IV.9.1 Extraction des paramètres

Nous réalisons plusieurs types de fusion par combinaison des scores (mean, max, min) et les résultats obtenus sont recensés sur le tableau suivant :

- Nous étudions d'abord le cas où le nombre de paramètres $N(\text{PCA})$ est varié de **20** à **200**. Les résultats obtenus sont présentées sur le tableau suivant :

Norme 1(L1)

M N	Mean			Min			Max		
	TFA	TFR	TR	TFA	TFR	TR	TFA	TFR	TR
20	20.60	21	79.20	21.40	21	78.80	21.5	25.5	74.74
40	20.84	21	79.08	21.55	22	78.22	27.56	28	72.22
60	21	20	79.50	21.45	21	78.77	28	28.5	72
80	21	21	79	21.5	21.5	78.50	28.65	29	71.42
100	21.14	21.5	78.68	21.4	21	78.80	29	29.5	71
120	21.29	21.5	78.60	21.3	21	78.85	29.5	30	70.50
140	21.05	21	78.97	21.29	21	78.85	30	30	70
160	21.45	21	78.77	21.25	21.5	78.85	30	30	70
180	22	22	78	21.21	21.5	78.62	30.5	30.5	69.50
200	22	22	78	21.16	21.5	78.67	30.5	30.5	69.50

TableauIV.1 : Taux d'erreurs de norme L1

Dans ce tableau nous constatons que le résultat obtenu avec la méthode de fusion MEAN est le meilleur avec un TFA de l'ordre de 20.60% avec taux de réussite TR=79.50% en utilisant seulement 60 paramètres de l'ACP.

- Taux de réussite $TR = (100 - (TFR + TFA))$
- Taux égale d'erreur $TEE = TFA + TFR / 2$

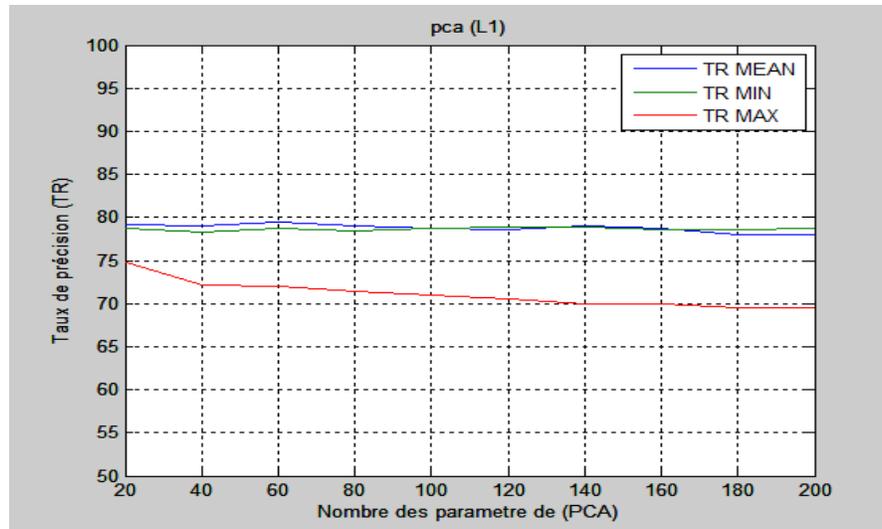


Fig. IV.9: Taux de réussite (TR) en fonction de N(ACP)

La figure suivante illustre une courbe de trois taux de précision (mean couleur bleu ,et min couleur vert et max couleur rouge) on constate que TR mean et TR min sont stables proche de la valeur 80% . Alors que dans le cas TR max la stabilité de ce paramètre se trouve aux environs de 75% .

Nous pouvons dire que la fusion mean est maintenue dans le cas de la distance L1.

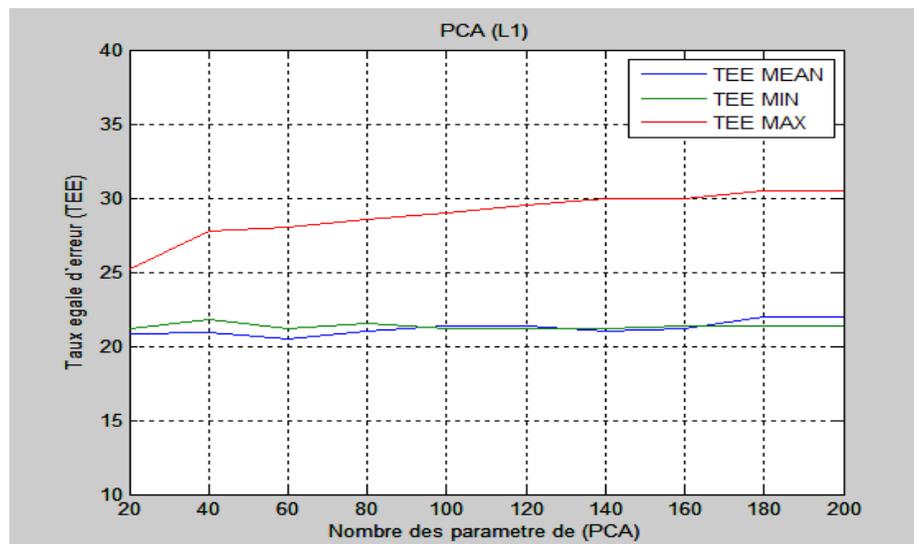


Fig. IV.10 : Taux d'égale erreur (TEE) en fonction de N(ACP)

