

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
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 :

***Identification de reconnaissance
faciale avec des expressions***

**Présentée par :
OUAMANE Hanane
Soutenu le : 07 Juin 2012**

Devant le jury composé de :

Mlle : *Hendaoui Mounira*

MAB

Président

Mme: *BENATIA Mébarka*

MAA

Encadrice

Mlle : *Fediase Meriam*

MAB

Examinateur

Promotion: 2011/2012

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
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 :

*Identification de reconnaissance
faciale avec des expressions*

Présenté par :

OUAMANE Hanane

Avis favorable de l'encadreur:

BENATIA Mébarka

Avis favorable du président du Jury

Nom et prénom

Hendaoui Mounira

Signature

Cachet et signature



Résumé du mémoire de fin d'étude
En vue d'optention de diplôme:
Master
Option : Télécommunication

Thème :
Identification de reconnaissance
faciale avec des expressions

Proposée par : Mme BENATIA Mébarka
Dirigée par : Mme BENATIA Mébarka

RESUME

Notre objectif est l'étude des techniques d'identification de visages en présence d'expression dans un premier temps. Puis la mise en œuvre d'un système de reconnaissance en choisissant celui qui nous présente les meilleures performances au sens biométrique. Un état de l'art des approches récentes dans le domaine est établi. Ce qui nous a guidés dans notre choix des techniques utilisées dans la conception de notre système de reconnaissance de visages. Notre approche est basée sur un prétraitement qui nous assure la préparation de l'image pour la soumettre à l'analyse. Cette dernière est basée sur une étape d'extraction de paramètres qui sont les régions d'intérêts du visage qui sont fusionnées. Puis projetées sur un espace réduit réalisé par Le modèle discriminant Linéaire amélioré de Fisher l'EFM. Ensuite ces descripteurs sont soumis à un classifieur connu pour son efficacité c'est le SVM. Et finalement nous avons tenté la concaténation des meilleurs filtres pour améliorer le taux de réussite. Les résultats obtenus sont encourageants et prometteurs. Nos résultats montrent une précision de taux de reconnaissance automatique des cinq expressions faciales de 81.30% sur la base de données d'expression faciale CASIA 3D.

Chapitre I

Reconnaissance de visage

I.1 Introduction

Dans son environnement quotidien, un individu a besoin de s'identifier dans une multitude de contextes : pour entrer dans son immeuble ou accéder à son lieu de travail, pour retirer de l'argent à un distributeur ou payer en magasin, pour demander un service social... Autant de codes et de mots de passe à mémoriser et à protéger. Afin de développer les moyens de reconnaissance, la recherche connaît depuis quelques années un renouveau spectaculaire et manifeste un intérêt majeur aux données "biométriques", c'est-à-dire aux caractéristiques propres à chaque personne : sa voix, ses empreintes digitales, les traits de son visage, la forme de sa main, sa signature et même son ADN. Ainsi, les techniques basées sur la biométrie jouissent d'un engouement général favorisé par un phénomène de mode, principalement véhiculé par les films au cinéma et à la télévision. Cependant, plus récemment, l'augmentation de la fraude à l'identité a créé un besoin croissant de la technologie biométrique dans un certain nombre d'applications nécessitant un haut degré de sécurité : accès à des sites sensibles, surveillance d'aéroport.

Il existe traditionnellement deux manières d'identifier une personne :

1. Méthodes basées sur une connaissance (knowledge-based). Cette connaissance correspond par exemple à un mot de passe pour ouvrir une session ou un code SIM pour un téléphone portable.
2. Méthodes basées sur une possession (token-based). Il peut s'agir d'une pièce d'identité, d'un badge, d'une clé ...

Ces deux modes d'identification peuvent être utilisés d'une manière complémentaire afin d'obtenir une sécurité accrue comme pour la carte bleue. Cependant elles ont leurs faiblesses respectives. Le mot de passe peut être oublié ou deviné par une autre personne. On estime ainsi qu'une personne sur quatre écrit son mot de passe sur sa carte bleue pour ne pas l'oublier. D'autre part, la pièce d'identité peut être perdue ou volée. La biométrie exploite à l'encontre de ces deux méthodes les caractéristiques d'une personne qu'elles soient innées comme les empreintes digitales ou acquises comme la signature. Ces caractéristiques sont attachées à chaque individu et ne souffrent donc pas des faiblesses des méthodes basées sur une connaissance ou une possession. En effet les caractéristiques biométriques ne peuvent être oubliées ou perdues. De plus, elles sont très difficiles à deviner, à voler, ou à dupliquer [Cha09].

I.2 La reconnaissance biométrique

La reconnaissance biométrique fait référence à l'utilisation de différentes caractéristiques physiologiques (empreintes digitales, visage, rétine, iris, voix) et comportementales (manière de marcher, dynamique de la signature), appelées caractéristiques biométriques (voir Figure I.1), ou simplement biométrie, pour reconnaître automatiquement les individus. Toutes les caractéristiques biométriques d'une personne sont, en fait, une combinaison de caractéristiques physiologiques et comportementales. Par exemple, les empreintes digitales sont de nature physiologique, mais la façon dont un utilisateur présente son doigt au capteur dépend du comportement de la personne. Ainsi, l'entrée d'un système de reconnaissance biométrique est une combinaison de caractéristiques physiologiques et comportementales. De même, la voix est en partie déterminée par la structure biologique qui produit la parole chez un individu et aussi par la façon dont une personne parle. Souvent, une similitude peut être constatée entre les parents, les enfants, et les frères et sœurs dans leurs voix, leurs démarches, et même dans leurs signatures. Le même argument s'applique aux visages: les visages de jumeaux peuvent être très similaires à la naissance, mais au cours du développement, les visages changent en fonction du mode de vie et du comportement de la personne. De plus, les sources de variabilité entre deux vues d'un même visage sont multiples, et peuvent même engendrer des dissimilarités plus importantes que celles observés entre deux visages différents. Ceci se traduit par une baisse significative des performances de la reconnaissance automatique dans les conditions réelles pour le visage ainsi que pour la majorité des modalités biométriques. Ainsi la reconnaissance automatique des caractéristiques biométriques reste un problème complexe et non efficacement résolu [Cha09].

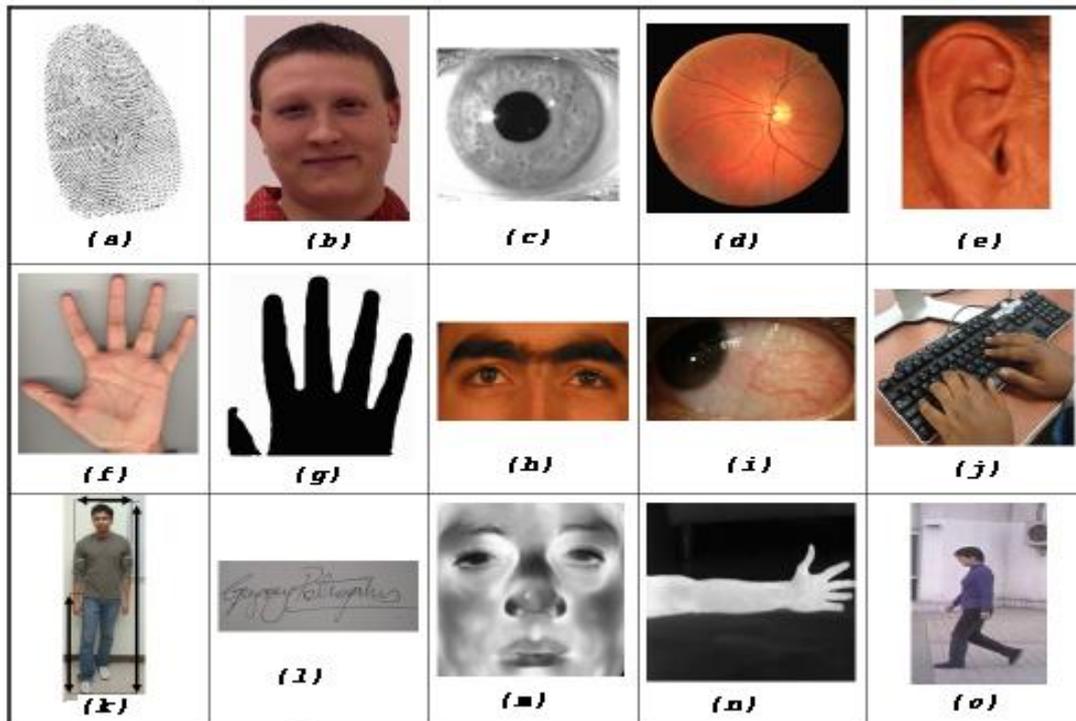


Fig I.1 : Les principales caractéristiques biométriques pouvant servir comme moyen d'identification : a) Empreinte digitale, b) Visage, c) Iris, d) Rétine, e) Forme de l'oreille, f) Paume, g) Géométrie de la Main, h) Paire oculaire, i) Conjonctivale Vasculaire, j) Dynamique de frappe, k) Anthropométrie, l) Signature, m) Thermo-gramme du Visage, n) Thermo-gramme de la main, o) Démarche.

I.2.1 Les différentes techniques biométriques

La comparaison entre les différentes biométries permet de choisir une technologie en fonction des contraintes liées à l'application. En effet, chaque caractéristique (ou modalité) biométrique a ses forces et ses faiblesses, et faire correspondre un système biométrique spécifique à une application dépend du mode opérationnel de l'application et des caractéristiques biométriques choisies. En a proposé une comparaison (avantages / inconvénients) des principales technologies biométriques en se basant sur la facilité ou l'ergonomie d'utilisation, la vulnérabilité aux attaques et aux contournements et la fiabilité relative à la précision et à l'efficacité de la reconnaissance (voir tableau I.1) [Cha09].

| Techniques | Avantages | Inconvénients |
|-----------------------------|---|--|
| Empreintes digitales | Coût, ergonomie moyenne, facilité de mise en place, taille du capteur | Fiabilité des appareils de mesure, acceptabilité moyenne, possibilité d'attaques (rémanence de l'empreinte,...) |
| Forme de la main | Très ergonomique, bonne acceptabilité | Système encombrant, coût, perturbation possible par des blessures et l'authentification des membres d'une même famille, permanence des données |
| Visage 2D | Coût, peu encombrant, bonne acceptabilité | Jumeaux, psychologie, déguisement, vulnérabilité aux attaques |
| Rétine | Fiabilité, pérennité | Coût, acceptabilité faible, installation difficile |
| Iris | Fiabilité | Acceptabilité très faible, contrainte d'éclairage |
| Voix | Facilité | Vulnérable aux attaques |
| Signature | Ergonomie | Dépendant de l'état émotionnel de la personne fiabilité |
| Frappe au clavier | Ergonomie | Dépendant de l'état physique de la personne |

Tableau I.1 : Avantages et inconvénients des technologies biométriques.

I.2.2 Le marché mondial de la biométrie

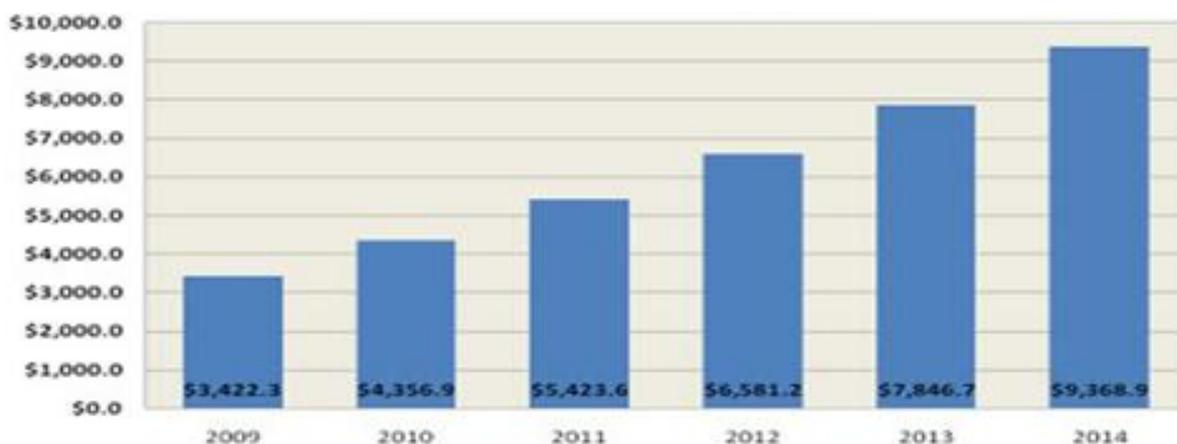


Fig I.2 : Revenus de la vente de technologies biométriques de 2009 à 2014 selon IBG.

IBG (International Biometric Group) édite régulièrement un rapport sur le marché de la biométrie. Cette étude est une analyse complète des chiffres d'affaires, des tendances de croissance et des développements industriels pour le marché de la biométrie actuel et futur. La figure I.2 montre les revenus du marché de la biométrie prévu entre l'année 2009 et 2014.

I.3 Le mode de reconnaissance

La reconnaissance peut être une vérification ou une identification [Lou10].

I.3.1 Le mode de vérification ou d'authentification

La vérification est une comparaison "un à un", dans laquelle le système valide l'identité d'une personne en comparant les données biométriques saisies avec le modèle biométrique de cette personne stocké dans la base de données du système. Dans un tel mode, le système doit alors répondre à la question suivante: *«Suis-je réellement la personne que je suis en train de proclamer?»*

I.3.2 Le mode d'identification

C'est une comparaison "un à N", dans lequel le système reconnaît un individu en l'appariant avec un des modèles de la base de données. Ce mode consiste à associer une identité à une personne. En d'autres termes, il répond à des questions de type: *« Qui suis-je ? »*.

Les schémas d'un système de vérification et d'un système d'identification sont illustrés dans la figure I.3; le processus d'enrôlement, qui est commun à ces deux tâches est également illustré. Le module d'enrôlement correspond à l'enregistrement biométrique des individus dans la base de données du système. Pendant la phase d'enrôlement, la caractéristique biométrique d'un individu est capturée par un lecteur biométrique pour produire une représentation numérique.

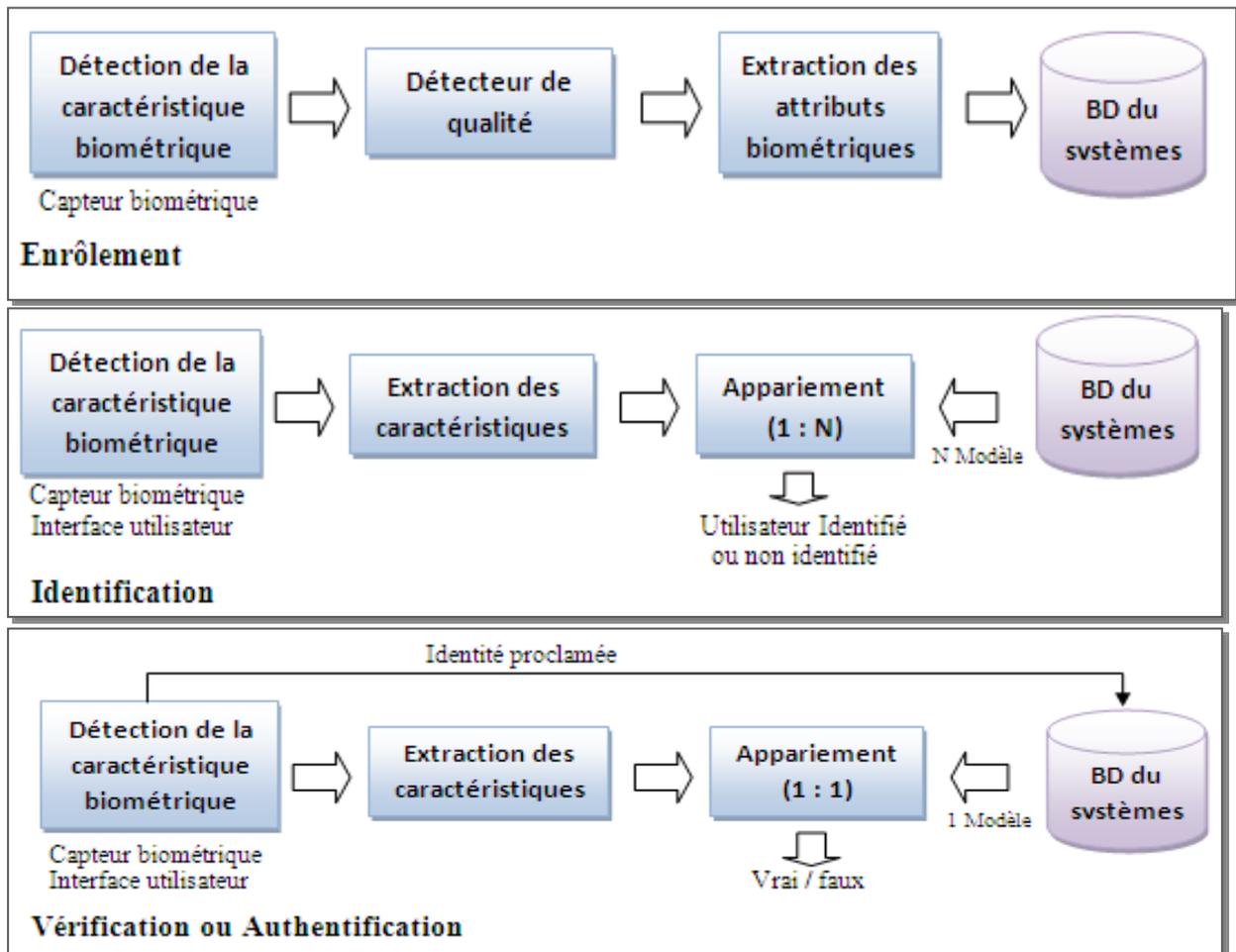


Fig I.3 : Diagrammes des processus d'enrôlement, de vérification et d'identification.