Norme L2

M N /	Mean			Min			Max		
	TFA	TFR	TR	TFA	TFR	TR	TFA	TFR	TR
20	34.42	34	65.79	21.5	21	78.75	37.26	37.5	62.62
40	35	35	65	22.5	22.5	77.5	37.08	37.5	62.71
60	35.06	35.5	64.72	22.89	22.5	77.30	37.06	37.5	62.72
80	35	35	65	23.33	23.01	76.83	36.84	36.5	63.33
100	35.07	35.5	64.71	23.71	23.66	76.31	39.96	37	61.52
120	35.06	35	64.97	23.99	23.60	76.20	37.03	37	62.98
140	35.06	35	64.92	24.03	23.77	76.10	37	37	63
160	35.07	35.5	64.71	24.50	24	75.75	37.02	37	62.99
180	35.09	35.5	64.70	24	24	76	36.98	37	63.01
200	35.11	35	64.94	24.5	24.22	75.64	37.09	37	62.95

Tableau IV.2 Taux d'erreurs pour la norme L2

Dans ce tableau aussi nous remarquons que la différence entre les erreurs pour les méthodes MEAN et MAX est importante. La méthode de fusion MIN donne le meilleur résultat. Nous constatons que le résultat obtenu avec la méthode de fusion MEAN est le meilleur avec un TFA de l'ordre de 21% avec taux de réussite TR=78.75% en utilisant seulement 20 paramètres de l'ACP.

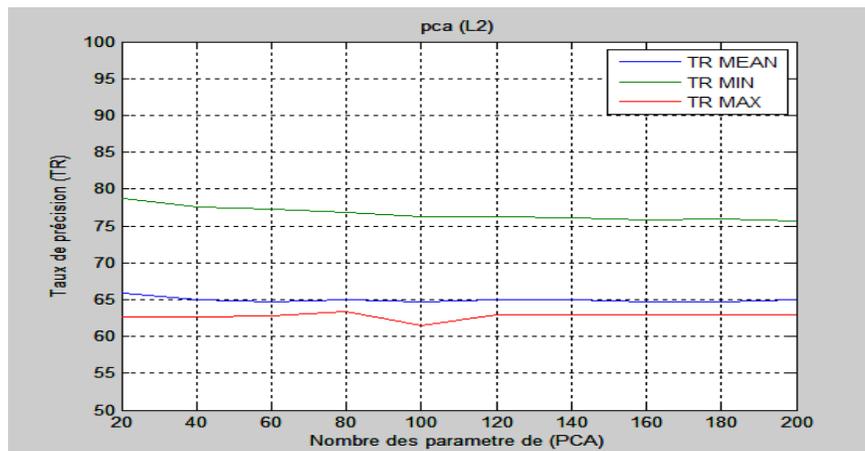


Fig. IV.11 : Taux de réussite (TR) norme L2

D’après cette figure on voit que TR MIN donne un bon taux de réussite par rapport aux TR(TR MAX,TR MEAN).

On peut dire que TR MIN est retenu dans le cas de la norme L2 avec un $N(PCA)=20$,

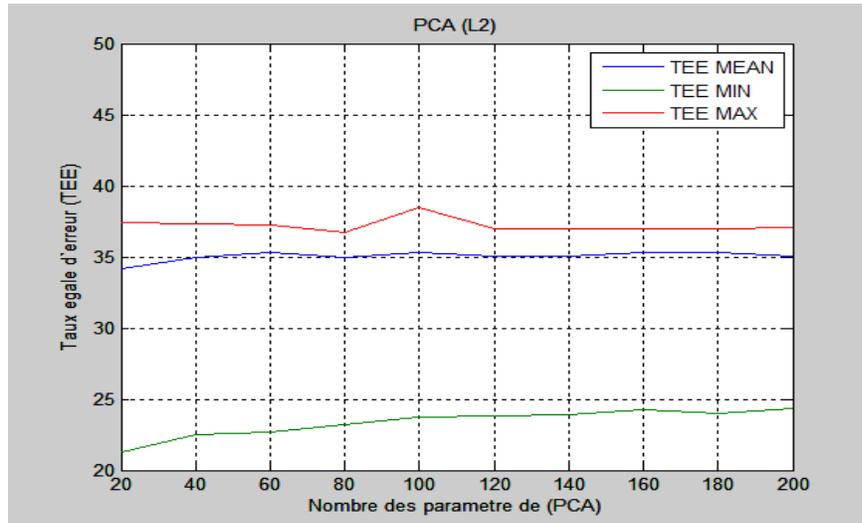


Fig. IV.12 : Taux d’égale erreur(TEE) norme L2

Covariance L3

M \ N	Mean			Min			Max		
	TFA	TFR	TR	TFA	TFR	TR	TFA	TFR	TR
20	38.74	37.74	61.76	39.74	39.74	60.26	35.87	37.25	63.44
40	35	34	65.50	36	36	64	35.50	37.02	6374
60	33.24	33.24	66.76	34.24	35.26	65.25	35.25	36.65	64.05
80	28	27	72.5	28	34.01	68.99	34.28	36.25	64.73
100	27.24	26.24	73.26	27.33	34.25	69.21	34	36.05	64.97
120	26.56	26.20	73.62	27	32.65	70.17	33.27	35.69	65.52
140	26.03	26.01	73.98	26.88	30.20	71.46	33	35.09	65.95
160	25.01	25.32	74.83	25	28	73.50	32.14	34.55	66.65
180	23.24	24.24	76.26	25.24	25.25	74.75	31.68	34.12	67.10
200	23	24.09	76.45	23.24	23.25	76.75	31.23	33.80	67.48

Tableau IV.3 : Taux d’erreurs et de réussite pour la covariance L3

Dans ce tableau le bon résultat de TR pour la COV L3 est donné par la méthode MEAN, TFA=23% et TFR=24.09%, TR=76.75% avec N=200.

La covariance L3 ne présente pas un bon résultat relativement aux autres normes.

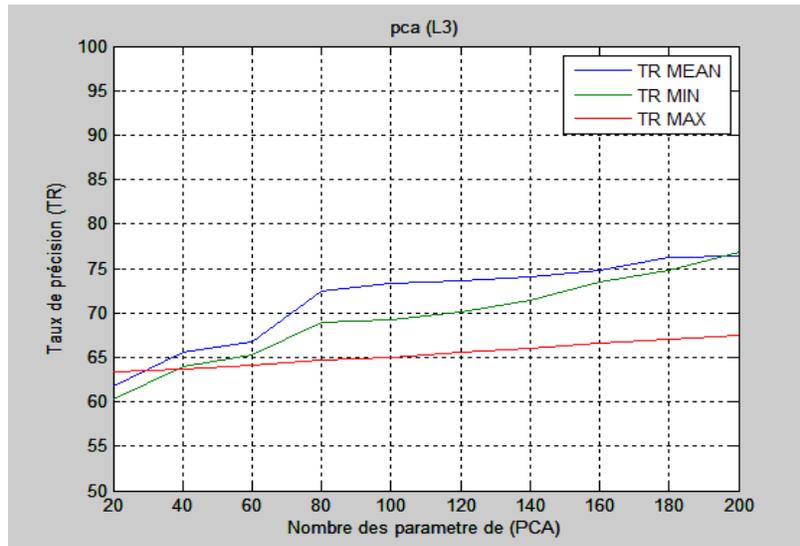


Fig. IV.13 : Taux de réussite(TR) Cov L3

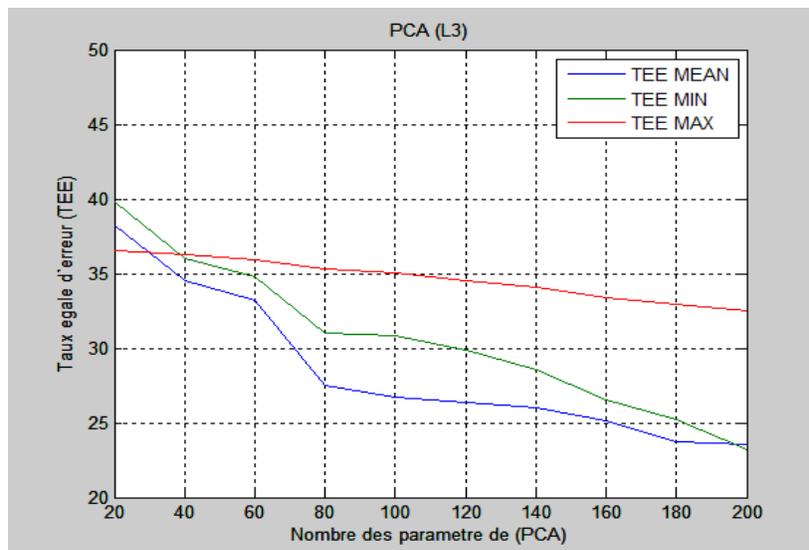


Fig. IV.14 : Taux d'égale erreur(TEE) Cov L3

Nous présentons la comparaison de TR entre les normes : L1, L2 et covariance L3 :

	Mean	Min	Max
L1	79.50	78.85	74.74
L2	65.79	78.75	63.33
COV	76.45	76.75	67.48

Tableau IV.4 : Comparaison de TR pour L1, L2 et COV

D'après ces résultats nous pouvons dire que dans le cas de la norme L1 la méthode de fusion par moyenne pour les images de visages 2D couleur et profondeur donne le meilleur taux de précision **79.50%** avec le nombre de paramètre égal à **N= 60**. Par contre la fusion par min donne un taux égal à **78.80%** avec le nombre de paramètres égal à **N= 20**.

Malgré que le taux d'égale erreur reste élevé pour dire que notre système est stable. Mais nous pouvons affirmer que c'est un bon résultat.