I.4 Pourquoi choisir le visage ?

La reconnaissance de visages est la technique la plus commune et populaire. Elle reste la plus acceptable puisqu'elle correspond à ce que les humains utilisent dans l'interaction visuelle; et par rapport aux autres méthodes, la reconnaissance du visage s'avère plus avantageuse, d'une part c'est une méthode non intrusive, c'est-à-dire elle n'exige pas la coopération du sujet (en observant les individus à distance), et d'une autre part les capteurs utilisés sont peu coûteux (une simple caméra) contrairement à l'empreinte digitale et l'iris où le sujet devra être très proche du capteur et devra coopérer pour l'acquisition de l'image sans oublier le coût de l'équipement nécessaire pour l'acquisition (équipement spécial coûteux). Malgré que certains disent que la reconnaissance de visages est une biométrie relativement peu sûre, sur le fait que le signal acquis est sujet à des variations beaucoup plus élevées que d'autres caractéristiques, comme la variation de l'éclairage, le changement de la position du visage, la présence ou l'absence de lunettes et autres; mais, au cours de ces dernières années

plusieurs techniques de traitements d'images sont apparues, telle que la détection du visage, la normalisation de l'éclairage, etc. Sans oublier le développement considérable des technologies des caméras numériques, ce qui néglige l'effet de ces problèmes [Adj07].

I.5 Reconnaissance du visage

La reconnaissance faciale est une tâche que les humains effectuent naturellement et sans effort dans leurs vies quotidiennes. La grande disponibilité d'ordinateurs puissants et peu onéreux ainsi que des systèmes informatiques embarqués ont suscité un énorme intérêt dans le traitement automatique des images et des vidéos numériques au sein de nombreuses applications, incluant l'identification biométrique, la surveillance, l'interaction homme-machine et la gestion de données multimédia. La reconnaissance faciale, en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de la recherche, ceci étant dû aux **avances rapides** dans des technologies telles que les appareils photo numériques, Internet et les dispositifs mobiles, le tout associé à des besoins en sécurité sans cesse en augmentation.

La reconnaissance faciale possède plusieurs avantages sur les autres technologies biométriques : elle est naturelle, non intrusive et facile à utiliser. Parmi les six attributs biométriques considérés, les caractéristiques faciales marquent un score de compatibilité le plus élevé dans un *système MRTD* ("Machine Readable Travel Documents"), ce score étant basé sur **plusieurs facteurs d'évaluation** tels que l'enrôlement, le renouvellement des données, les requis matériels et la perception des utilisateurs (Figure.I.4) [Mor09].

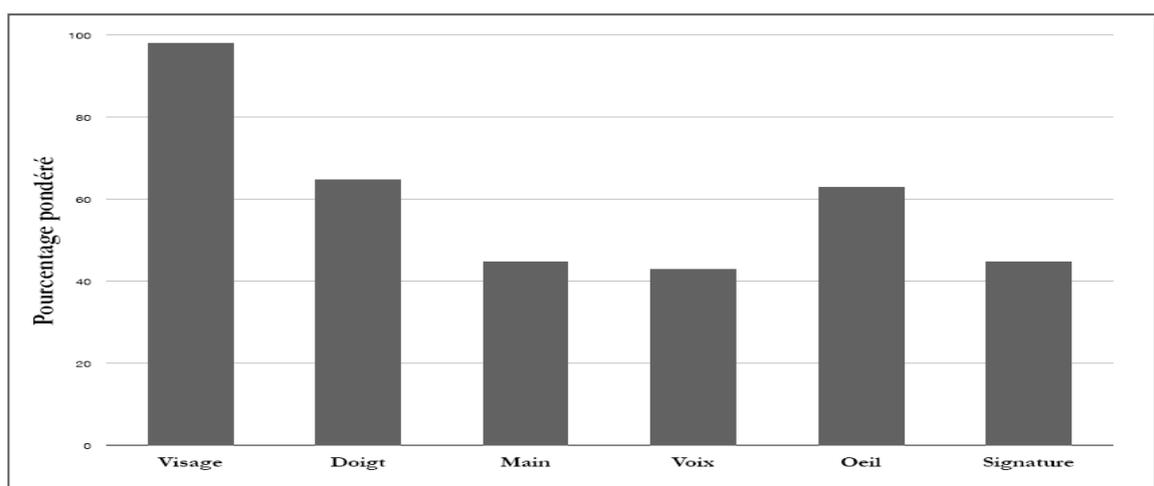


Fig I.4: Scores de compatibilité pour différentes technologies biométriques dans un système MRTD.

Idéalement, un système de reconnaissance faciale doit pouvoir identifier des visages présents dans une image ou une vidéo de manière automatique. Le système peut opérer dans les deux modes suivants : authentification ou identification ; on peut également noter qu'il existe un autre type de scénario de reconnaissance faciale mettant en jeu une vérification sur une liste de surveillance, où un individu est comparé à une liste restreinte de suspects. Le principe de fonctionnement de base d'un système de reconnaissance faciale (Figure. I.5) peut être résumé **en trois étapes**.

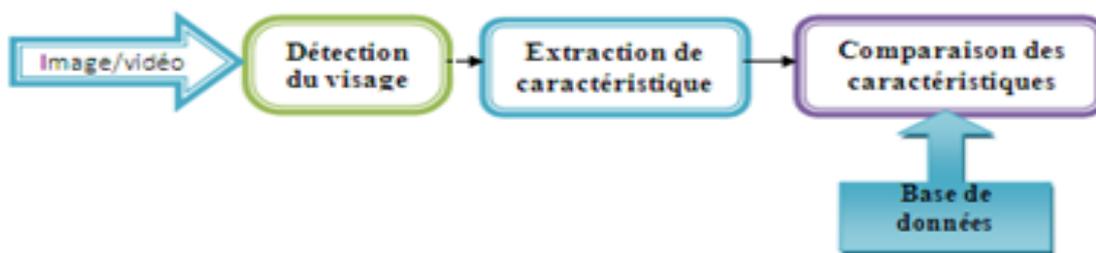


Fig I.5: Principe de fonctionnement de base d'un système de reconnaissance faciale.

I.5.1 La détection du visage

Cette étape consiste à l'extraire de l'image de visage, La détection de visage peut se faire par détection de la couleur de la peau, la forme de la tête ou par des méthodes détectant les différentes caractéristiques du visage. Cette étape est autant plus délicate quand l'image acquise contient plusieurs objets de visage ou un fond non uniforme qui crée une texture perturbant la bonne segmentation du visage. Cette étape est dépendante de la qualité des images acquises. Les méthodes de la détection de visages peuvent être classifiées en quatre catégories :

- **Méthodes basées sur les connaissances** : Ces méthodes basées sur des règles tentent de modéliser la connaissance de ce qui caractérise un visage. Par exemple un visage apparaît souvent dans une image avec deux yeux symétriques entre un nez, et une bouche. Classiquement, ces règles représentent des relations en caractéristiques faciales.
- **Méthodes basées sur les caractéristiques invariantes** : Ces méthodes se basent sur des caractéristiques structurelles (Traits faciaux, Texture, Couleur de la peau) qui existent même quand la pose, le point de vue, ou les conditions d'illumination varient.
- **Méthodes basées sur la mise en correspondance modèle** : Plusieurs modèles standards de visages sont prédéfinis manuellement ou paramétrés per des fonctiins.

L'un de ces modèles est comparé à une image en entrée. La corrélation entre une image présentée et la base des modèles est évaluée pour détecter la présence de visage.

- **Méthodes basées sur l'apprentissage** : les modèles sont ici appris à partir d'un ensemble d'images d'apprentissage qui doivent permettre de caractériser la variabilité de l'apparence d'un visage. Ces méthodes se basent sur des techniques tels que l'analyse statistique et l'apprentissage automatique pour trouver les caractéristiques appropriées des images de visage et de non-visage.

I.5.2 L'extraction des caractéristiques

Après la détection de la zone de visage, on doit extraire les informations utiles qui reviennent à établir un modèle du visage (vecteur de caractéristiques). Ces informations nécessaires pour que le visage d'une personne ne ressemble pas à celui d'une autre personne et en même temps qu'il ressemble à lui-même dans d'autres conditions d'acquisition. L'efficacité de cette étape a une influence directe sur la performance du système de reconnaissance de visage. On distingue les trois catégories suivantes :

- **Les méthodes globales** : Ces méthodes utilisent la région entière du visage comme entrée du système de reconnaissance. L'une des méthodes la plus largement utilisée pour la représentation du visage dans son ensemble est ACP.
- **Les méthodes locales** : Précisément, ces méthodes extraient les caractéristiques locales de visage comme les yeux, le nez et la bouche, puis utilisent leur géométrie et/ou l'apparence comme donnée d'entrée du classificateur.

I.5.3 La comparaison des caractéristiques (classification)

Selon les caractéristiques extraites précédemment, il faut introduire une mesure de similarité afin d'estimer la différence entre deux visages. Plusieurs méthodes sont utilisées pour la comparaison de visages. On trouve dans la littérature: calcul de distance, calcul de similarité. D'autres méthodes se basent sur la classification des caractéristiques par un seul classificateur (SVM, classificateur bayésien, etc.) ou par plusieurs (Adaboost) [Suo08].

I.6 Les méthodes de reconnaissance faciales

Les méthodes de reconnaissance faciales peuvent être séparées en deux grandes familles, les méthodes globales (ou holistiques) et les méthodes locales, basées sur des modèles. Le choix a été fait de se concentrer sur ces deux types d'approches fondamentales et

de n'aborder ni les réseaux neuronaux (NN) (plus adaptés à la détection des visages), ni les modèles cachés de Markov (HMM) (plus utilisés en reconnaissance de la parole) car ces deux dernières techniques rencontrent des problèmes lorsque le nombre d'individus augmente (les calculs deviennent très importants) ; de plus elles ne conviennent pas pour les systèmes de reconnaissance basés sur une seule "image modèle" car de nombreuses images par personne sont nécessaires pour entraîner les systèmes afin de configurer leurs paramètres de façon "optimale" [Mor09].

I.6.1 Méthodes globales

Les méthodes globales sont basées sur des techniques d'analyse statistique bien connues. Il n'est pas nécessaire de repérer certains points caractéristiques du visage (comme les centres des yeux, les narines, le centre de la bouche, etc.) à part pour normaliser les images. Dans ces méthodes, les images de visage (qui peuvent être vues comme des matrices de valeurs de pixels) sont traitées de manière globale et sont généralement transformées en vecteurs, plus faciles à manipuler.

L'avantage principal des méthodes globales est qu'elles sont relativement rapides à mettre en œuvre et que les calculs de base sont d'une complexité moyenne. En revanche, elles sont très sensibles aux variations d'éclairage, de pose et d'expression faciale.

Ceci se comprend aisément puisque la moindre variation des conditions de l'environnement ambiant entraîne des changements inéluctables dans les valeurs des pixels qui sont traités directement.

Ces méthodes utilisent principalement une analyse de sous-espaces de visages. Cette expression repose sur un fait relativement simple : une classe de "formes" qui nous intéresse (Dans notre cas, les visages) réside dans un sous-espace de l'espace de l'image d'entrée. Par exemple, considérons une petite image de 64×64 , en niveaux de gris codés sur 8 bits (donc de 0 à 255). Cette dernière possède 4096 pixels qui peuvent exprimer un grand nombre de classes de formes, comme des arbres, des maisons ou encore des visages.

Cependant, parmi les $256^{4096} > 10^{9864}$ "configurations" possibles, seulement une petite quantité peut correspondre à des visages. Ainsi, la représentation de l'image originale est très redondante et la dimensionnalité de cette représentation pourrait être grandement réduite si l'on se concentre uniquement sur les formes qui nous intéressent. L'utilisation de techniques de modélisation de sous-espace a fait avancer la technologie de reconnaissance faciale de manière significative.

Nous pouvons distinguer deux types de techniques parmi les méthodes globales : Les techniques linéaires et les techniques non linéaires.

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

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

Cependant, le PCA classique nécessite que les images de visage soient mises sous formes de vecteurs, ce qui a pour effet de détruire la structure géométrique de l'image. Pour ne pas perdre les informations de voisinage lors du passage de l'image en vecteur, une méthode PCA bidimensionnelle (2D PCA). Cette méthode prend en entrée des images et non plus des vecteurs.

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

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

L'algorithme ICA, quant à lui, est une généralisation de l'algorithme PCA avec lequel il coïncide dans le cas de données gaussiennes. L'algorithme ICA est basé sur le concept intuitif de contraste et permet d'éliminer la redondance statistique des données de départ. Ce dernier a été rendu célèbre notamment avec l'expérience de la "cocktail party" [Bro01] [Sag01] [Hay05] qui met en avant la résolution d'un problème de séparation (dé corrélation) de sources audio. Bartlett et al. ont fourni deux architectures différentes pour l'ICA : une

première architecture (ICA I) qui construit une base d'images statistiquement indépendantes et une deuxième architecture (ICA II) qui fournit une représentation en code factoriel des données.

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

Afin de pouvoir traiter ce problème de **non-linéarité** en reconnaissance faciale, de telles méthodes linéaires ont été étendues à des techniques non linéaires basées sur la notion mathématique de noyau ("kernel") comme le Kernel PCA et le Kernel LDA [Mik99]. Ici, une projection non linéaire (réduction de dimension) de l'espace de l'image sur l'espace de caractéristiques ("feature space") est effectuée ; les variétés présentes dans l'espace de caractéristiques résultant deviennent simple, de même que les subtilités des variétés qui sont préservées. Bien que les méthodes basées sur le noyau peuvent atteindre une bonne performance sur les données d'entraînement, il ne peut pas en être de même pour de nouvelles données en raison de leur plus grande flexibilité ; contrairement aux méthodes linéaires.

I.6.2 Méthodes locales

Les méthodes locales, basées sur des modèles, utilisent des connaissances a priori que l'on possède sur la morphologie du visage et s'appuient en général sur des points caractéristiques de celui-ci. Kanade présenta un des premiers algorithmes de ce type [Kan73] en détectant certains points ou traits caractéristiques d'un visage puis en les comparant avec des paramètres extraits d'autres visages. Ces méthodes constituent une autre approche pour prendre en compte la non-linéarité en construisant un espace de caractéristiques local et en utilisant des filtres d'images appropriés, de manière à ce que les distributions des visages soient moins affectées par divers changements.

Les approches Bayésiennes, les machines à vecteurs de support (SVM) [Guo00], la méthode des modèles actifs d'apparence (AAM) [Coo01] ou encore la méthode "local binary pattern" (LBP) [Aho04] ont été utilisées dans ce but.

Toutes ces méthodes ont l'avantage de pouvoir modéliser plus facilement les variations de pose, d'éclairage et d'expression par rapport aux méthodes globales. Toutefois, elles sont plus lourdes à utiliser puisqu'il faut souvent placer manuellement un assez grand nombre de points sur le visage alors que les méthodes globales ne nécessitent de connaître que la position des

yeux afin de normaliser les images, ce qui peut être fait automatiquement et de manière assez fiable par un algorithme de détection [Arc06].

I.6.3 Méthodes hybrides

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

L'analyse de caractéristiques locales (LFA) et les caractéristiques extraites par ondelettes de Gabor, sont des algorithmes hybrides typiques.

Plus récemment, l'algorithme LogGabor PCA (LG-PCA) [Per05] effectue une convolution avec des ondelettes de Gabor orientées autour de certains points caractéristiques du visage afin de créer des vecteurs contenant la localisation et la valeur d'amplitudes énergétiques locales ; ces vecteurs sont ensuite envoyés dans un algorithme PCA afin de réduire la dimension des données. La figure. I.6 fournit une classification des algorithmes principaux de reconnaissance faciale.

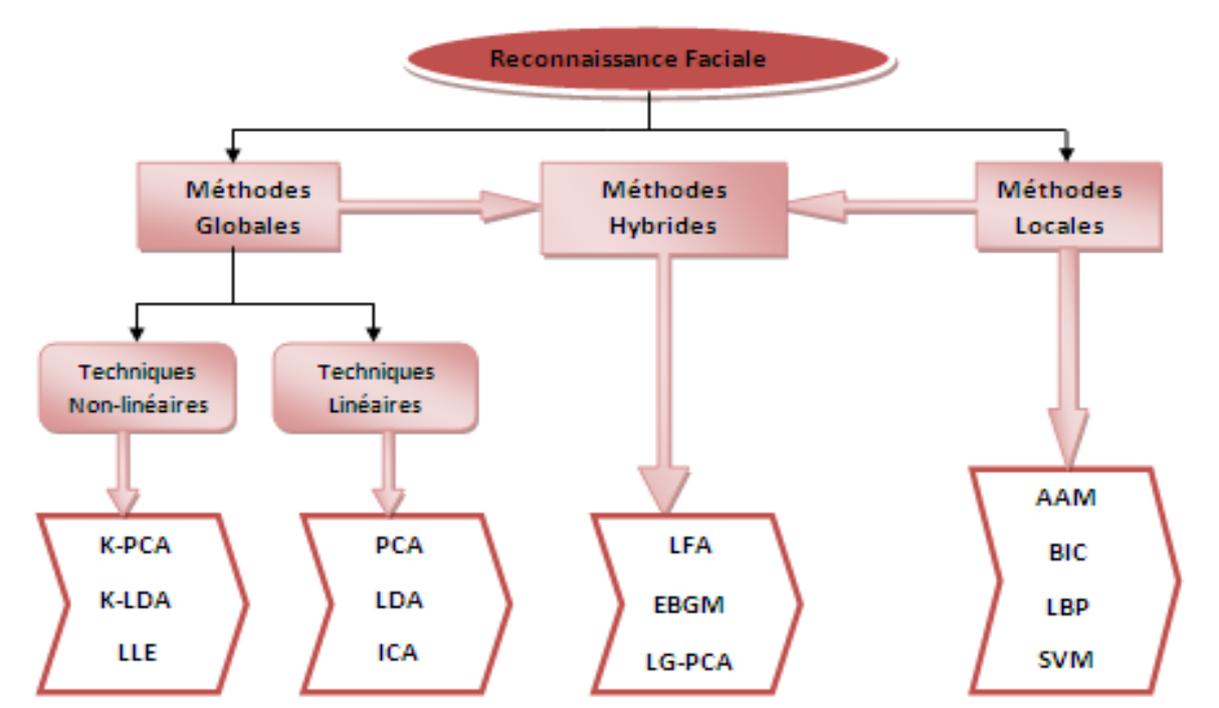


Fig I.6 : Classification des algorithmes principaux utilisés en reconnaissance faciale.