IV. 10 Conclusion

D'après l'étude faite et toutes les expériences réalisées nous pouvons dire que la fusion des scores améliore certes les performances du système. La norme L1 est la meilleure pour le calcul de scores par rapport à la norme L2 et la covariance L3. La fusion des scores par la moyenne est la méthode qui a donné les meilleurs résultats. Ceci dit une étude plus poussée sur les distances reste envisageable. Des améliorations au niveau du prétraitement sont aussi à tenir en compte. Ainsi qu'une meilleure méthode de réduction.

Conclusion générale

Jusqu'à maintenant il n'existe pas de système de reconnaissance de visages performant à 100%, par la multitude de contraintes à qui l'image du visage est confrontée, ce qui altère l'information donnée aux techniques de reconnaissances et donc non stabilité des performances.

Dans ce travail, il était question de réaliser un système biométrique de reconnaissance de visages 3D et 2D basé sur la fusion des scores. La fusion de ces deux systèmes biométriques était une fusion parallèle en scores, avec des différentes méthodes de combinaison pour les comparer. L'évaluation du système est effectuée sur une base de données standard à savoir la CASIA-3D V1. La clarté des textures de ses images de visages ainsi que ses différentes poses, illumination, mouvements de la tête...était le principal motif de notre choix.

Plusieurs tests ont été effectués sur cette base pour étudier les paramètres pertinents du système et ressortir les valeurs optimales. Ces tests ont révélé l'efficacité de l'utilisation de la méthode de filtrage par rapport à la méthode directe. Ceci s'explique par le fait que la méthode de filtrage n'utilise que l'information utile (information discriminante).

Les résultats de ces tests ont montré l'apport de la fusion aux performances, néanmoins on a constaté que la supériorité d'une telle technique sur une autre dépend étroitement des caractéristiques de la base de données étudiée.

Dans ce mémoire nous avons étudié la mise en place et l'évaluation de systèmes biométriques multimodaux. Après une présentation générale du contexte de ce mémoire, la biométrie multimodale, ses caractéristiques, ses avantages et ses inconvénients, les détails de la reconnaissance des visages 3D et 2D, effectuée au chapitre 1, nous avons présenté les différentes méthodes de fusion dans le chapitre 2.

Ensuite, nous avons présenté les étapes de la réalisation de notre système, les méthodes utilisées, les différentes techniques d'extraction des paramètres et les types de calculs des distances, puis les types de fusion des scores.

Tous les résultats obtenus lors de notre expérience sont discutés, en commençant par les résultats de prétraitement des images de visages 3D et 2D puis

l'extraction des paramètres et le calcul des distances, puis la fusion des distances et le calcul des erreurs et des taux de précision.

Après plusieurs investigations et essais nous obtenons un taux de précision **TR=79.50%** et un **TEE= 21.5%** pour la fonction de mesure de score l'Euclidien norme **L1**, et cela avec un **N (ACP) = 60** pour la reconnaissance de visage 3D et 2D. Ce résultat est acceptable dans les SRV à 3D (Système de Reconnaissance de Visages 3D) c'est un modeste apport aux travaux antérieurs. Nous pensons que des améliorations restent à réaliser notamment dans les étapes de détection et de prétraitement 3D. D'autres techniques de classification et fusion peuvent être abordées pour assurer la bonne performance et la stabilité du système SRV 3D.

Perspectives

Pour améliorer de plus les performances du système d'identification nous proposons plusieurs types d'amélioration :

- Améliorer les méthodes d'acquisition des images pour l'identification et/ou l'authentification.
- Améliorer la détection.
- Améliorer le prétraitement des images pour l'identification et l'authentification
- Améliorer la classification.
- Appliquer plusieurs modèles de fusion sur des différents niveaux pour chaque système biométrique.

Bibliographie

- [AHC 09] A. Chaari, thèse, Reconnaissance de visages par réseaux d'ondelettes de Gabor, pour obtenir le grade de docteur, thèse, Université de Sfax, 2009
reconnaissance des personnes, pour obtenir le diplôme master, 2011.
- [AHO 04] Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J.(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
- [AMB 06] Adjout Mohamed & Benaïssa Abdelhak, Institut National de formation en Informatique (I.N.I) Oued-smar Alger. Thème Fusion de la DCT-PCA et la DCT-LDA appliquée à la reconnaissance de visages. pour obtenir un diplôme de l'ingénieur d'état en informatique. 2006.
- [ARO 11] Alismail Mohamed Raouf & Ourchani Nor Elhouda, Université Mohamed Khider Biskra, thème fusion multimodale des scores pour la reconnaissance de personne. Diplôme master. 2011.
- [AUD 06] Audrey POINSOT, Fan YANG, Michel PAINDAVOINE, Laboratoire LE2I CNRS 5158, Faculté des Sciences Mirande, Université de Bourgogne, 21078 Dijon, France 2006.
- [BAR 02] Bartlett, M., Movellan, J., Sejnowski, T.: Face Recognition by Independent Component Analysis. IEEE Trans. Neural Networks 13(6), 1450–1464 (2002)
- [BBA06] Boulbaba BEN AMOR, Ecole Doctorale Informatique et Information pour la Société.(Lyon), thèse Contributions à la reconstruction, la reconnaissance et l'authentification faciale 3D. pour obtenir le grade de Docteur. 2006.
- [BEB 06] G. N. Bebis, A. Gyaourova, S. Singh, and I. T. Pavlidis. Face recognition by fusing thermal infrared and visible imagery. Image and Vision Computing, 24(7) July 2006.