I.7 Principales difficultés de la reconnaissance de visage

Pour le cerveau humain, le processus de la reconnaissance de visages est une tâche visuelle de haut niveau. Bien que les êtres humains puissent détecter et identifier des visages dans une scène sans beaucoup de peine, construire un système automatique qui accomplit de telles tâches représente un sérieux défi. Ce défi est d'autant plus grand lorsque les conditions d'acquisition des images sont très variables. Il existe deux types de variations associées aux images de visages : inter et intra sujet. La variation inter sujet est limitée à cause de la ressemblance physique entre les individus. Par contre la variation intra sujet est plus vaste. Elle peut être attribuée à plusieurs facteurs que nous analysons ci-dessous [Sou08].

I.7.1 Changement d'illumination

L'apparence d'un visage dans une image varie énormément en fonction de l'illumination de la scène lors de la prise de vue (voir figure I.7). Les variations d'éclairage rendent la tâche de reconnaissance de visage très difficile. En effet, le changement d'apparence d'un visage dû à l'illumination, se révèle parfois plus critique que la différence physique entre les individus, et peut entraîner une mauvaise classification des images d'entrée. Ceci a été expérimentalement observé dans Adini et al [Adi97] où les auteurs ont utilisé une base de données de 25 individus. L'identification de visage dans un environnement non contrôlé reste donc un domaine de recherche ouvert. Les évaluations FRVT ont révélé que le problème de variation d'illumination constitue un défi majeur pour la reconnaissance faciale.



Fig I.7 : Exemple de variation d'éclairage.

I.7.2 Variation de pose

Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images. Cette difficulté a été démontrée par des tests d'évaluation élaborés sur les bases FERET et FRVT [Phi03]. La variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Quand le visage est de profil dans le plan image (orientation $< 30^\circ$), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux). Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible (voir figure I.8).



Fig I.8 : Exemples de variation de poses.

I.7.3 Expressions faciales

Un autre facteur qui affecte l'apparence du visage est l'expression faciale (voir figure 1.10). La déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. L'information faciale se situant dans la partie supérieure du visage reste quasi invariable. Elle est généralement suffisante pour effectuer une identification. Toutefois, étant donné que l'expression faciale modifie l'aspect du visage, elle entraîne forcément une diminution du taux de reconnaissance. L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui reste non résolu. L'information temporelle fournit une connaissance additionnelle significative qui peut être utilisée pour résoudre ce problème [Gro01].



Fig I.9 : Exemples de variation d'expressions.

I.7.4 Présence ou absence des composants structurels

La présence des composants structurels telle que la barbe, la moustache, ou bien les lunettes peut modifier énormément les caractéristiques faciales telles que la forme, la couleur, ou la taille du visage. De plus, ces composants peuvent cacher les caractéristiques faciales de base causant ainsi une défaillance du système de reconnaissance. Par exemple, des lunettes opaques ne permettent pas de bien distinguer la forme et la couleur des yeux, et une moustache ou une barbe modifie la forme du visage.

I.7.5 Occultations partielles

Le visage peut être partiellement masqué par des objets dans la scène, ou par le port d'accessoire tels que lunettes, écharpe... Dans le contexte de la biométrie, les systèmes proposés doivent être non intrusifs c'est-à-dire qu'on ne doit pas compter sur une coopération active du sujet. Par conséquent, il est important de savoir reconnaître des visages

partiellement occultés. Gross et al [Gre01] ont étudié l'impact du port de lunettes de soleil, et du cache-nez occultant la partie inférieure du visage sur la reconnaissance faciale. Ils ont utilisé la base de données AR [Mar98]. Leurs résultats expérimentaux semblent indiquer que, dans ces conditions, les performances des algorithmes de reconnaissance restent faibles.

I.8 Performances d'un système de reconnaissances de visage

Les performances d'un système de reconnaissance dépendent de plusieurs facteurs qui interviennent à plusieurs niveaux et qui peuvent limiter le degré de précision. Cependant, il serait judicieux de s'intéresser à ces facteurs avant de mesurer la performance d'un système de reconnaissance. Nous citons ici les principaux facteurs:

- L'environnement au moment de l'acquisition.
- Les différentes positions des capteurs.
- La qualité des capteurs.
- La mauvaise interaction entre l'utilisateur et les capteurs.

I.8.1 Performance d'un système d'identification

Pour évaluer les performances d'un système d'identification, on calcule le taux de reconnaissance du système. 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 : [Lou10]

$$TIC = \frac{\text{nombre de tests correspondant à une identification correcte}}{\text{le nombre total de tests}} \quad (\text{équ 1.1})$$

I.8.2 Performance d'un système de vérification

Un système de vérification ne peut donner, lors d'une comparaison entre deux échantillons qu'un résultat sous forme de probabilité de coïncidence ou score (S), ce score est comparé à un seuil de décision (T) que l'on détermine grâce à de nombreux tests; Si S est supérieur à T le sujet en question est accepté et dans le cas contraire le sujet est considéré comme imposteur. Dans ce cas, deux types d'erreurs peuvent être commis:

- **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.

$$FRR = \frac{\text{nombre de FR}}{\text{le nombre de clients}} \quad (\text{équ 1.2})$$

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

$$FAR = \frac{\text{nombre de FA}}{\text{le nombre d'imposteurs}} \quad (\text{équ 1.3})$$

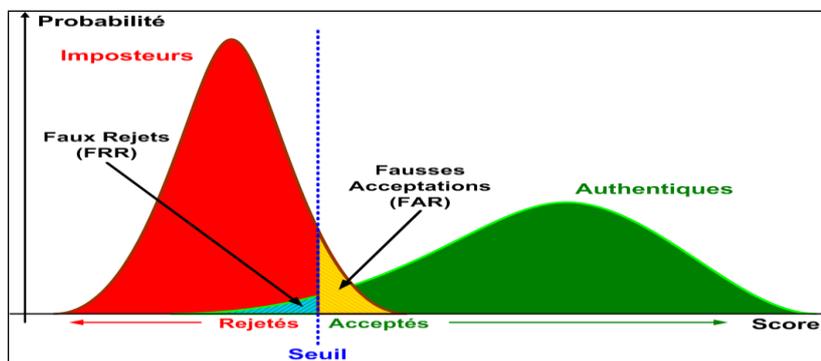


Fig I.10 : La distribution du taux de vraisemblances.

On peut représenter graphiquement les performances d'un système de reconnaissance de visage à l'aide de la courbe **ROC (Receiver Operating Characteristic)**. Les **FRR** sont donnés en fonction des **FAR** pour les différentes valeurs du seuil Θ



Fig 1.11: La courbe ROC.

Il existe d'autres critères statistiques pour mesurer les performances d'un système biométrique :

- **L'EER (Equal Error Rate ou EER):** Il correspond à l'intersection de la courbe ROC avec la première bissectrice : **FAR = FRR**

- **Le HTER (Half Total Error Rate):** Il représente la moyenne de **FAR** et **FRR** donc :

$$HTER = \frac{FAR + FRR}{2} \quad (\text{équ 1.4})$$

- **Le TER (Total Error Rate):** représente le taux d'erreur total.

$$TER = \frac{\text{nombre de clients rejetés} + \text{nombre d'imposteurs acceptés}}{\text{nombre total d'accès}} \quad (\text{équ 1.5})$$

I.9 Nouvelle technologie améliorant la reconnaissance faciale

I.9.1 L'utilisation de la technologie infrarouge

Les systèmes de reconnaissance de visage infrarouges utilisent des ondes infrarouges pour mesurer les radiations thermiques émises dans la gamme du spectre infrarouge. Ces systèmes ne sont pas sensibles à la couleur de la peau, bronzage, usage de produits de beauté, les lentilles de l'œil coloré, la chirurgie plastique et les photos peuvent être prises dans l'obscurité. Néanmoins, le coût extrêmement élevé de cette technologie rend son utilisation prohibitive pour les applications [Lou10].

I.9.2 L'utilisation de la technologie 3D

La quasi-totalité des systèmes de reconnaissance de visages travaillent à partir d'images fixes. Pour améliorer cette technique, les chercheurs s'orientent vers l'utilisation du modèle 3D du visage, à l'aide de plusieurs caméras, scanners ou caméras spécialisées. L'objectif est de conserver les avantages de la reconnaissance faciale avec des scores de fiabilité proches de ceux obtenus en empreintes digitales par exemple [Pet 09].

I.9.3 Avantage de la technologie 3D

a) Insensibilité à la variation lumineuse Lorsque la personne se présente de biais, un logiciel ferait « tourner » numériquement le modèle 3D pour le faire correspondre à l'image 2D prise en conditions réelles. De même, ce logiciel modifierait les conditions lumineuses de l'image 3D pour les rendre plus proches de l'éclairage existant.

b) Invariance de l'angle La diversité des paramètres 3D fait que le dispositif d'enregistrement en temps réel est en mesure de reconnaître des personnes se trouvant dans un champ correspondant à une rotation de 30° de part et d'autre de son axe.

c) Unicité des images 3D La quantité de points de mesure et de données biométriques saisies est telle qu'elle permet de distinguer des jumeaux monozygotes.

I.10 Conclusion

Dans ce chapitre, nous avons présenté les technologies utilisées dans les systèmes biométriques pour l'identification de personnes. Cette étude nous a permis de constater que la reconnaissance de visage suscite de plus en plus l'intérêt de la communauté scientifique, car elle présente plusieurs challenges et verrous technologiques. Nous avons mis en évidence les différentes difficultés inhérentes à la reconnaissance automatique de visages, notamment l'invariance à l'illumination. Enfin, nous avons aussi donné un aperçu sur les techniques de mesure de performances. Les techniques utilisées aux différentes étapes de la reconnaissance des expressions faciale sont détaillées dans le chapitre suivant.

Chapitre II

Etat de l'art de reconnaissance d'expression faciale

II.1 Introduction

La reconnaissance des expressions faciales a débuté depuis une vingtaine d'années. Les expressions faciales jouent un rôle important dans la conversation humaine. Donc, l'interface homme machine du futur utilisera non seulement la souris, le clavier, mais aussi de nouvelles sources d'informations telles que les six expressions faciales (**colère**, dégoût, **joie**, **peur**, **surprise**, **tristesse**), la posture humaine ou parole. Dans ce chapitre, nous présenterons un état de l'art sur les techniques de reconnaissance d'expressions faciales.

II.2 Qu'est-ce qu'une expression faciale?

L'expression faciale est un ensemble des signes du visage qui traduisent un sentiment de changement dans le visage, perceptible visuellement, dû à l'activation (volontaire ou non) de l'un ou de plusieurs des 44 muscles composant le visage, il y a 250000 expressions possibles.



Fig II.1 : Mouvements faciaux globaux.

Pourquoi les expressions faciales existent-elles?

Trois points de vue différents :

- Les émotions sont au centre de l'explication des expressions faciales
Sélection naturelle \Rightarrow émotions de base \Rightarrow expressions faciales universelles (colère, dégoût, joie, peur, surprise, tristesse), les étiquettes pourraient être remises en question car elles dépendent du langage et de la culture.
- Les expressions faciales sont des signaux de conversation qui dépendent des intentions de l'émetteur, du comportement du récepteur et du contexte de l'interaction, non des émotions.
Hausser les sourcils \leftrightarrow incrédulité, froncer les sourcils \leftrightarrow incompréhension, hausser le menton \leftrightarrow ignorance
- Les expressions faciales sont des activateurs et des régulateurs d'émotion dû à la plasticité

du cerveau, changent la température de l'hypothalamus pour faciliter ou inhiber la production de neurotransmetteurs reliés aux émotions.

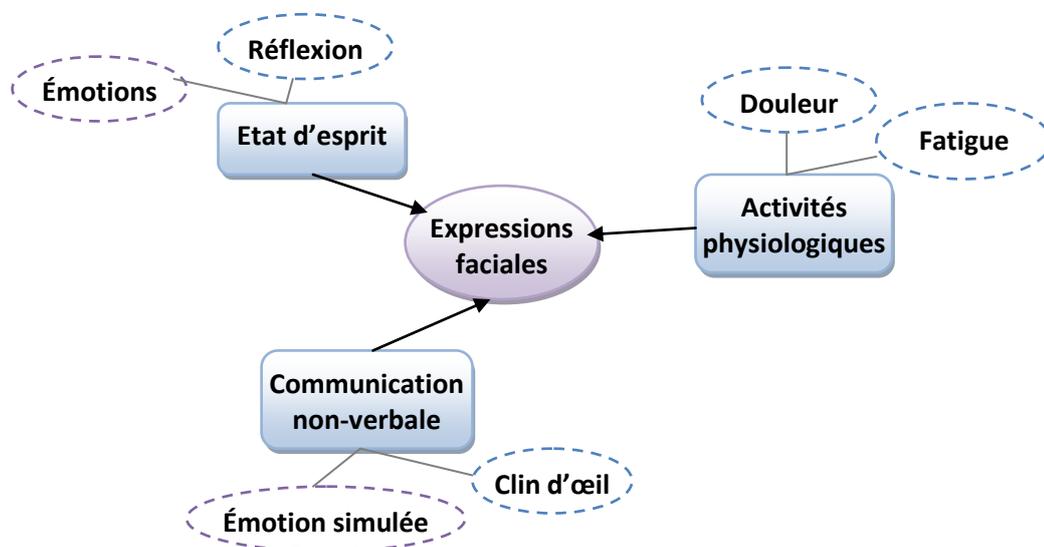


Fig II.2: Relation entre expression faciale et émotion.

Une émotion peut causer, une expression faciale ou non et une expression faciale peut être causée par une émotion ou non.

II.3 Le développement émotionnel

II.3.1 Comment étudie-t-on les expressions émotionnelles chez l'humain ?

a. Méthode subjective

Les études qui utilisent la "*méthode par les jugements*" sont les plus nombreuses. Elle consiste à chercher l'accord entre les observateurs sur l'état émotionnel exprimé par un visage. La face est alors considérée comme un stimulus engendrant un jugement chez des observateurs.

On peut se baser sur des photographies prises à l'improviste où les expressions faciales sont associées à des situations déterminées. Ou bien, on peut montrer aux juges des visages de sujets réagissant à des situations qu'on a contrôlées. Dans ces deux situations l'émotion exprimée est alors ressentie.

On peut à partir de photographies de sujets qui ont posé, tenter de traduire sur leurs visages une émotion, parmi celles proposées. Ici les résultats obtenus sont les plus satisfaisants. En effet lorsqu'une pose est faite dans un cadre artificiel, le sujet n'est ni soumis à une situation affective complexe, ni aux règles des manifestations communicatives, ainsi l'émotion mimée sera plus pure.

Quand on procède par jugements, l'observateur doit soit répondre librement soit répondre en faisant un choix parmi une liste de catégories d'émotion. L'approche par le jugement ne nous permet pas de savoir quelle modification de la face permet à l'observateur de déterminer l'émotion exprimée par le sujet [LAP07].

b. Méthode objective

L'"étude par les constituants" consiste à relever, pour les observateurs, les composantes de la face qui déterminent l'émotion exprimée par le sujet. Ici c'est la quantité et la qualité des paramètres faciaux qui sont déterminées.

Darwin utilisa cette méthode en adressant aux observateurs seize questions sur les expressions faciales : par exemple il demanda si l'étonnement s'exprime par l'ouverture de la bouche, des yeux et par l'élévation des sourcils.

Ekman et Friesen ont recherché les caractéristiques des groupes musculaires impliqués dans les mimiques faciales. Ils ont proposé, pour la joie, la tristesse, la peur, la surprise, la colère et le dégoût, une liste d'expressions faciales. Ces dernières sont décrites à partir d'un système de codification de changements faciaux, le *Facial Action Coding System* (FACS). Ce système comprend 44 composantes faciales de base appelées unités d'action (UA). Chaque unité correspond à une modification de l'expression du visage sous l'action d'un muscle ou d'un groupe musculaire. Les expressions émotionnelles comprennent une ou plusieurs UA variant en intensité (In Gosselin, 1995). Une des difficultés pour l'étude des expressions émotionnelles réside dans le fait que tous les états émotionnels ne donnent pas nécessairement lieu à des modifications d'expressions faciales. Ainsi, si on demande au sujet d'imaginer une situation, il peut ressentir l'émotion sans pour autant l'exprimer au niveau de la face.

c. Neuroimagerie fonctionnelle

L'objectif de cette méthode est de comparer les changements transitoires d'états émotionnels aux variations d'activité du système nerveux. Pour étudier les changements fonctionnels cérébraux associés à l'état émotionnel, on recourt à des paradigmes "d'activation émotionnelle" : ainsi on distingue l'induction externe et interne d'état émotionnel, respectivement le sujet exposé à un stimulus visuel ou auditif et le sujet se remémorant un événement personnel chargé affectivement. On se demande alors s'il existe une spécificité régionale cérébrale selon que le sujet regarde passivement un stimulus émotionnel ou qu'il se concentre sur l'éprouvé émotionnel subjectif d'un événement. Il semblerait que quelque soit l'émotion induite et indépendamment de la méthode d'induction, une même structure cérébrale serait activée, elle se situerait dans le lobe frontal.

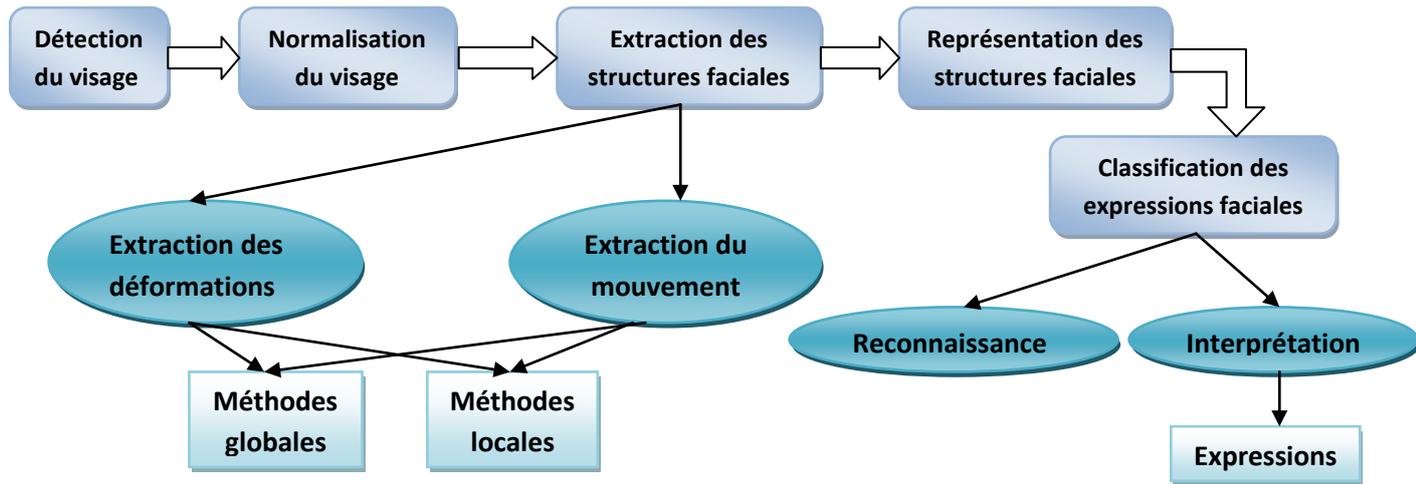


Fig II.3: Méthodes d'analyse automatique d'un système de reconnaissance d'expressions faciales.

II.4 Système de reconnaissance d'expression faciale existant

Ce système consiste à évaluer la possibilité de reconnaissance des six émotions universelles (citer les six classes considérées ici) en analysant les déformations des caractéristiques faciales telles que les yeux, les sourcils, et la bouche. On considère que ces caractéristiques contiennent suffisamment les informations pour la reconnaissance.

II.4.1 Méthode basée sur la segmentation

Ce système se présente comme l'indique la figure II.4 sous forme de quatre phases essentielles : la segmentation, extraction de données, analyse et classification.

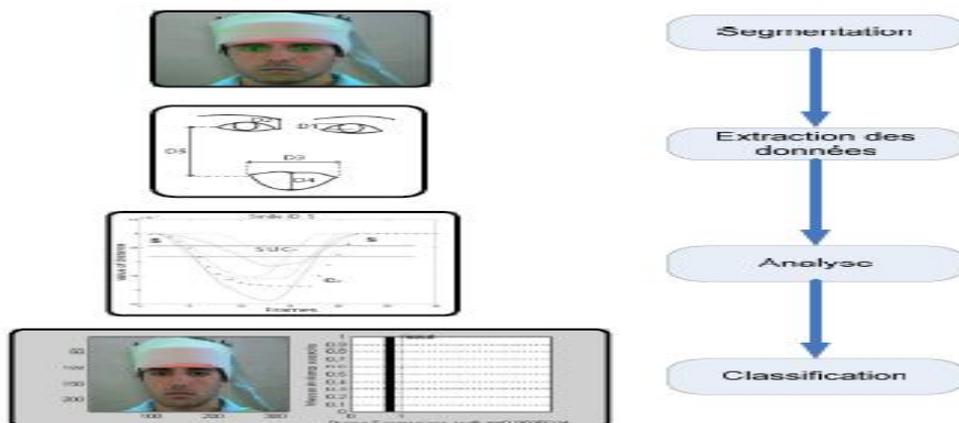


Fig II.4 : Les étapes dans le processus de reconnaissance des expressions faciales.

Dans l'étape de segmentation, le système va localiser les régions des yeux, des sourcils, et de la bouche. L'algorithme est présenté dans la thèse de Hammal [Ham06]. Dans l'étape d'extraction de données, les squelettes sont extraits. En fait, ce sont les contours des

yeux, de la bouche, des sourcils. A partir des squelettes, on va choisir cinq distances pour caractériser les déformations.

II.4.1.1 Segmentation des traits du visage

Dans cette partie, on décrit comment on peut extraire le squelette d'émotion à partir d'une image. Une hypothèse indispensable est que l'utilisateur doit faire face à la caméra. Après l'utilisation de l'algorithme de segmentation de contours de l'œil, des sourcils [Cap04] et de la bouche, les contours de caractéristiques faciales sont extraits automatiquement.

II.4.1.2 Extraction de données caractéristiques

Les caractéristiques considérées contiennent les contours des yeux, des sourcils et des lèvres. Finalement, on obtient les squelettes d'expressions.

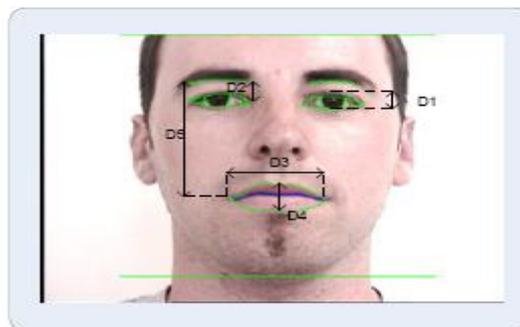


Fig II.5 : Extraction des contours et définition des 5 distances.

À partir du squelette d'expression, on peut déterminer les déformations de caractéristiques faciales. Cinq distances sont définies pour chaque squelette.

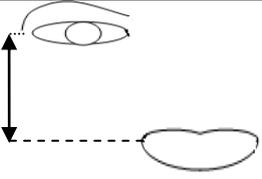
| | | |
|----|---|---|
| D1 |  | Distance entre la paupière supérieure et inférieure |
| D2 |  | Distance entre le coin intérieur de l'œil et le celui du sourcil |
| D3 |  | Largeur de bouche |
| D4 |  | Ouverture de bouche en hauteur |
| D5 |  | Distance entre le coin de la bouche et celui du coin extérieur de l'œil |

Tableau II.1 : Tableau de définition de distances.

II.4.1.3 Classification par le Modèle de Croyance Transférable

À partir de ces distances caractéristiques, on va utiliser le modèle de croyance transférable pour la reconnaissance d'expressions faciales.

États symboliques : On associe l'un des trois états symboliques suivants à chacune des valeurs de distances

- état C^+ pour lequel la distance D_i est plus grande que celle pour l'expression neutre,
- état S pour lequel la distance D_i est du même ordre de grandeur que celle pour l'expression neutre
- état C^- pour lequel la distance D_i est plus petite que celle pour l'expression neutre.

Après cela, chaque expression faciale est caractérisée par une combinaison d'états symboliques. Une combinaison d'états est définie pour chaque expression [Hon06].

| | D1 | D2 | D3 | D4 | D5 |
|----|--------|--------|--------|--------|--------|
| E1 | C^- | SC^- | C^+ | C^+ | |
| E2 | C^+ | C^+ | C^- | C^+ | C^+ |
| E3 | C^- | C^- | SC^+ | C^+ | SC^- |
| E4 | C^+ | C^- | S | SC^- | S |
| E5 | C^- | C^+ | S | S | S |
| E6 | SC^+ | SC^+ | SC^- | SC^+ | S |
| E7 | S | S | S | S | S |

Tableau II.2 : États symboliques associés à chaque expression.

II.4.2 Méthode par masque pour le visage

L'idée c'est d'utiliser un masque pour l'enregistrement dans une reconnaissance de visage. Ce système n'est pas nouveau. Par exemple, [Bla03] utilise un masque 3D pour enregistrer la vue frontale avec la vue de profil. Certains masques existants, sont bien conçus pour le montage surface déformable selon l'une des actions musculaires. Toutefois, lorsqu'on travaille avec des images face planes, il est difficile d'atteindre d'un bon enregistrement pour tous les sommets densément placés. En outre, bien que ces masques sont à morphologie souple. Les masques triangulaires sont avantageux dans un mappage de texture.

Par conséquent, on utilise une procédure simplifiée et à masque triangulaire, comme représenté dans la Figure. Sur ce masque, les triangles gris correspondent à des régions des sourcils, yeux, nez et bouche.

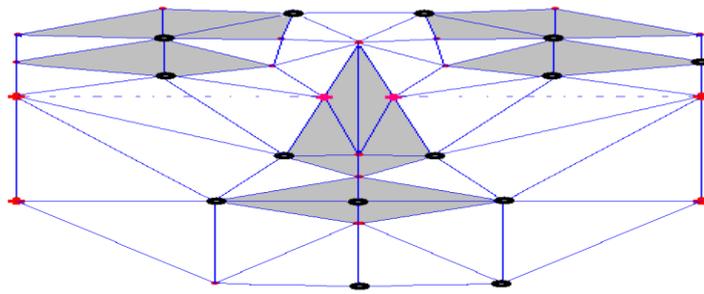


Fig II.6 : Masque pour le visage.

Ils sont intentionnellement à de plus petites façon à saisir les caractéristiques plus détaillées. Ce masque contient seulement 34 sommets et 51 triangles. La Figure ci dessous donne deux exemples d'un tel processus d'enregistrement [Xio08].



Fig II.7 : Masques de montage. (a), (c): manuellement marqueurs placés (points blancs), (b), (d): équipés masques.

II.4.3 Méthode par le modèle d'apparence (AAM)

Franck et al [Fra04] ont travaillé sur reconnaissance et synthèse d'expressions faciales par modèle d'apparence. Et dans ce travail ils traitent une application nouvelle des modèles actifs d'apparence pour l'analyse et la synthèse de visages expressifs, ainsi que pour la reconnaissance d'expressions faciales. Ils considèrent les six expressions universelles que sont les expressions de joie, colère, peur, dégoût, tristesse et surprise. Après une description de ce type de modèle (calculé à l'aide de trois ACP ou d'une seule), ils proposent une méthode originale d'analyse et de synthèse permettant, à partir d'une seule photo, d'annuler l'expression d'un visage quelconque, puis de synthétiser une expression faciale artificielle sur ce même visage. Ils proposent pour cela deux approches permettant la modélisation des expressions faciales. Enfin, l'intérêt du modèle actif d'apparence pour la reconnaissance automatique d'expressions faciales est montré.

II.4.4 Autres méthodes pour la reconnaissance d'expressions faciales

D'après notre étude bibliographique nous réalisons qu'il y'a plusieurs façons d'utiliser les étapes du système de reconnaissance. Tous les systèmes de reconnaissance d'expressions faciales se composent de trois modules principaux :

- Extraction des informations,
- Reconnaissance d'expressions,
- Synthèse d'expressions,

a) Extraction des informations : Black et Yacoob][**Bla97**] utilisent des modèles locaux paramétriques pour représenter le mouvement des visages. Ils estiment le mouvement relatif des traits caractéristiques dans le repère du visage. Les paramètres de ce mouvement servent par la suite à représenter l'expression faciale. De manière identique, Cohn *et al.*[**Coh98**] utilisent un algorithme hiérarchique pour effectuer le suivi des traits caractéristiques par estimation du flux optique. Les vecteurs de déplacement représentent l'information sur les changements d'expression faciale. De même, Padgett et Cottrell [**Pad98**] utilisent des gabarits d'oeil et de bouche, calculés par Analyse en Composantes Principales d'un ensemble d'apprentissage, en conjonction avec des réseaux de neurones. D'autre part, Hong *et al.* [**Hon98**] utilisent un modèle global basé sur des graphes étiquetés construits à partir de points de repère distribués sur le visage. Les noeuds de ces graphes sont formés par des vecteurs dont chaque élément est la réponse à un filtrage de Gabor extraite en un point donné de l'image. Finalement, Cootes *et al.* [**Coo01**] utilise une représentation par modèle actif d'apparence (AAM) pour extraire automatiquement des paramètres caractérisant le visage.

b) Reconnaissance d'expressions : Après avoir détecté le visage et extrait les informations pertinentes, l'étape suivante consiste à identifier l'expression faciale affichée. Pour classer l'expression faciale dans l'une des six catégories de base citées dans [**Ekm99**] en plus de la catégorie neutre, Hong *et al.* [**Hon98**] part du principe que deux personnes qui se ressemblent affichent la même expression faciale de manière similaire. Un graphe étiqueté est attribué à l'image de test puis la personne connue la plus proche est déterminée à l'aide d'une méthode de mise en correspondance de graphes élastiques. La galerie personnalisée de cette personne est alors utilisée pour reconnaître l'expression faciale de l'image de test. Un graphe étiqueté par des réponses de filtres de Gabor est par ailleurs utilisé par Lyons *et al.* [**Lyo99**] et Bartlett *et al.* [**Bar03**]. L'ensemble des graphes construits sur un ensemble d'apprentissage est ensuite soumis à une ACP puis analysé à l'aide d'une analyse discriminante linéaire (ADL) afin de séparer les vecteurs dans des classes ayant des attributs faciaux différents. Le graphe étiqueté de l'image testée sera alors projeté sur les vecteurs discriminants de chaque classe

afin de déterminer son éventuelle appartenance à cette classe. Dans une finalité identique, Essa et Pentland [Ess97] extraient des gabarits spatiotemporels de l'énergie du mouvement du visage pour chaque expression faciale. Le critère de similarité repose sur la distance Euclidienne entre ces gabarits et l'énergie du mouvement de l'image observée. Heisele, Ho et Poggio [Hei01] utilisent des machines à vecteurs de support (SVM) dans le cadre de la reconnaissance de visages par des méthodes globales ainsi que par des méthodes reposant sur des traits caractéristiques. De manière identique, l'algorithme de reconnaissance de visages FaceIt est basé sur une technique d'analyse locale des traits caractéristiques (LFA) développée par Penev et Atick [Pen96]. Draper *et al.* [Dra00] compare les performances de l'analyse en composantes principales et de l'analyse en composantes indépendantes (ICA) pour la reconnaissance de visages et d'expressions faciales en se basant sur le système de codage FACS. Par contre, Yang [Yan02] utilise une représentation par noyaux (KPCA) pour la reconnaissance de visages. Finalement, Edwards, Cootes et Taylor [Edw98] utilisent le modèle actif d'apparence pour reconnaître l'identité d'un individu observé de manière robuste par rapport à l'expression faciale ainsi que l'illumination et la pose. Pour ceci, le critère de similarité utilisé repose sur la distance de Mahalanobis, et une ADL est appliqué afin de maximiser la séparation des classes.

c) Synthèse d'expressions : La synthèse d'expressions faciales est une tâche difficile compte tenu de la complexité de la forme et de la texture des visages. De plus le visage présente un grand nombre de rides et de plis ainsi que des variations subtiles de forme et de texture qui ont une importance cruciale dans la compréhension et la représentation des expressions faciales. Dans cette perspective, les techniques d'interpolation offrent une approche intuitive pour l'animation de visages. Pighin *et al.* [Pig96] utilise des techniques de morphing 2D combinées avec des transformations d'un modèle géométrique 3D, pour créer des modèles faciaux réalistes tridimensionnels à partir de photographies, et pour construire des transitions lisses entre les différentes expressions faciales. Dans la même optique, Chen *et al.* [Che95] applique le morphing au cas 3D. En outre, dans le cadre du "Video-Rewrite", Bregler *et al.* [Bre97] utilise des techniques de suivi de points 2D sur la bouche d'un orateur dans une séquence d'apprentissage et des techniques de morphing pour combiner ces mouvements dans une vidéo finale montrant une personne différente prononçant les mêmes paroles. Dans une finalité analogue, Ezzat *et al.* [Ezz02] utilise une représentation par modèle déformable multidimensionnel et une technique de synthèse de trajectoire pour modifier les mouvements de la bouche d'un visage parlant. Cette représentation est capable de synthétiser des configurations inconnues de lèvres parlantes "vidéo-réalistes" à partir d'une séquence initiale, en utilisant des techniques de morphing. Dans la même optique, Kang *et al.* [Kan02]

utilise le modèle actif d'apparence en conjonction avec des techniques de régression linéaire pour annuler l'expression faciale d'un visage dans le but d'améliorer les performances de la technique de reconnaissance de visages par AAM décrite dans. On propose une variante au modèle AAM décrit dans [Coo01] pour la représentation de visages en n'utilisant qu'une seule ACP. On montre que cette représentation plus directe donne des résultats sensiblement comparables au modèle AAM standard pour la représentation et la reconnaissance d'expressions faciales. On propose également une nouvelle approche de la reconnaissance d'expressions faciales basée sur le modèle AAM ainsi que sur sa variante. Finalement, on propose une extension de la méthode décrite dans [Kan02] pour l'annulation de l'expression faciale, à une application de synthèse de nouvelles expressions faciales. De plus on introduit une nouvelle méthode d'interpolation, basée sur la représentation par le modèle actif d'apparence AAM, pour la synthèse et l'annulation d'expressions faciales.