- [BEV 03] J. Beveridge, D. Bolme, M. Teixeira, and B. Draper. “The CSU Face Identification Evaluation System User’s Guide : Version 5.0”. Tech.Rep., Colorado State University, May 2003.
- [CER 07] Cerrada, C., Salamanca, S., Perez, E., Cerrada, J.A., Abad, I.: Fusion of 3D Vision Techniques and RFID Technology for Object Recognition in Complex Scenes. In: IEEE Int. Symp. on Intelligent Signal Processing, WISP 2007, October 3-5, pp. 1–6 (2007)
- [CHA 03] K. Chang, K. Bowyer, and V. Barnabas, “Comparison and combination of ear and face images in appearance–based biometrics”, *IEEE T. Pattern Anal.* **25**, 1160–1165 (2003).
- [CHE 05] W. Chen, M. J. Er and S. Wu, “PCA and LDA in DCT domain”, *Pattern Recognition Letters*, Volume 26, Issue 15, 2005, pp. 2474 – 2482.
- [CLA 92] ClaudeTouzet « Les réseaux de neurones artificiels» Cours, Juillet 1992.
- [CRI 04] Cristinacce, D., Cootes, T.F., Scott, I.: A multi-stage approach to facial feature detection. In: 15th British Machine Vision Conference, pp. 231–240 (2004).
- [DUD 01] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley& Sons, 2001.
- [DEU] Deutscher, J., Blake, A., Reid, I.: *Articulated body motion capture*
- [ELB 07] E. Mohamed & B. Mohamed, thèse, Conception et réalisation d’une plateforme biométrique multimodale basée sur la fusion en scores. Pour l’obtention du diplôme d’ingénieur. 2007-2008.
- [ETE 97] Etemad, K., Chellappa, R.: Discriminant Analysis for Recognition of Human Face Images. *Journal of the Optical Society of America A* 14, 1724–1733 (1997).
- [GEA 10] <http://geai.univ-brest.fr/~carpentier/2009-2010/Ana-mult-4-2010.pf>.

00:30 ; 26/04/2012

- [GET09] M. Gettler Summa, C. Pardoux « La classification automatique », Université Paris Dauphine Ecole Doctorale de Gestion.2009.
- [GUA 12] Feature fusion of palmprint and face via tensor analysis and curvelet transform School of Electronics and Information Engineering, Xi'an Jiaotong University, 28 Xian'ning West Road, Xi'an 710049, People' s Republic of China 2012
- [GUE11] GUEDDOUH Soumia «*L'apprentissage des SVMs incrémentales* » Mémoire en vue de l'obtention du diplôme d'Ingénieur d'Etat en Informatique, juin 2011.
- [HAN 09] Hanifi Majdoulayne « Extraction de caractéristiques de texture pour la classification d'images satellites », mémoire en vue de l'obtention du doctorat en informatique et télécommunications, université toulouse ,2009.
- [HAO 06] Hao, Anderson, and Daugman. Combining crypto with biometrics efficiently.IEEE Transactions on Computers, 55, 2006.
- [HAR 12].Haris Baltzakis · Maria Pateraki · Panos Trahanias Visual tracking of hands, faces and facial features of multiple persons2012.
- [HER 98] Herpers, R., Sommer, G.: An attentive processing strategy for the analysis of facial features. In: Face recognition: From Theory to Applications, pp. 457–468 (1998).
- [HOG 11] Hong Pan, Siyu Xia, Lizuo Jin, and Liangzheng Xia Efficient Face Recognition Fusing Dynamic Morphological Quotient Image with Local Binary Pattern School of Automation, Southeast University, Nanjing, China, 210096 2011.
- [HUB 00] B. Hubbard. Ondes et ondelettes. La Saga d'un outil mathématique. Pour la Science, July 2000.
- [HUL 94] T. K. Ho, J. J. Hull, and S. N. Srihari. Decision combination in multiple classifier systems. IEEE Transactions on Pattern Analysis and Machine In-telligence, 16(1) :66–75, January 1994

- [JAI 02] A. K. Jain and A. Ross. Fingerprint mosaicking. In *International Conference on Acoustics, Speech and Signal Processing*, pages IV : 4064–4067, 2002.
- [JAI 07] K. Jain, “Biometric recognition”, *Nature* **449**, 38–40 (2007).
- [JEN 08] R. Jenkins and A.M. Burton, “100% accuracy in automatic face recognition”, *Science* **319**, 435 (2008).
- [JIA 07] Jia, S., Sheng, J., Chugo, D., Takase, K.: Human recognition using RFID technology and stereo vision. In: *IEEE Int. Conf. on ROBIO 2007*, December 15-18, pp. 1488–1493 (2007).
- [JON 02] Jones, M.J., Rehg, J.M.: Statistical color models with application to skin detection. *Int. J. Comput. Vis.* **46**(1), 81–96 (2002).
- [KAW 05] Kawaguchi, T., Rizon, M., Hidaka, D.: Detection of eyes from human faces by hough transform and separability filter. *Electron. Commun. Japan* **88**(5), 29–39 (2005)
- [KHE 06] Khebab Djame, Belkhefha Mohamed Nadir, « Système de détection du visage humain basé sur la méthode SVM »”, Mémoire de fin d’études pour l’obtention du diplôme d’ingénieur d’état en informatique, 2006/2007.
- [KRZ 92] L. Xu, A. Krzyzak, and C. Y. Suen. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(3) :418–435, 1992.
- [MAN10] Manu Carricano, Fanny, Poujol, Laurent Bertrandias « analyse de données avec SPSS », Pearson Education France, Paris 2010 .
- [MAS11] Massimiliano Dibiato^{1,2}, Antonio Buonaiuto¹, Gian Luca Marcialis¹, Fusion of Radio and Video Localization for People Tracking Daniele Muntoni³, Carlo Maria Medaglia², and Fabio Roli¹ ¹ University of Cagliari, Dept. of Electrical and Electronic Engineering 2011