La reconnaissance faciale est l'un des sujets les plus étudiés de manière intensive dans la vision par ordinateur et reconnaissance de formes. L'expression du visage, qui change la géométrie du visage, a généralement un effet négatif sur la performance d'un système de reconnaissance faciale. D'autre part, la géométrie du visage est un indice utile pour la reconnaissance. Prenant en compte ces paramètres, on utilise l'idée de séparer la géométrie de l'information de texture dans une image de visage et le modèle des deux types des information en les projetant dans des espaces distincts de l'APC qui sont spécialement conçus pour capturer les traits distinctifs entre les différents individus. Par la suite, les attributs de texture et la géométrie sont recombinaés pour former un classificateur qui est capable de reconnaître des faces avec différentes expressions. Enfin, en étudiant la géométrie du visage, on est en mesure de déterminer quel type d'expression faciale a été réalisée, ainsi construire une classificateur expression.

La sélection des fonctionnalités est l'étape la plus importante à succès d'analyser et de reconnaître les expressions faciales automatiquement. Deuxièmement, il est également très important pour extraire les caractéristiques de régions de surface qui contient les informations discriminantes. Les caractéristiques optimales devraient réduire au minimum au sein de la classe des variations d'expressions tout en maximisant entre les variations de la classe. Dans la littérature, différentes méthodes sont utilisés pour extraire les traits du visage et ces méthodes peuvent être classées soit comme méthodes en fonction apparence ou géométriques basées sur les caractéristiques des méthodes où les formes et les emplacements de composantes faciales extraites forment un vecteur [Tin05] de fonctionnalité. L'une des méthodes largement étudié pour extraire les informations apparence sont basées sur le Gabor ondelettes [Lit05 ,Eva04, Don99]. Littlewort et al. [Lit05] a montré une précision de

reconnaissance élevé (97% pour Cohn-Kanade base de données d'expression du visage [Kan00]) pour le visage expressions utilisant les fonctions de Gabor. Ils ont proposé d'extraire les caractéristiques de Gabor sur l'ensemble du visage. AdaBoost a été utilisé pour sélectionner les sous-ensemble de fonctionnalités. Tian [Eva04] a utilisé les ondelettes de Gabor de multi-échelles et multi-orientation à la «différence» des images. Les images de différence ont été obtenues en soustrayant une trame expression neutre du reste des trames de la séquence. Donato et al. [Don99] a utilisé la technique de la division de l'image faciale dans deux: face supérieure et inférieure pour extraire une représentation plus fine de Gabor pour la classification. En règle générale, l'inconvénient d'utiliser les filtres de Gabor est qu'il produit de très grand nombre de caractéristiques et nécessite à la fois du temps et beaucoup de mémoire pour convoluer les images de visage avec une banque de filtres de Gabor pour extraire les coefficients multi-échelles et multi-orientation. Récemment, les méthodes de description texture et de classification à savoir Local Motif Binary (LBP) [Oja96] et de la phase locale de quantification(PLQ) [Hei08] sont également étudiés pour extraire l'apparence du visage à base de caractéristiques. Zhao et al. [Zha07] a proposé l'utilisation du modèle de texture de modèles de volume binaires locales (VLBP) une extension de LBP, pour la reconnaissance des expressions. Les auteurs ont proposé d'utiliser uniquement les co-occurrences de motifs binaires locales sur trois plans orthogonaux (LBP-TOP), afin d'améliorer l'applicabilité de la méthode de calcul en réduisant la complexité. Moyenne précision de reconnaissance faciale expression de 96,26% a été atteinte de six expressions universelles avec leur modèle proposé utilisant Cohn-Kanade visage base de données d'expression. Liao et al. [Lia06] ont proposé d'utiliser deux ensembles de caractéristiques pour la classification d'expression. Le premier ensemble a été obtenu par LBP et le second ensemble de caractéristiques est obtenu par l'analyse discriminante linéaire (LDA). Ils ont testé leur modèle sur la base de données JAFFE [Lio99] et la moyenne précision atteinte de la reconnaissance de 94,59% pour les sept expressions faciales (Six universelle et une expression neutre). Jiang et al. [Jia11] a étendu l'idée de LBP-TOP au PLQ-TOP et ont montré que la performance de système basé sur LPQ vaut mieux que le système basé sur LBP. Pour les méthodes géométriques basées sur les caractéristiques [Act05, Pat05, Bai09], des formes et emplacements des composants faciaux sont extraites pour former un vecteuriel de caractéristiques. Pour la reconnaissance des expressions, Zhang et al. [Act05] a mesuré et suivi le mouvement du visage à l'aide de Filtres Kalman. Pour atteindre la tâche de reconnaissance, ils ont également modélisé les comportements temporels des expressions faciales à l'aide réseaux bayésiens dynamiques (DBN). Dans [Pat05] les auteurs ont présenté l'unité d'action de détection régime (UA) en classant les caractéristiques, calculées à partir de

"filtre à particules" suivi de repère points de visage. Ils ont formé le système MMI sur l'expression du visage-base de données [Rad05] et testés sur la base de données Cohn-Kanade [Kan00] et atteint un taux de reconnaissance de 84%. Bai et al. [Bai09] extrait des informations de forme que l'aide Histogramme Pyramidal des gradients d'orientation (Phog) et a montré l'exactitude "sourire" de détection plus élevé que 96,7% à l'aide Cohn-Kanade base de données.

II.5 État de l'art des techniques récentes de reconnaissance d'expression faciale

Hong-Viet [Han06] on travaillé en 2006 sur la **reconnaissance multimodale de gestes de communication non verbale**. Le contexte général est de présenter leurs contributions en reconnaissance automatique des expressions faciales à partir d'information vidéo. Premièrement, on doit extraire les caractéristiques de l'image du visage, puis on va les utiliser comme entrée du système de classification. La performance attendue est d'environ 87%. Mais ils restent des limitations. Toutes les méthodes nécessitent une intervention manuelle pour la détection et la normalisation. Et toutes les méthodes vont mettre en correspondance les expressions faciales avec six expressions basiques. Le problème est qu'il existe des déformations faciales qui ne correspondent pas à ces six expressions basiques. Donc, le système doit tenir compte d'une nouvelle classe d'expressions à savoir la classe des expressions inconnues. Dans le travail de Hammal [Ham06], la reconnaissance des expressions faciales utilise le Modèle des Croyances Transférable.

Ligang et al. [Lig11] ont étudiés en 2011 la **Géométrie vs Apparence pour discrimination posée par rapport version spontanée d'une émotion**. Les émotions spontanées des expressions faciales diffèrent de celles qui se posent dans l'apparence (séquences et mouvements de tête). Les images fixes ne peuvent pas fournir des informations de mouvement directement. Cependant, indirectement, les distances entre les points clés sur un visage extraites d'une image fixe à l'aide de modèles à forme active peuvent capter une partie en mouvement et changements de pose. Cette information est superposée sur l'information non-rigide du mouvement du visage qui fait également partie de l'expression. L'information géométrique permet d'améliorer la discrimination entre spontanées et les expressions posées du visage. On étudie les performances d'un système de vision artificielle pour discrimination entre les versions qui se posent et spontanées des six émotions de base qui utilise les fonctions d'apparence et les fonctionnalités relatives aux EIPDFAP géométriques. Des résultats expérimentaux sur la base de données NVIE démontrent que la fusion de l'information

géométrique ne conduit qu'à une amélioration marginale par rapport à l'apparence caractéristique. Utilisation des fonctions de fusion, la surprise est le plus facile émotion à distinguer (83,4% de précision), tandis que le dégoût est le plus difficile (76,1%). Les résultats trouvés sur différentes régions importantes du visage entre discrimination posé par rapport à version spontanée d'une émotion et la classification par rapport à la même émotion et d'autres émotions. La répartition des fonctions relatives aux EIPD sélectionnés montre que la bouche est plus importante pour la tristesse, tandis que le nez est plus important pour la surprise. Cependant, à la fois le nez et la bouche sont importants pour le dégoût, la peur et le bonheur. Sourcils, yeux, nez et la bouche sont importants pour la colère.

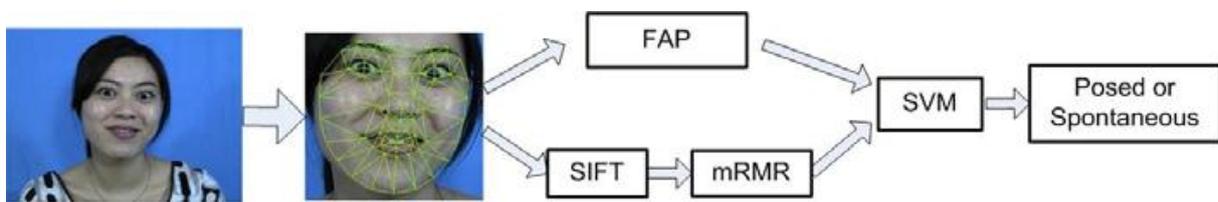


Fig II.8 : Cadre du système d'évaluation

- Paramètres D'animation Faciale (**FAP**)
- minimal Redundancy Maximal Relevance (**mRMR**)
- Machine Vecteur de Support (**SVM**)

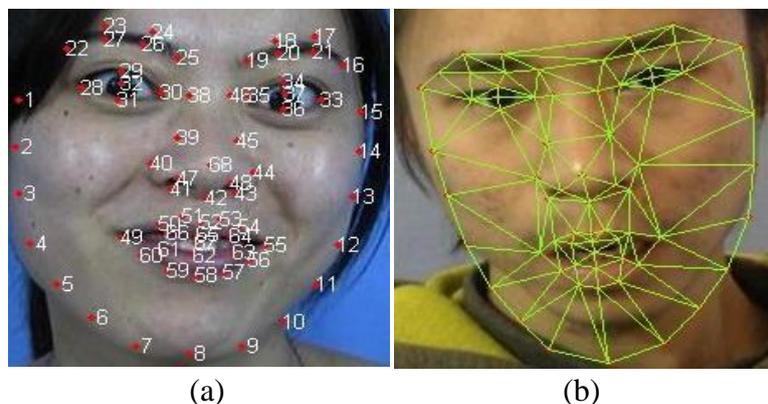


Fig II.9 : (a) 68 points de référence pour la formation et (b) les résultats de détection ASM

Amit et al. [Ami12] ont étudié une méthode intitulée *type-2 Intervalle modèle flou pour reconnaissance d'émotion de visage expressif*. Le document propose une nouvelle approche pour la reconnaissance des émotions de l'expression du visage d'un sujet en construisant un intervalle de type-2 modèle flou. Un type-2 intervalle flou de l'espace visage est d'abord construit avec les connaissances de base des traits du visage de sujets différents pour les émotions différentes. Le flou de l'espace visage ainsi créé comprend des distributions d'adhésion primaire pour les fonctions m visage, obtenus à partir de n sujets, chacun ayant 1-

cas de l'expression du visage pour une émotion donnée. Deuxièmement, l'émotion d'une expression du visage inconnu est déterminée sur la base du consensus des caractéristiques mesurées du visage avec le flou face space. La précision de la classification de la méthode proposée est aussi élevée que 88,66%. La figure II.10 explique les traits du visage sur une image sélectionnée visage.

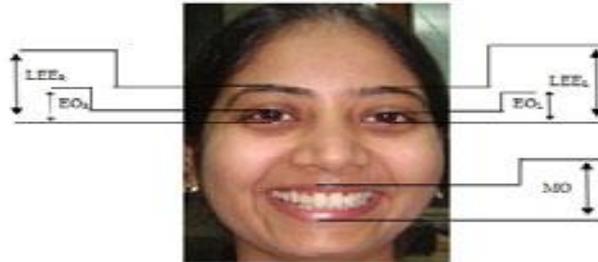


Fig II.10 : Caractéristiques Faciales

La recherche a été effectuée avec succès ces derniers temps de **combiner les méthodes caractéristiques extraites à l'aide de méthodes basées sur l'apparence et basées sur les caractéristiques géométrique [Kot08, Dha11]**. Dans l'ensemble des méthodes pour la reconnaissance des expressions faciale automatique aucune d'entre elles ne peut imiter le système visuel humain. Plutôt toutes les méthodes, utilisent beaucoup de temps de calcul sur l'ensemble du visage ou divise l'image de visage basée sur une heuristique mathématique ou géométrique pour l'extraction de caractéristiques. La tâche d'analyse d'expression et la reconnaissance devrait être faite de manière favorable : seules certaines régions sont sélectionnées pour le traitement ultérieur (par exemple régions saillantes), comme il arrive dans le système visuel humain. Ainsi, on essaye de déterminer statistiquement quelles sont les régions faciales saillantes selon la vision humaine en effectuant une expérience psychovisuelle. L'expérience a été réalisée à l'aide d'eye-tracker qui enregistre les fixations et saccades des observateurs de l'homme comme ils regardent la collection des vidéos montrant six expressions faciales universelles. Les régions saillantes du visage pour les expressions spécifiques ont été déterminés par l'analyse de la fixation des données.

On propose d'extraire les caractéristiques de la région saillante du visage seulement en utilisant l'histogramme pyramidal de gradients d'orientation [Bos07]. La précision de la reconnaissance FER réalisé peut atteindre 99,5% en utilisant Cohn-Kanade base de données [Kan00]. L'avantage de l'extraction de caractéristiques des régions saillantes du visage, c'est que le cadre peut être utilisé pour des applications en temps réel. Projet de cadre traite 4 fps (frames per second) en utilisant Matlab 7,6 pc sous Windows avec 1,8 GHz et 1 Go de

RAM. Les processus de la machine mêmes 2fps pour extraire les mêmes caractéristiques à partir de l'image de visage en entier [Riz12].

Rizwan et al.[Riz12] ont aussi travaillé en **2012** sur **une méthode de reconnaissance des expressions du visage basée sur la perception**. Les humains peuvent reconnaître très facilement les expressions du visage en temps réel. Toutefois, la reconnaissance fiable et rapide des expressions faciales en temps réel est une tâche difficile pour un ordinateur. Nous présentons une nouvelle approche de reconnaissance de trois types d'expressions faciales qui se base sur l'idée de ne considérer que de petites régions du visage bien définies pour en extraire les caractéristiques. Cette proposition est basée sur une étude psycho-visuel expérimental menée avec un eye-tracker. Les mouvements des yeux de quinze sujets ont été enregistrés dans des conditions de visualisation libre d'une collection de 54 vidéos montrant six expressions faciales universelles. Les résultats de cette étude montrent que pour certaines expressions du visage une unique région est perceptuellement plus attractive que les autres. Les autres expressions montrent une attractivité pour deux ou trois régions du visage. Cette connaissance est utilisée pour définir une méthode de reconnaissance se concentrant uniquement sur certaines régions perceptuellement attrayantes du visage et ainsi réduire par un facteur de deux les temps de calcul. Les résultats montrent une précision de reconnaissance automatique de trois expressions de 99.5% sur la base de données d'expression faciale Cohn-Kanade.

II.6 Tableau comparatif des travaux récents sur la REF

REF : Reconnaissance d'expressions faciales.

| Source | Année | Base | Méthode | Taux de reconnaissance |
|--|-------|---|------------------|------------------------|
| Zilu Ying Beihang university, Beijing | 2009 | JAFFE | LBP et LE | 70.48% |
| Ligang Zhang. Queensland University of Technology. Australia | 2011 | Le naturel d'expression visible et infrarouge du visage NVIE | FAP SIFT+mRMS | 76.1% 83.4 % |
| AnishaHalder. Jadavpur University | 2011 | | GT2FS | 96.67% |

| | | | | |
|---|------|-------------------------|--|--------------------------------|
| Amit Konar1. Jadavpur University | 2012 | | IT2FS | 88.66% |
| Rizwan Ahmed Khan. Université de Lyon | 2012 | CK (Cohn-Kanade) | SVM C4.5 Decision Tree RF | 97.3% 97.6% 99.5% |

Tableau II.3 : Comparaison des performances des algorithmes proposés sur les bases de visages.

- Japanese Female Facial Expression(**JAFFE**).
- facial animation parameters(**FAP**).
- Scale-invariant feature transform(**SIFT**).
- minimal redundancy maximal relevance(**mRMS**).
- General Type-2 Fuzzy sets(**GT2FS**).
- Interval Type-2 Fuzzy Face-Space(**IT2FS**).
- Random Forest(**RF**)

II.7 Conclusion

Ce chapitre nous a permis de tirer un certain nombre d'enseignements du processus de reconnaissance d'expressions faciales. Après le développement des principales difficultés inhérentes à la reconnaissance d'expressions faciales, nous avons présenté les principales le système de reconnaissance d'expressions faciales et un état de l'art des techniques proposées jusqu'à ce jour. Nous présentons, dans le chapitre suivant.

Chapitre III

Etude du modèle de reconnaissance de visage avec expression

III.1 Introduction

Comme nous l'avons évoqué au chapitre précédent, un système automatique de reconnaissance des expressions faciales s'effectue en cinq étapes principales : prétraitement, extraction des données, analyse, classification, décision. Certaines techniques de traitements d'images peuvent être communes à plusieurs étapes. Par exemple, l'extraction des caractéristiques faciales (yeux, nez, bouche) est utilisée aussi bien pour la détection que pour l'identification de visages. Par ailleurs, les étapes de détection de visage et d'extraction de caractéristiques peuvent être exécutées simultanément. Cela dépend notamment de la nature de l'application, de la taille de la base d'apprentissage, et des conditions de prise de vue (bruit, occultation, etc.). Enfin, les techniques de traitement utilisées dans chaque étape (Figure III.1) sont très critiques pour les applications biométriques, et doivent, par conséquent, être optimisées pour améliorer les performances du système global.