- [MAT 05] Matthieu, A., Crowley, J.L., Devin, V., Privat, G.: Localisation intra-bâtiment multitechnologies: RFID, Wi-Fi, vision. In: Ubimob 2005, Grenoble, France, pp. 29–35 (2005)
- [MAR 02] G.L. Marcialis and F. Roli, “Fusion of LDA and PCA for Face Recognition”, Proceedings of the Workshop on Machine Vision and Perception, 8th Workshop of the Italian Association for Artificial Intelligence (2002).
- [MES 06] Fusion de la DCT-PCA et la DCT-RLDA pour la Reconnaissance de Visages Université Saad Dahlab de Blida Laboratoire Traitement de signal et d’imagerie Route De Soumaa BP 270 BLIDA, Algérie 2006
- [MUH 05] Muhammad ARIF. Fusion de Données : Ultime Etape de Reconnaissance de Formes, Applications à l’Identification et à l’Authentification. Thèse pour obtenir le grade de Docteur de l’université de Tours Soutenue Le 25 novembre 2005.
- [MOO 04] Y. Moon, H. Yeung, K. Chan, and S. Chan. Template synthesis and imagemosaicking for fingerprint registration : An experimental study. In International Conference on Acoustics, Speech and Signal Processing, pages 409–412, 2004.
- [NED 04]. Nedjem Eddine Ayat « Sélection de modèle automatique des machines à vecteurs de support : Application à la reconnaissance d’image de chiffres manuscrits », mémoire en vue de l’obtention du doctorat en génie P.H.D, montréal , janvier 2004.
- [NGU 11] Nguyen, T., Nguyen, V., Kim, H.: Robust feature extraction for facial image quality assessment. In: Chung, Y., Yung, M. (eds.) Information Security Applications. Lecture Notes in Computer Science, vol. 6513, pp. 292–306. Springer, Berlin (2011).
- [OLI 01] Olivier Bousquet « Introduction aux Support Vector Machines (SVM) », Centre de Mathématiques Appliquées, Ecole Polytechnique, Palaiseau, 2001.

- [OUA11]. Ouamane Abdelmalik Etude de la fusion de modalités pour l'authentification en biométrie (visage, voix), université de mohamed khider biskra, pour l'obtention du diplôme de magister en Electronique. 2011
- [PAR 01] Pardas, M., Losada, M.: Facial parameter extraction system based on active contours. In Proceedings of the International Conference on Image Processing (ICIP'01), Thessaloniki, Greece, October 2001, vol. 1, pp. 1058–1061
- [PAT 09] Pateraki, M., Baltzakis, H., Kondaxakis, P., Trahanias, P.: Tracking of facial features to support human-robot interaction. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '09), pp. 3755–3760 (2009)
- [RAC 06] Rachid Benmokhtar and Benoit Huet. Classifier Fusion: Combination Methods For Semantic Indexing in Video Content, Springer-Verlag Berlin Heidelberg pp: 65–74, 2006.
- [RAG 09] B. Raghavendra, R. and Dorizzi, A. Rao, and G.K Hemantha. Pso versus adaboost for feature selection in multimodal biometrics. In Biometrics : Theory, Applications and Systems, 2009
- [RAG 10] R. Raghavendra, B. Dorizzi, A. Rao, and G.H. kumar. Designing efficient fusion schemes for multimodal biometric systems using face and palmprint. Pattern Recognition, 2010.
- [RAG 11] R. Raghavendra, Bernadette Dorizzi, Ashok Rao, and G. Hemantha Kumar. Particle swarm optimization based fusion of near infrared and visible images for improved face verification. Pattern Recognition, 44(2) :401–411, 2011.
- [RIC 08] Richa Singh, Mayank Vatsa, and Afzel Noore. Hierarchical fusion of multi-spectral face images for improved recognition performance. Information Fusion, 9(2) :200–210, 2008.
- [ROS 09] A. Ross. Multibiometrics. In Stan Z. Li and Anil K. Jain, editors, Encyclopedia of Biometrics, pages 967–973. Springer US, 2009.

- [SAM 07] C. Samir. Analyse des déformations des visages 3D utilisant le chemin géodésique dans l'espace des surfaces faciales. Thèse, Université de Lille1 et Institut National des Télécommunications, Lille, 2007.
- [SAN 02] C. Sanderson and K. K. Paliwal. Information fusion and person verification using speech and face information. In Technical Report, 2002.
- [SAN 11] Sanjeev Sharma, Ritu Tiwari, Anupam Shukla, and Vikas Singh. Fusion of Gait and Facial Feature Using PCA ABV-Indian Institute of Information Technology and Management, Gwalior, India 2011.
- [SAY10] M. Sayad, A. Amara «Modèle discriminant pour la classification de documents XML à l'aide des réseaux bayésiens et le noyau de Fisher », Mémoire en vue de l'obtention du diplôme d'Ingénieur d'Etat en Informatique, juin 2010.
- [SIG 09] Sigalas, M., Baltzakis, H., Trahanias, P.: Visual tracking of independently moving body and arms. In: Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS '09), St. Louis, MO, USA, October 2009
- [SIN 08] R. Singh, M. Vatsa, and A. Noore. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition. *Pattern Recognition*, 41(3) :880–893, March 2008.
- [SIT 10] Sitalakshmi Venkataraman, Siddhivinayak Kulkarni, Risk-Based Neuro-Grid Architecture for Multimodal Biometrics, T. Sobh, K. Elleithy (eds.), *Innovations in Computing Sciences and Software Engineering*, Springer Science+Business Media B.V. 2010.
- [SNI 05] Snidaro, L., Micheloni, C., Chiavedale, C.: Video security for ambient intelligence. *IEEE Trans. on Systems. Man and Cybernetics, Part A: Systems and Humans* 35(1), 133–144 (2005)
- [SOB 98] Sobottka, K., Pitas, I.: A novel method for automatic face segmentation, facial feature extraction and tracking. *Signal Process. Image Commun.* 12(3), 263–281 (1998)

- [SOU 08] Souhila Guerfi Ababsa , Universite D'evry Val D'essonne.Thème
Authentification d'individus par reconnaissance de caractéristiques
biométriques liées aux visages 2D/3D obtenir le diplôme de doctorat .
- [STE 07] Stéphane Tuffery « Data mining et statistique décisionnelle : l'intelligence
des données », Edition Techip, Paris, 2007.
- [TAN 07] Tan, X., Triggs, B.: Fusing Gabor and LBP Feature Sets for Kernel-based
Face Recognition. In: Proc. of Analysis and modeling of faces and gestures,
pp. 235–249 (2007)
- [TUR 91] Turk, M., Pentland, A.: Eigenfaces for Recognition. *J. Cogn. Neurosci.* 13,
71–86 (1991)
- [TOU 06] A. Toufik & B. Kamel, Thème, Système de reconnaissance de visage basé
sur les GMM, pour l'obtention du diplôme d'ingénieur. (I.N.I) Oued-
Smar Alger.2006-2007
- [VER 99] P. Verlinde, P. Druyts, G. Cholet, and M. Acheroy. “Applying Bayes based
classifiers for decision fusion in a multi-modal identity verification system”.
In :Proceedings of International Symposium on Pattern Recognition “In
Memoriam Pierre Devijver”, Brussels, Belgium, February 1999.
- [VOL 00] Volker Blanz Thomas Vetter, A Morphable Model For The Synthesis Of
3D Faces, Max-Planck-Institut für biologische Kybernetik,Tubingen,
Germany.2000.
- [WAL09] : Walid Hizem, L'Institut National des Télécommunications et l'Université
Pierre et Marie Curie - Paris 6, Capteur Intelligent pour la Reconnaissance
de Visage, pour obtenir le diplôme de Docteur De L'institut National Des
Télécommunications .2009.
- [WIL 06] Wilson, P., Fernandez, J.: Facial feature detection using haar classifiers. *J.*
Comput. Sci. Coll. **21**(4), 127–133 (2006)
- [YAN 02] Yang, M.H., Kriegman, D., Ahuja, D.: Detecting faces in images: a survey.
IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), 34–58 (2002)

[YUI 92] Yuille, A.L., Hallinan, P.W., Cohen, D.S.: Feature extraction from faces using deformable templates. *Int. J. Comput. Vis.* **8**(2), 99–111 (1992).

[ZHA 98] Zhang, X., Xu, Y., Du, L.: Locating facial features with color information. In: *Proceedings of the Fourth International Conference on Signal Processing (ICSP)*, vol. 2, pp. 889–892 (1998)

Annexe A

A.1 Classification de machines à vecteur

A.1.1 Historique

L'origine des machines à vecteurs de support (SVM) remonte à 1975 lorsque **Vapnik** et **Chervonenkis** proposèrent des théories d'apprentissage statistique pour caractériser la capacité d'une machine d'apprentissage. A cette époque, ces principes n'ont pas trouvé place et il n'existait pas encore un modèle de classification solidement appréhendé pour être utilisable. Il a fallu attendre jusqu'à l'an 1982 pour que Vapnik propose un premier classificateur basé sur la minimisation du risque structurel baptisé SVM.