III.2 Présentation du système d'analyse

Le système de reconnaissance de visage en présence d'expression, comme tout système d'analyse d'image se présente sous les étapes suivantes:

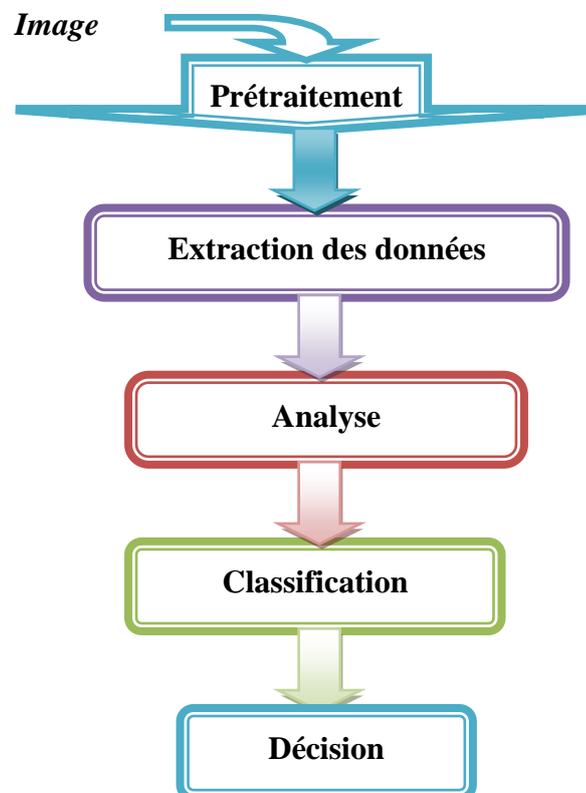


Fig III.1 : Les étapes dans le processus de reconnaissance des expressions faciales.

Dans ce qui suit nous allons détailler chaque étape du système de reconnaissance de l'expression faciale, et nous présenterons les principales difficultés rencontrées.

III.2.1 Prétraitement

L'efficacité des systèmes biométriques basés sur l'identification de visage dépend essentiellement de la méthode utilisée pour localiser le visage dans l'image. Dans la littérature scientifique, le problème de localisation de visages est aussi désigné par la terminologie "détection de visages". Plusieurs travaux de recherches ont été effectués dans ce domaine. Ils ont donné lieu au développement d'une multitude de techniques allant de la simple détection du visage, à la localisation précise des régions caractéristiques du visage, tels que les yeux, le nez, les narines, les sourcils, la bouche, les lèvres, les oreilles, etc. Nous détaillerons quelques unes de ces méthodes dans la section. Cependant, les solutions proposées jusqu'à maintenant sont loin d'être satisfaisantes car elles fonctionnent uniquement dans des environnements contrôlés, et par conséquent elles ne gèrent pas la variabilité des conditions d'acquisition de la vie quotidienne, notamment :

- **La pose** où les images d'un visage changent en fonction de l'orientation de ce dernier.
- **La présence ou absence des composantes structurales** les caractéristiques faciales tels que la barbe, la moustache, et les lunettes causent une grande variabilité des composantes structurales du visage, notamment au niveau de la forme, de la couleur, et de la taille.
- **Les occultations** les visages peuvent être partiellement occultés par d'autres objets. En effet, dans une image contenant un groupe de personnes par exemple, des visages peuvent partiellement masquer d'autres visages.
- **Les conditions d'illumination** des facteurs tels que l'éclairage (distribution de la source de lumière, son intensité, son spectre) et les caractéristiques de l'appareil photographique affectent l'aspect d'un visage dans l'image acquise [Sou08].

III.2.2 Extraction de caractéristiques des visages avec expressions

L'extraction des caractéristiques telles que les yeux, le nez, la bouche est une étape prétraitement nécessaire à la reconnaissance d'expression faciale. On peut distinguer deux pratiques différentes : la première repose sur l'extraction de régions entières du visage, elle est souvent implémentée avec une approche globale de reconnaissance de visage. La deuxième pratique extrait des points particuliers des différentes régions caractéristiques du visage, tels

que les coins des yeux, de la bouche et du nez. Elle est utilisée avec une méthode locale de reconnaissance et aussi pour l'estimation de la pose du visage.

Par ailleurs, plusieurs études ont été menées afin de déterminer les caractéristiques qui semblent pertinentes pour la perception, la mémorisation et la reconnaissance d'un visage humain. Par exemple, les caractéristiques pertinentes rapportées sont : les cheveux, le contour du visage, les yeux et la bouche. Cette étude a également démontré le rôle important que joue le nez dans la reconnaissance faciale à partir des images de profil. En effet, dans ce cas de figure, il est évident que la forme distinctive du nez est plus intéressante que les yeux ou la bouche. Les auteurs ont particulièrement établi que la partie supérieure du visage est plus utile pour la reconnaissance faciale que la partie inférieure [Oua11].

Dans le cadre de notre travail nous optons pour le descripteur suivant pour caractériser les images de visage de la base de données utilisées **CASIA**.

III.2.2.1 Les ondelettes de Gabor

Une alternative à l'usage de la morphologie mathématique est l'application d'une transformée d'ondelette. Les ondelettes de Gabor sont connues comme le moyen d'analyse espace-fréquence qui minimise l'incertitude de Heisenberg à savoir si on est précis dans l'analyse de l'information dans l'espace du pixel on le sera moins dans l'espace fréquentiel et vice versa. Cette caractéristique fait des ondelettes de Gabor un moyen puissant pour l'analyse de la texture. Elles ont été développées au début pour l'analyse des signaux à une dimension. Leurs expression générale est :

$$W(t, t_0, \omega) = e^{-\sigma(t-t_0)^2} e^{-i\omega(t-t_0)} \quad (\text{éq. 3.1})$$

Où ω est la fréquence et $i^2 = -1$. Cette ondelette est en fait une gaussienne modulée par une fonction sinusoïdale. La généralisation de cette fonction en 2D donne une fonction gaussienne à deux dimensions modulée par une harmonique. Les ondelettes de Gabor 2D ont la caractéristique d'être modulables par différents paramètres qui contrôlent leurs orientations, fréquences, phases de la sinusoïde, le support temporel et l'enveloppe de la gaussienne [Wal09]. L'équation générale d'une ondelette de Gabor en 2D est :

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{x'}{\lambda} + \varphi)} \quad (\text{éq. 3.2})$$

Ou : $x' = x \cos \theta + y \sin \theta$ et $y' = -x \sin \theta + y \cos \theta$

Il y a donc 5 paramètres qui contrôlent l'analyse par ondelette. Ce jeu de paramètres permet dès lors une analyse complète de la texture d'une région de l'image.

1) L'orientation de l'ondelette (θ)

Ce paramètre fait pivoter l'ondelette autour de son centre. L'orientation de l'ondelette (figure III.2) détermine l'angle des contours ou des lignes de l'image auxquelles l'ondelette va être sensible.

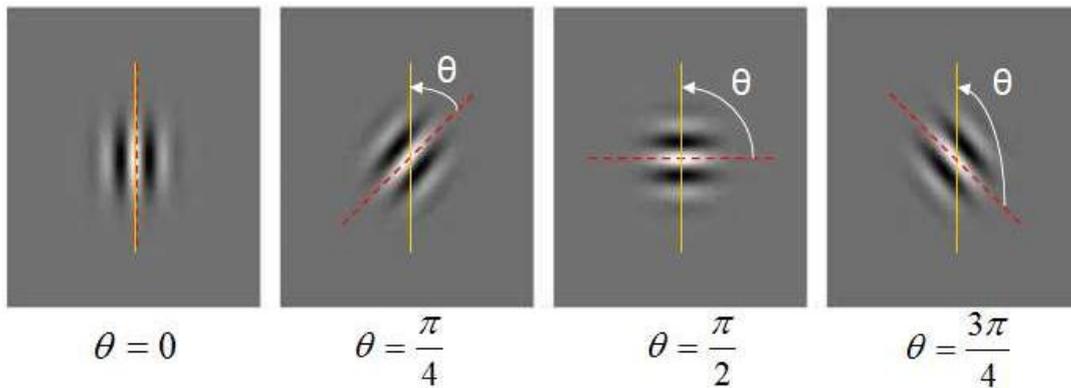


Fig III.2 : Orientation de l'ondelette.

2) La fréquence centrale de l'ondelette (λ) :

Ce paramètre spécifie la longueur d'onde du cosinus ou inversement la fréquence centrale de l'ondelette (figure III.3). 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,

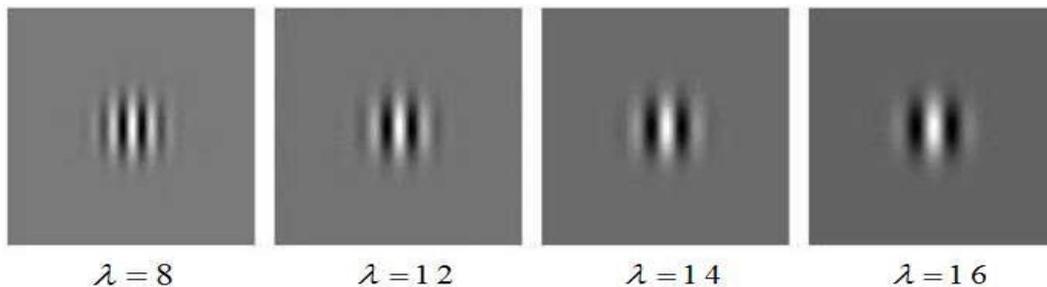


Fig III.3 : Fréquence centrale de l'ondelette.

Dans notre travail nous utilisons 40 filtres : 5 fréquences et 8 orientations. Ces filtres seront présentés dans le chapitre suivant.

III.2.3 Les algorithmes de réduction l'espace de données

Une image du visage est un signal à 2 dimensions, acquis par un capteur digital (caméra numérique, scanner...). Ce capteur codera la couleur ou l'intensité des différents points de l'image dans une matrice de pixels à deux dimensions.

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

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

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

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

L'appartenance à un même sous-espace et la redondance statistique qui caractérise les visages expliquent l'intérêt des chercheurs pour les méthodes de reconnaissance basées sur les approches par réduction d'espace, appelées également "les approches globales". Le fait que les visages soient confinés dans un sous-espace permet de définir une base (orthogonale, ou non orthogonale) dans cet espace. Cette base devrait représenter de façon plus pertinente n'importe quel point de l'espace des visages. L'avantage de la redondance statistique est qu'elle permet une extraction d'une structure simple des caractéristiques importantes et pertinentes de l'image du visage. Cette structure permettrait de représenter le visage tout en gardant l'information la

plus importante, et par conséquent, de réduire la dimensionnalité de l'espace visage. Tout l'intérêt des approches globales est la construction de cette base de projection qui permettra de comparer, de reconnaître ou d'analyser l'information essentielle des visages [Ano09].

III.2.3.1 Analyse en Composantes Principales (PCA)

L'analyse en composantes principales (**PCA**) est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard (Eigen Picture). Ces poids sont obtenus en projetant l'image dans un espace de visage engendré par les visages propres (Eigen faces). Ainsi au lieu de stocker un vecteur image de taille Résolution N , on stocke un vecteur de taille $R \leq M$ nombre d'image d'apprentissage. Dans l'approche **PCA** la normalisation d'éclairage est toujours indispensable.

Les visages propres sont des images de la même taille que les nombres d'images d'apprentissage et qui montrent des visages ayant un aspect fantomatique. Mathématiquement, elles sont les composantes principales de la distribution des visages, ou les vecteurs propres de la matrice de covariance de l'ensemble des images de visage. Chaque image de visage de l'ensemble d'apprentissage peut être exactement représentée en termes de combinaison linéaire des visages propres et du visage moyen.

Le nombre possible des visages propres est égal au nombre d'images de visage dans l'ensemble d'apprentissage. Néanmoins les visages peuvent être approximés, en utilisant seulement les meilleurs visages propres (ayant les plus larges valeurs propres qui représentent en fait, la plupart de variance dans l'ensemble d'images de visage), permettant ainsi de diminuer les calculs.

La **PCA** est une technique rapide, simple et populaire dans l'identification de modèle, c'est l'une des meilleures techniques. Les projections de la **PCA** sont optimales pour la reconstruction d'une base de dimension réduite. Cependant, la **PCA** n'est pas optimisée pour la séparabilité (discrimination) de classes. Une alternative qui est l'analyse discriminante linéaire **LDA** tient compte de ceci.

1) Le processus de reconnaissance (Karhunen-Loève) :

Au début de la phase d'apprentissage, on fait l'acquisition, la lecture et la normalisation des images d'apprentissage (de taille \mathbf{N}). Puis on calcule le visage moyen de ces images. On soustrait ensuite le visage moyen de chaque image d'apprentissage normalisée (ces nouvelles images représentent les colonnes d'une matrice $\bar{\mathbf{X}}$). On fait le calcul de la matrice de covariance. On a :

Chaque image est représentée par un vecteur de taille \mathbf{N} .

$$x^i = [x_1^i, x_2^i, \dots, x_N^i]^T \quad (\text{équ 3.3})$$

Où \mathbf{N} est le nombre de pixels dans l'image.

Toutes les images sont centrées en soustrayant l'image moyenne de chaque vecteur image,

$$\bar{x}^i = x^i - m \quad \text{ou} \quad m = \frac{1}{p} \sum_{i=1}^p x^i \quad (\text{équ 3.4})$$

m : La moyenne de toutes les images d'apprentissage.

Ces vecteurs sont combinés, côte à côte, pour créer une matrice de données d'apprentissage de taille $\mathbf{N} \times \mathbf{P}$ (où \mathbf{P} est le nombre d'images d'apprentissage, \mathbf{N} la taille de la vecteur image).

$$\bar{\mathbf{X}} = [\bar{x}^1 \ \bar{x}^2 \ \bar{x}^3 \ \dots \ \bar{x}^p] \quad (\text{équ 3.5})$$

Le calcul de la matrice de covariance qui est définie par la formule :

$$\mathbf{\Omega} = \bar{\mathbf{X}} \bar{\mathbf{X}}^T \quad (\text{équ 3.6})$$

On calcule ensuite les vecteurs propres \mathbf{V} et les valeurs propres \mathbf{D} de la matrice $\mathbf{\Omega}$.

Les valeurs propres de la matrice $\mathbf{\Omega}$ représentent le taux de variation le long de l'axe du vecteur propre associé. Dès lors, on pourra éliminer les vecteurs propres les moins importants sans perte conséquente d'information.

Puis on passe à l'étape de calcul des visages propres selon la formule

$$\mathbf{U} = \bar{\mathbf{X}} * \mathbf{V} * (\mathbf{abs})(\mathbf{D})^{-0.5} \quad (\text{équ 3.7})$$

Et finalement on calcule les poids des visages de la base (de taille \mathbf{M}) : en les projetant dans le sous-espace engendré par les visages propres “*Face Space*”.

Pendant la phase de vérification, on fait l’acquisition, la lecture et la normalisation de l’image de vérification (de taille \mathbf{N}). Puis on soustrait le visage moyen (calculé auparavant lors de la phase d’apprentissage) de l’image de vérification normalisée. Ensuite on calcule le poids de l’image (de taille \mathbf{M}) en utilisant les visages propres comme une base de projection. Et finalement on utilise les distances de similitude [Ano09].

2) Le choix de la dimension \mathbf{M} de l'espace de visage

En général le choix de la dimension \mathbf{M} ou le nombre des vecteurs propres n'est pas facile. En effet, plusieurs techniques ont été utilisées pour déterminer cette dimension intrinsèque \mathbf{M} de l’espace des visages.

- L'utilisation d'un critère d'énergie : \mathbf{M} est choisi tels que les valeurs propres maintenues contiennent 95% de toute l'énergie.

$$\sum_{j=1}^m \lambda_j = 0.95 \sum_{j=1}^{\min(l,n)} \lambda_j \quad (\text{équ 3.8})$$

- La sélection est basée sur les caractéristiques des vecteurs propres : les auteurs notent qu'au-dessus du rang 300, les vecteurs propres encodent principalement le bruit et devraient être jetés pour des tâches de classification de visages.
- Les résultats empiriques basé sur une large gamme de bases de données de visages : il est suggéré de choisir un nombre \mathbf{M} qui représente le 40% du nombre de tous les vecteurs propres.
- Une sélection de type bayésienne est utilisée dans [Min00] pour trouver une dimension correcte. Une approche pragmatique consiste en choisissant \mathbf{M} de telle sorte de minimiser les taux d'erreur de classification sur l'ensemble de données.
- Le choix dépend de l'étape du prétraitement des images. Par exemple, si les visages sont normalisés par une transformation souple de sorte que l'information de la forme soit uniforme comme dans le sous-espace de visages est de dimension basse car seule la variation de l'information de niveau de gris doit être représentée par le modèle.
- La partie du visage utile pour la classification est déterminée en étudiant les rangs statistiques (rank statistics) entre les pixels dans l'image. La localisation bonne de visage donne un région étroite de visage en éliminant le front et le fond et ça donc permet de réduire la dimension \mathbf{M} [Dja06].

- **Inconvénient**

L'**ACP** ne tient pas compte de l'aspect (discrimination) des classes et on ne devrait pas s'attendre ainsi une performance optimale pour des tâches telles que l'authentification de visage en utilisant un tel **ACP** pour le codage.

- **Solution**

Une solution qui a été proposée pour résoudre ce nouveau problème est d'employer l'analyse discriminante linéaire de Fisher (**FLD**) pour le but même de réaliser une séparabilité élevée entre les différentes classes dans lesquels la classification est intéressante.

III.2.3.2 Analyse Discriminante Linéaire(LDA)

L'analyse discriminante linéaire(**LDA**) et les discriminants linéaires de Fisher (**FLD**) relatifs, sont utilisés pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement. Les combinaisons résultantes peuvent être employées comme classificateur linéaire, ou généralement dans la réduction de caractéristiques.

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

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

Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixels. Les combinaisons linéaires obtenues en utilisant **FLD** s'appellent les Fisher faces , en analogie avec les visages propres.

LDA est une technique qui cherche les directions qui sont efficaces pour la discrimination entre les données. La figure III.4 représente un exemple de classification de deux nuages de points. L'axe principal de la LDA est l'axe de projection qui maximise la séparation entre les deux classes. Il est clair que cette projection est optimale pour la séparation des deux classes par rapport à la projection sur l'axe principal calculé par **ACP**.