Ce modèle était toutefois linéaire et l'on ne connaissait pas encore le moyen d'induire des frontières de décision non linéaires. En 1992, **Boser et al** proposent d'introduire des noyaux non-linéaires pour étendre le SVM au cas non-linéaire. En 1995, Cortes et al proposent une version régularisée du SVM qui tolère les erreurs d'apprentissage tout en les pénalisant [Ned 04]

A.1.2. Principe de la technique SVM

Le modèle le plus simple de SVM est celui appelé linéaire de marge maximale. Il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables. Ce SVM cherche à séparer les deux classes de données par un hyperplan qui est équidistant des « frontières » de chaque classe.

Dans la plupart des cas réels, SVM consiste à projeter les données de l'espace d'entrée (appartenant à deux classes différentes) non linéairement séparables dans un espace de plus grande dimension appelé espace de caractéristiques (futur space) de façon à ce que les données deviennent linéairement séparables. Dans cet espace, on construit un hyperplan optimal séparant les classes tel que [Han 09]:

- Les vecteurs appartenant aux différentes classes se trouvent de différents côtés de l'hyperplan.
- La plus petite distance entre les vecteurs et l'hyperplan (la marge) soit maximale.

Le principe de la technique SVM est représenté dans la figure ci-dessous :

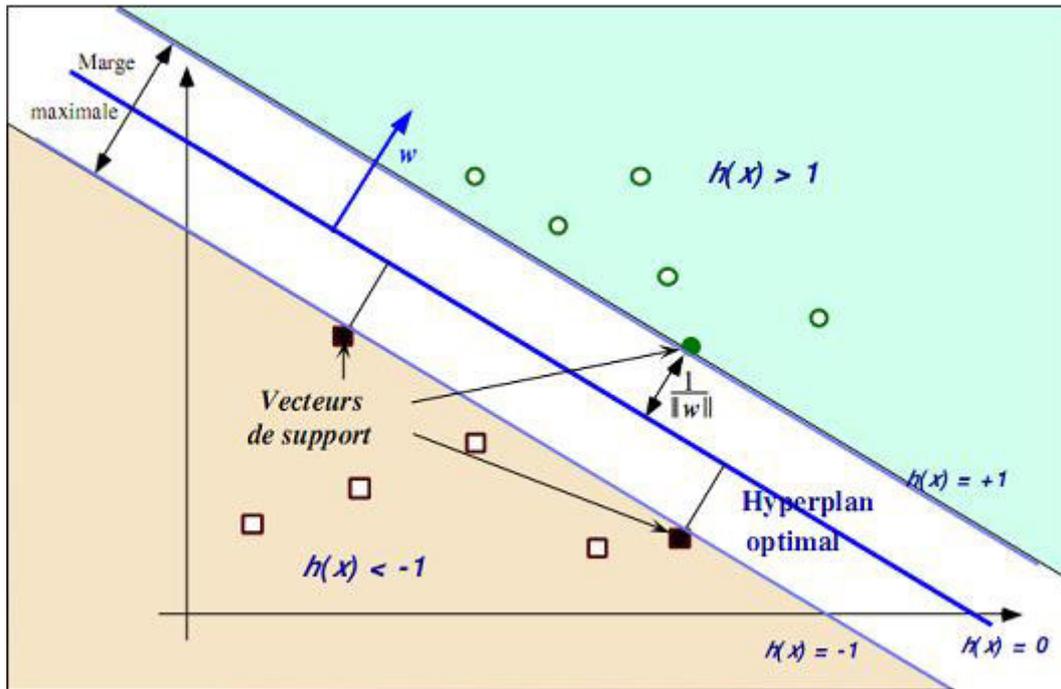


Figure A.1: Principe de la technique SVM (hyperplan optimal, vecteurs de supports, marge maximale).

Hyperplan optimal : est un Hyperplan qui classé correctement les données (lorsque c'est possible) et qui se trouve le plus loin possible de tous les exemples, on peut dire aussi que cet hyperplan maximise la marge .[Oli 01]

Vecteurs de support : ce sont Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan.

La marge : est la distance entre l'hyperplan et les exemples. La marge est calculée à partir du produit scalaire entre les vecteurs situés la frontière de chaque classe et le vecteur unitaire normal de l'hyperplan séparateur.

Annexe B

B.1 Réseaux de neurones artificiels

Le principe général des Réseaux de Neurones Artificiels (RNA) est à l'origine inspiré de certaines fonctions de base des neurones naturels du cerveau.

Un réseau de neurones artificiel est généralement organisé en plusieurs couches :

- une couche d'entrée,
- une couche de sortie,
- des couches intermédiaires appelées couches cachées.

La présence de couches cachées permet de discriminer des classes d'objets non linéairement séparables.

En général, un réseau de neurones est fondamentalement un classifieur, il réalise un travail de classification pendant la phase d'apprentissage, et de classement lors de la reconnaissance.

Mais on peut l'utiliser pour réaliser la fusion de données [Sit 10], [Rac 06] pour séparer deux populations données, à savoir les clients et les imposteurs dans notre cas.

Les grands avantages des réseaux de neurones résident dans leur capacité d'apprentissage automatique (approximation universelle (Cybenko, Hornik)), ce qui permet de résoudre des problèmes sans nécessiter l'écriture de règles complexes, tout en étant tolérant aux erreurs. Ils résident aussi dans leur capacité à prendre une décision à partir de critères non formalisables explicitement[Muh05].