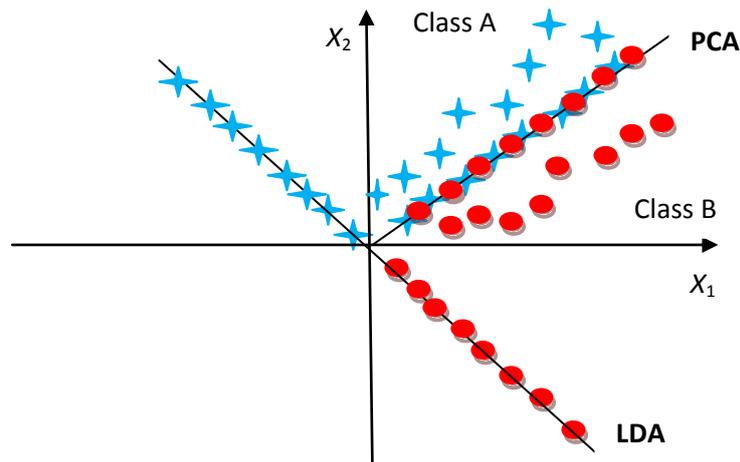


Fig III.4 : Comparaison entre les projections de deux classes de points ("class 1" et "class 2") sur les axes principaux construits par ACP et par LDA.

- **Discriminant linéaire de Fisher (Méthode de la base orthonormée)**

Deux problèmes se présentent en employant le discriminant de Fisher. D'abord, les matrices nécessaires pour le calcul sont très grandes, il en entraîne un temps de calcul très lent et des problèmes possibles dans la précision numérique. Deuxièmement, puisqu'il y a moins d'images d'apprentissages que le nombre des pixels par image ; il est possible de résoudre le problème des vecteurs propres et les valeurs propres d'un rang de matrice déficiente en employant la méthode de décomposition singulière généralisée.

Une solution plus simple est de projeter la matrice de données d'images d'apprentissage dans une base orthonormée de taille $\mathbf{P} \times \mathbf{P}$ (où \mathbf{P} est le nombre d'images d'apprentissage).

Cette projection produit une matrice de données de plein rang qui est beaucoup plus petite ce qui entraîne une diminution du temps de calcul.

Les étapes à suivre pour trouver le discriminant de Fisher pour une base orthonormée sont :

1) Calcul des moyennes

Nous calculons la moyennes des images dans chaque classe (\mathbf{m}_i) et la moyenne de toute les images m

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{j=1}^i x_j^i ; \quad i=1,2,\dots,c \quad (\text{équ 3.9})$$

$$\mathbf{m} = \frac{1}{p} \sum_{i=1}^p \mathbf{x}_i \quad (\text{équ 3.10})$$

P : Le nombre d'images d'apprentissage.

n_i : Le nombre d'images dans chaque classe.

C : Le nombre de classes.

2) *Centrer les images dans chaque classe*

Soustraire la moyenne de chaque classe de l'image de cette classe

$$\forall \mathbf{x} \in X_i, \mathbf{x}_i \in X, \hat{\mathbf{x}} = \mathbf{x} - \mathbf{m}_i \quad (\text{équ 3.11})$$

3) *Centrer les moyennes de chaque classe*

Soustraire la moyenne totale des moyennes de classes.

$$\tilde{\mathbf{m}}_i = \mathbf{m}_i - \mathbf{m} \quad (\text{équ 3.12})$$

4) *Créer une matrice de données*

Combiner tous les vecteurs images, côte à côte, dans une matrice de données.

5) *Trouver la base orthonormée pour la matrice de données*

Cela peut être calculé par la méthode « **Orthogonal-triangular décomposition.** » QR ou en calculant les vecteurs propres de la matrice de covariance (formée par la matrice de données), soit U cette base orthonormée.

6) *Projeter toutes les images centrées dans la base orthonormée*

C'est le produit des images centrées et la matrice transposée des vecteurs de la base orthonormée.

$$\tilde{\mathbf{x}} = \mathbf{U}^T \hat{\mathbf{x}} \quad (\text{équ 3.13})$$

7) *Projeter les moyennes centrées dans la base orthonormée*

$$\tilde{\mathbf{m}} = \mathbf{U}^T \hat{\mathbf{m}}_i \quad (\text{équ 3.14})$$

8) *Calculer la matrice de dispersion intra-classes S_W (within class scatter matrix)*

La matrice de dispersion interclasses (S_W) est la somme des matrices de dispersion des classes. Pour la 1^{ème} classe la matrice de dispersion S_i est calculée par la somme des

matrices de covariance des images centrées projetées.

$$\mathbf{S}_i = \sum_{x \in X_i} \tilde{x} \tilde{x}^T \quad (\text{équ 3.15})$$

La matrice de dispersion interclasses (\mathbf{S}_w) est la somme de toutes les matrices de dispersion.

$$\mathbf{S}_w = \sum_{i=0}^C \mathbf{S}_i \quad (\text{équ 3.16})$$

Où C est le nombre de classe.

9) Calculer de la matrice de dispersion inter-classes \mathbf{S}_B (Between class scatter matrix)

La matrice de dispersion interclasses (\mathbf{S}_B) est la somme de dispersion entre classes.

$$\mathbf{S}_B = \sum_{i=1}^C n_i \tilde{m}_i \tilde{m}_i^T \quad (\text{équ 3.17})$$

Où n_i : Le nombre d'images dans la classe.

\tilde{m}_i : La moyenne des classes projetées dans la base orthonormée ;

Le but est de maximiser les distances inter-classes tout en minimisant les distances intra-classes, ce qui revient à retrouver la matrice de transformation W qui maximise le critère

$J(W) = \frac{W^T \mathbf{S}_B W}{W^T \mathbf{S}_w W}$ donc W est optimale pour :

$$W_{opt} = \mathbf{arg} \max_W \left(\frac{|W^T \mathbf{S}_B W|}{|W^T \mathbf{S}_w W|} \right) = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m] \quad (\text{équ 3.18})$$

10) Résoudre le problème de valeurs propres généralisé

Résoudre le problème généralisé de vecteurs propres (V) et les valeurs propres (Λ) de la matrice de dispersion \mathbf{S}_w et la matrice de dispersion \mathbf{S}_B .

$$\mathbf{S}_B V = \Lambda \mathbf{S}_w V \quad (\text{équ 3.19})$$

La solution est rendue par une analyse en composant principales en calculant les vecteurs propres et les valeurs propres de la matrice $\mathbf{S}_w^{-1} * \mathbf{S}_B$.

11) *Tenir les C-1 premiers vecteurs propres*

Triez les vecteurs propres par leurs valeurs propres associés dans l'ordre décroissant et retenir les premiers **C-1** vecteurs propres. Ceux-ci sont les vecteurs de base de Fisher.

12) *Projeter les images sur les vecteurs propres*

Projeter toutes les images originales dont les visages présentent une déviation (rotation) (non centrées) sur les vecteurs de base de Fisher. D'abord, projeter les images originales (non centrées) dans la base orthonormée et projeter ensuite ces images sur les vecteurs de base de Fisher.

Les images originales dont les visages présentent une déviation (rotation) sont projetées sur cette ligne parce que ceux-ci sont les points que la ligne crée pour les discriminer.

- **Inconvénient**

FLD exige un grand nombre d'échantillons de l'ensemble d'apprentissage pour la bonne généralisation. Quand un tel besoin n'est pas répondu, l'FLD crée un problème de sur-ajustement aux données d'apprentissage et généralise ainsi mal aux nouvelles données de test.

- **Solution**

Le modèle discriminant Linéaire amélioré de Fisher (Enhanced Fisher Model).

III.2.3.3 Le modèle discriminant Linéaire amélioré de Fisher (EFM)

Le modèle discriminant Linéaire amélioré de Fisher la capacité de généralisation de la FLD en décomposant la procédure FLD en diagonalisation simultanée des deux matrices de dispersion intra-classe et inter-classe. La diagonalisation simultanée est une étape sagement équivalente à deux opérations comme l'a souligné Fukunaga [Fuk91] Blanchiment la matrice de dispersion intra-classe et l'application du ACP sur la matrice de dispersion intrre-classe en utilisant les données transformées. Durant l'opération du blanchiment de la matrice de dispersion intra-classe apparaisse dans le dénominateur de la séparabilité des petites valeurs propres qui tendent à capturer du bruit. Pour atteindre des performances améliorées l'EFM préserve un équilibre approprié entre la sélection des valeurs propres (correspondant à la composante principale de l'espace de l'image originale) qui tiennent compte de la plupart de

l'énergie spectrale des données brutes ,c.-à-d., représentation adéquate , et l'exigence que les valeurs propres de la matrice de dispersion intra-classe (de l'espace ACP réduit)ne sont pas trop petites, c.-à-d., meilleure généralisation.

Le choix de rang des composantes principales (m) pour la réduction de la dimension, prend en compte de l'ordre de grandeur de l'énergie spectral. Les valeurs propres de la matrice de covariance fournit un bon indicateur pour répondre au critère de l'énergie ; Il faut ensuite calculer les valeurs propres de la matrices de dispersion intra-classe dans l'espace ACP réduit pour faciliter le choix du rang des composantes principales de sorte que l'exigence de grandeur est respecté. A cette fin, on effectue le FLD par des étapes comme décrit ci-dessous. En particulier, ces étapes FLD permettent de trouver les valeurs propres et les vecteurs propres de $S_W^{-1}S_b$ comme résultat de la diagonalisation simultanée de S_W et S_b . Alors l'étapes de EFM est présenter comme suit :

- Blanchissons d'abord la matrice de dispersion intra-classe :

$$S_W V = V \Lambda \text{ et } V^T V = I \quad (\text{équ 3.20})$$

$$\Lambda^{-1/2} V^T S_W V \Lambda^{-1/2} = I \quad (\text{équ 3.21})$$

Où $V, \Lambda \in \mathbb{R}^{m \times m}$ sont la matrice des vecteurs propres et la matrice diagonale des valeurs propres de S_W respectivement.

Les valeurs propres de la matrice de dispersion intra-classe dans l'espace ACP réduit peut être obtenu en (équation 3.20).

Donc, EFM diagonalise en premier lieu la matrice de dispersion intra-classe S_W (3.20) et (3.21). Notez que V et Λ sont les matrices des vecteurs propres et des valeurs propres correspondants aux vecteurs caractéristiques.

- En second lieu EFM procède à calculer la matrice de dispersion inter-classe comme suit:

$$\Lambda^{-1/2} V^T S_b V \Lambda^{-1/2} = B_b \quad (\text{équ 3.22})$$

- Diagonalisons maintenant la nouvelle matrice de dispersion inter-classe K_b :

$$K_b V_b = V_b \Lambda_b \text{ et } V_b^T V_b = I \quad (\text{équ 3.23})$$

Où $V_b, A_b \in \mathbf{R}^{m \times m}$ sont la matrice des vecteurs propres et la matrice diagonale des valeurs propres de K_b respectivement.

- La matrice de transformation globale de la EFM est définie comme suit :

$$U = V A^{-1/2} V_b \quad (\text{équ 3.24})$$

III.2.4 Comparaisons

III.2.4.1 Comparaisons entre deux vecteurs

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

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

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

$$L_p = \left(\sum_{i=1}^N |x_i - y_i|^p \right)^{1/p} \quad (\text{équ 3.25})$$

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 a :

$$L_1(x, y) = \sum_{i=1}^N |x_i - y_i| \quad (\text{équ 3.26})$$

- **Distance Euclidienne (L2)**

Pour $p=2$, on a :

$$L_2(x, y) = \sum_{i=1}^N |x_i - y_i| \sqrt{\sum_{i=1}^N |x_i - y_i|^2} \quad (\text{équ 3.27})$$

2) Distances dans l'Espace de Mahalanobis

Soit u et v deux vecteurs propres de J_m , issus de l'algorithme PCA, et m et n deux

vecteurs de 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 : $m_i = \frac{u_i}{\sigma_i} = \frac{u_i}{\sqrt{\lambda_i}}$ et $n_i = \frac{v_i}{\sigma_i} = \frac{v_i}{\sqrt{\lambda_i}}$

- **Mahalanobis L1 (MahL1)**

$L1$ est définie par :

$$Moh_{L1}(u, v) = \sum_{i=1}^N |m_i - n_i| \quad (\text{équ 3.28})$$

- **Mahalanobis L2 (MahL2)**

$L2$ est définie par :

$$Moh_{L2}(u, v) = \sqrt{\sum_{i=1}^N |m_i - n_i|^2} \quad (\text{équ 3.29})$$

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

III.2.4.2 Comparaisons entre deux matrices

Dans Yang a proposé une nouvelle mesure de similarité au plus proche voisin pour la reconnaissance de visages. La distance de Yang se base sur la classification de matrices caractéristiques obtenues par l'ACP2D. Cette distance a été, également, adoptée par Visaniet al. [Vis04] et Bengherabi [Ben08] Elle est définie pour deux matrices caractéristiques réduites $Y_i = [y_1^{(i)} y_2^{(i)} \dots y_{d_1}^{(i)}]$ et $Y_j = [y_1^{(j)} y_2^{(j)} \dots y_{d_1}^{(j)}]$ comme suit :

$$d(Y_i, Y_j) = \sum_{k=1}^{d_1} \|y_k^{(i)} - y_k^{(j)}\|_2 \quad (\text{équ 3.30})$$

Où $\|y_k^{(i)} - y_k^{(j)}\|_2$ désigne la distance Euclidienne entre les deux vecteurs $y_k^{(i)}$ et $y_k^{(j)}$ d'où:

$$d(Y_i, Y_j) = \sum_{k=1}^{d_1} \left(\sum_{h=1}^n (y_{hk}^{(i)} - y_{hk}^{(j)})^2 \right)^{1/2} \quad (\text{équ 3.31})$$

Avec $y_k^{(i)} = [y_{1k}^{(i)} y_{21}^{(i)} \dots y_{d_1 k}^{(i)}]$ et $y_k^{(j)} = [y_{1k}^{(j)} y_{21}^{(j)} \dots y_{d_1 k}^{(j)}]$

Dans [Zuo05], Zuo a proposé « l'Assembled Matrix Distance: AMD » définie par:

$$d(Y_i, Y_j) = \sum_{k=1}^{d_1} \left(\sum_{h=1}^n \left((y_{hk}^{(i)} - y_{hk}^{(j)})^2 \right)^{p/2} \right)^{1/p} \quad (\text{équ 3.32})$$

pour laquelle la distance de Yang est obtenue pour $p = 1$ et la distance de Frobenius pour $p = 2$.

III.2.4.3 Machine à Vecteurs de Support (SVM)

SVM (Support Vector Machines) est une nouvelle technique d'apprentissage statistique, proposée par V. Vapnik en 1995. Elle permet d'aborder des problèmes très divers comme le classement, la régression, la fusion, etc.

Depuis son introduction dans le domaine de la Reconnaissance de Formes (RdF), plusieurs travaux ont pu montrer l'efficacité de cette technique principalement en traitement d'image. L'idée essentielle consiste à projeter les données de l'espace d'entrée (appartenant à des classes différentes) non linéairement séparables, dans un espace de plus grande dimension appelé espace de caractéristiques, de façon à ce que les données deviennent linéairement séparables [Moh05]. Dans cet espace, la technique de construction de l'hyperplan optimal est utilisée pour calculer la fonction de classement séparant les classes tels que les vecteurs appartenant aux différentes classes se trouvent de différents côtés de l'hyperplan.

Le but du SVM est donc de trouver une séparatrice qui minimise l'erreur de classification sur l'ensemble d'apprentissage mais qui sera également performante en généralisation sur des données non utilisées en apprentissage. Pour cela le concept utilisé est celui de marge (d'où le nom de séparateurs à vaste marge). La marge est la distance quadratique moyenne entre la séparatrice et les éléments d'apprentissage les plus proches de celle-ci appelés vecteurs de support (figure III.6). Ces éléments sont appelés vecteurs de support car c'est uniquement sur ces éléments de l'ensemble d'apprentissage qu'est optimisée la séparatrice.

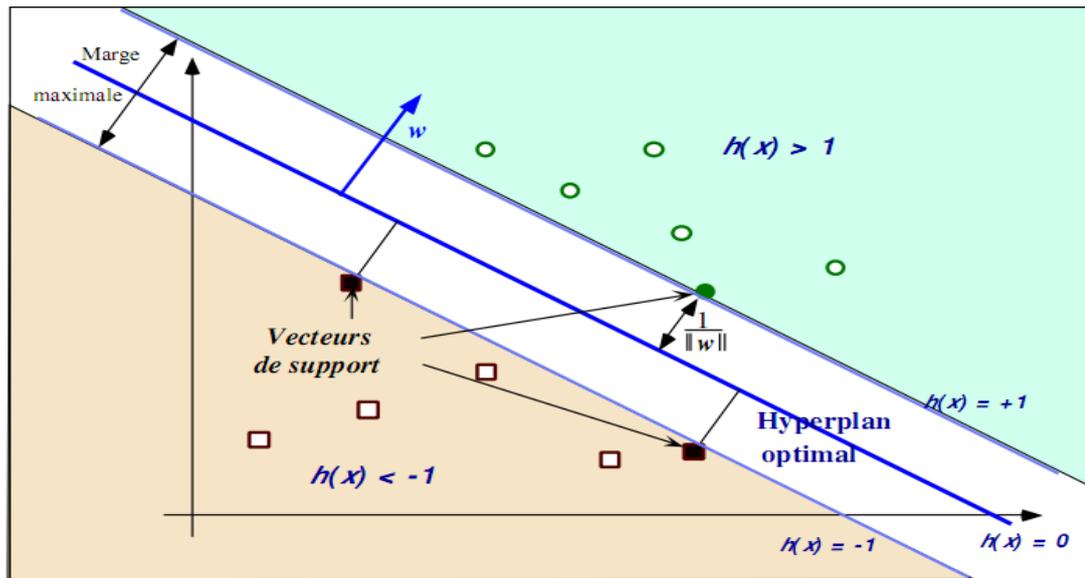


Fig III.5 : Principe de la technique SVM (hyperplan optimal, vecteurs de supports, marge maximale).

Hyperplan optimal : est un Hyperplan qui classe 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.

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 [Gue11].

Tout classifier a pour but de classifier un élément x , ici $x = (s_1, \dots, s_N)$ est un vecteur de scores de dimension N , dans l'une des classes possibles. Dans notre problème il y a deux classes, Client ou Imposteur, dont l'étiquette sera noté y avec $y = -1, 1, -1$ correspondant à la classe des Imposteurs et 1 à la classe des Clients. Le classifier a donc pour but de déterminer f telle que :

$$Y = f(x) \quad (\text{équ 3.33})$$

Le SVM a pour but de trouver la meilleure séparatrice linéaire (en terme d'émarge maximale, c'est à dire la meilleure généralisation) dans l'espace transformée par la fonction de noyau K , c'est à dire de déterminer le vecteur w et la constante b tels que la séparatrice ait pour équation :

$$w \cdot k(x) + b = 0 \quad (\text{équ 3.34})$$

La distance entre un point de l'espace x_i et l'hyperplan d'équation $w \cdot K(x) + b = 0$ est égal

$$\text{à :} \quad h(x_i) = \frac{w \cdot K(x_i) + b}{\|w\|} \quad (\text{équ 3.35})$$

Pour maximiser la marge, il faut donc minimiser $\|w\|$ tout en maximisant $w \cdot K(x_i) + b$ pour les x_i définis comme vecteurs de support. Ces vecteurs de supports sont les x_i pour $i = 1 : m$ de la base d'apprentissage tels que $w \cdot K(x_i) + b = \pm 1$.

La résolution de ce problème d'optimisation est faite par l'utilisation des multiplicateurs de Lagrange où le Lagrangien est donné par :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i (\mathcal{Y}_i(w \cdot k(x_i) + b) - 1) \quad (\text{équ 3.36})$$

Avec les coefficients α_i appelés multiplicateurs de Lagrange. Pour résoudre ce problème d'optimisation, il faut minimiser le Lagrangien par rapport à w et b et le maximiser par rapport à α .

Dans la pratique, il est souvent impossible de trouver une séparatrice linéaire (même dans l'espace transformé par la fonction noyau) car il y a toujours des erreurs de classification. Il a donc été introduit par Vapnik [Vap95] la technique de marge souple. Ce principe de marge souple tolère les mauvais classements par l'introduction de variables ressorts ξ_i qui permettent de relâcher les contraintes sur les éléments d'apprentissage qui ne doivent plus être à une distance supérieure ou égale à 1 de la marge (l'égalité correspondant aux vecteurs de support), mais à une distance supérieure ou égale à $1 - \xi_i$, c'est à dire :

$$\mathcal{Y}_i(w \cdot k(x_i) + b) \geq 1 - \xi_i \quad (\text{équ 3.37})$$

Avec $\xi_i \geq 0$ pour $i = 1 : M$, M étant le nombre d'éléments de la base d'apprentissage.

Le problème d'optimisation est donc modifié et le Lagrangien devient :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^M \xi_i - \sum_{i=1}^M \alpha_i (\mathcal{Y}_i(w \cdot k(x_i) + b) - 1) \quad (\text{équ 3.38})$$

Où C est une constante strictement positive qui permet de régler le compromis entre le nombre d'erreurs de classification et la largeur de la marge. Cette constante est en général déterminée empiriquement par validation croisée sur l'ensemble d'apprentissage [Lor09].

III.2.5 La Décision

Pour estimer la différence entre deux images, il faut introduire une mesure de similarité. Il est important de noter que le système d'identification automatique de visage se base en sa totalité sur la méthode de localisation.

III.3 Conclusion

Dans ce chapitre, nous avons présenté un certain nombre des méthodes utilisées dans les cinq étapes d'un système de reconnaissance des expressions faciales. Nous nous sommes limité aux méthodes proches du modèle de notre approche.

Toutes ces méthodes ont des avantages et des inconvénients selon la complexité, le besoin de données d'apprentissage ou les paramètres à optimiser. Le choix d'une méthode va donc dépendre de l'application visée selon le compromis entre performance, simplicité, confort et quantité de données d'apprentissage disponible. Toutes ces méthodes seront implémentées puis comparées sur la base de données **CASIA** dans le Chapitre IV.

Chapitre IV

Implémentation et résultats

IV.1 Introduction

De toutes les fonctionnalités biométriques, la reconnaissance du visage reste l'un des sujets de recherche les plus actifs dans la reconnaissance des formes. Dans les dernières décennies, la plupart des travaux se concentrent sur la source de l'intensité 2D ou des images en couleur. Comme la précision de la reconnaissance du visage 2D est influencée par les variations de poses, expressions, des illuminations et subordonnés, il est encore difficile de développer un solide système automatique de reconnaissance faciale 2D. Pour toutes ces raisons nous nous sommes intéressé à l'identification des personnes en présence des expressions sur le visage pour renforcer la stabilité du système et dans le but d'améliorer les performances dans le l'application 3D expressions.

Jusqu'ici nous avons approché la reconnaissance automatique des expressions faciales d'un point de vue théorique. Comme l'identification est un problème pratique, nous allons maintenant examiner les différentes alternatives de solutions et l'appliquer à une base de données universelle. Et finalement discuter nos résultats et essayer de tirer quelques conclusions.

IV.2 La base de données CASIA (3D expressions)

IV.2.1 Description de la BDD CASIA 3D expression

Entre **Août 2004** et **Septembre 2004**, on a recueilli une base de données de visages **3D** constituée de **4624 scans** de **123 personnes** qui utilisent le sans contact numériseur **3D, Minolta Vivid 910**, comme indiqué dans **la Fig.IV.1**. Au cours de la construction de la base de données, on considère non seulement les variations individuelles de poses, les expressions et les illuminations, mais aussi les variations combinées des expressions sous illumination et poses sous des expressions, comme montré sur la **Fig.IV.2**, **Fig.IV.3** et **Fig.IV.4**. Pour les sujets avec des lunettes, on a recueilli 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. On vise à construire une base de données complète du visage en 3D, qui est en outre amenés à être une plate-forme publique pour tester les algorithmes de reconnaissance faciale 3D ou autres [**Che08**].



Fig IV.1 : Capture des scènes de la base de données CASIA.

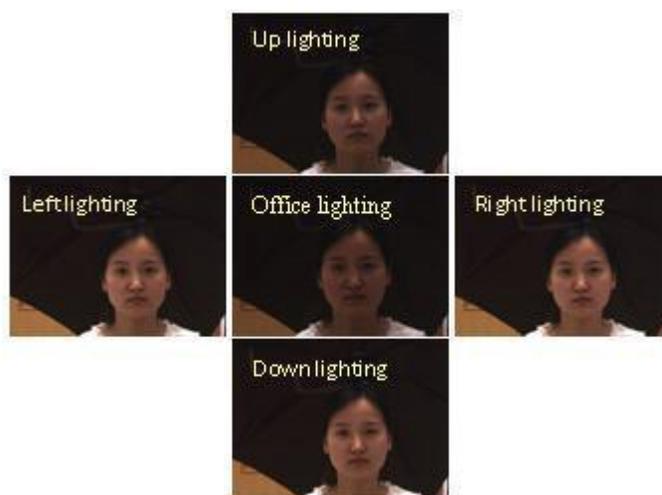


Fig IV.2: Variations d'éclairage de la base de données CASIA visage 3D.

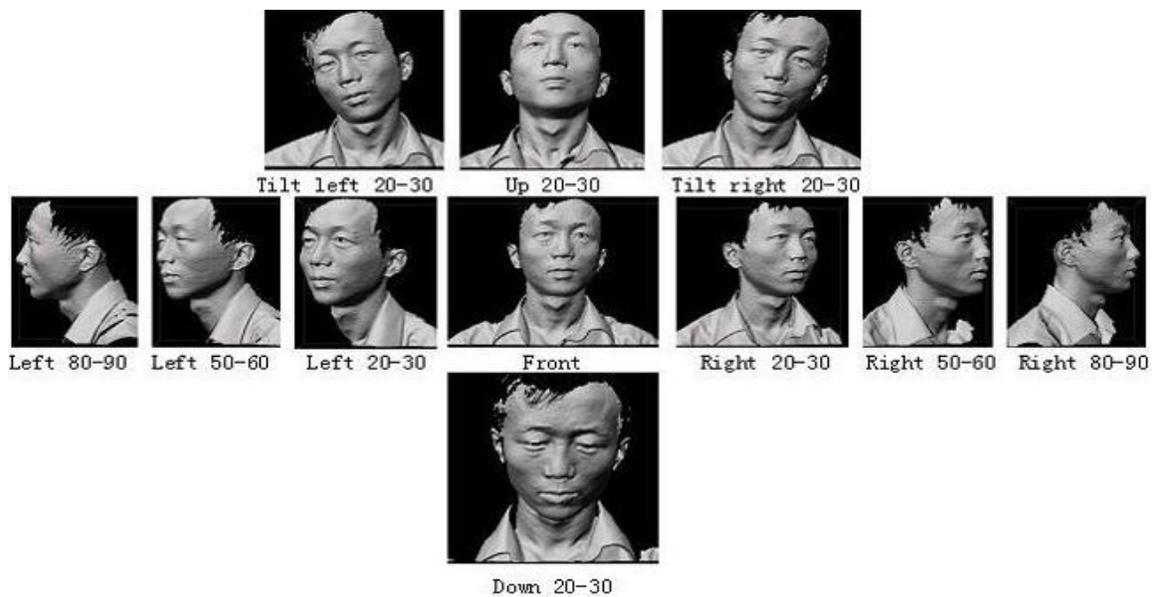


Fig IV.3: Pose des variations de la base de données CASIA visage 3D.

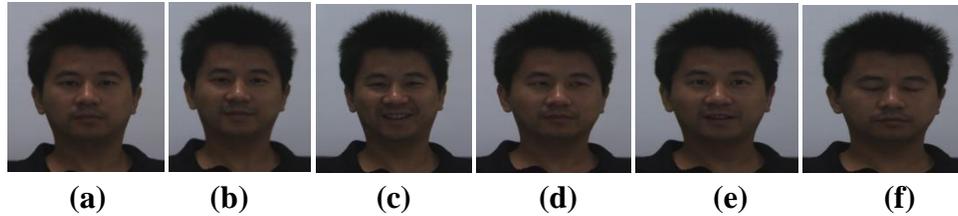
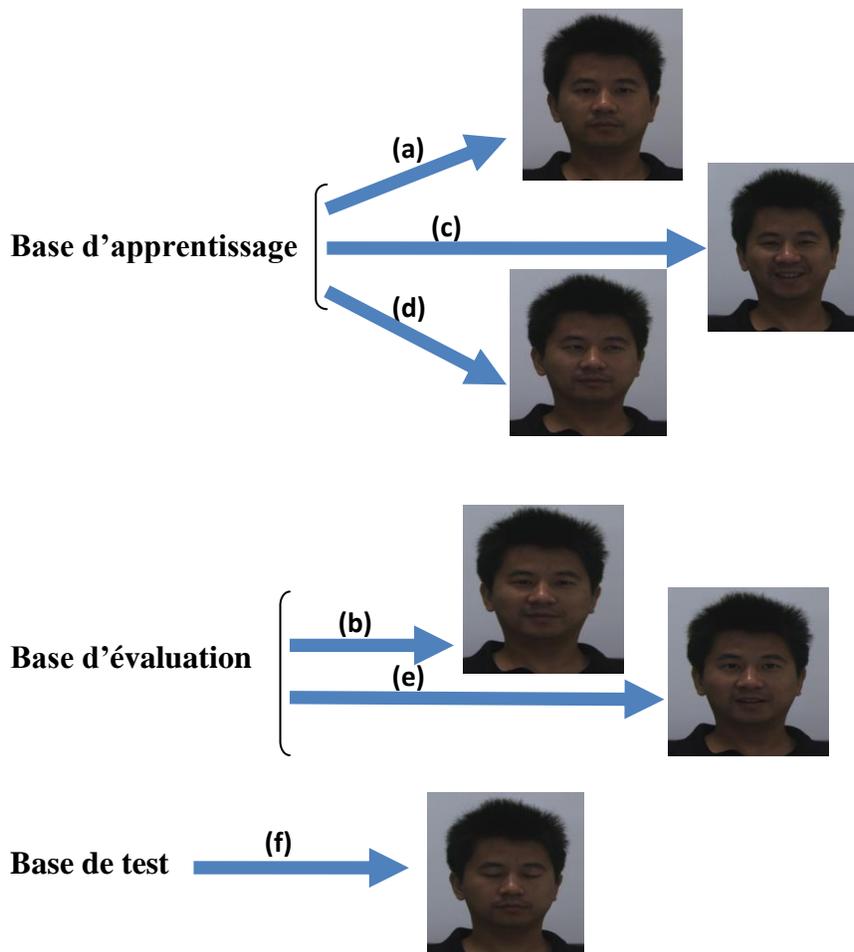


Fig IV.4: Variations d'expression de la base de données CASIA:(a)visage neutre, (b) sourire, (c) joie, (d) colère, (e) surprise, (f) prés des yeux.

IV.2.2 Protocole de la BDD CASIA 3D expressions

La BDD CASIA 3D expressions est partagée en trois sous ensembles selon le protocole suivants.

- *Apprentissage* de l'image (a),(c),(d) dans la base CASIA l'image 1,7,8.
- *Evaluation* de l'image (b),(e) dans la base CASIA l'image 6,9.
- *Test l'image* (f) dans la base CASIA l'image 10.



IV.3 Architecture structurelle du système de conception

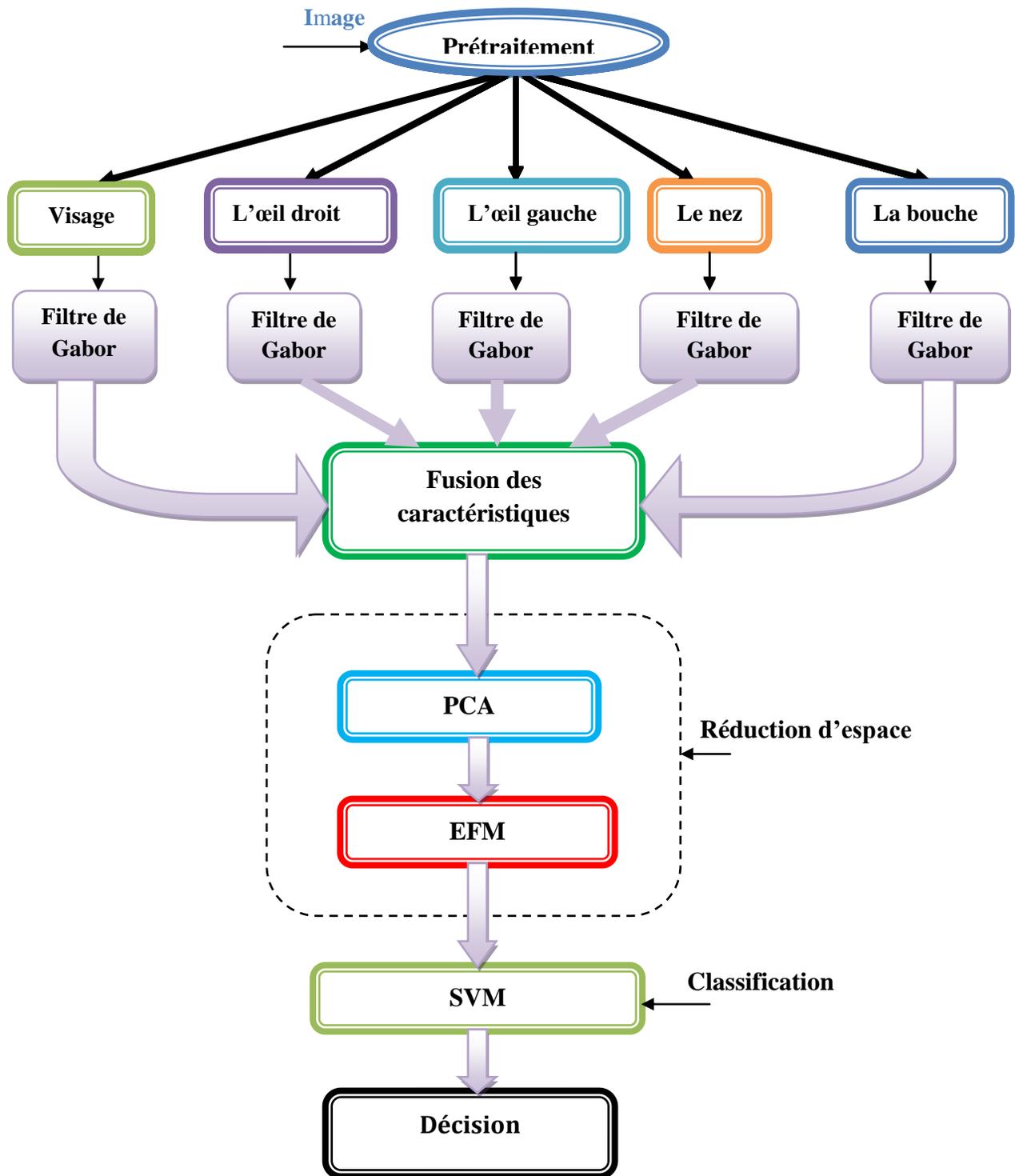


Fig IV.5 : Schéma de l'architecture structurelle du système de conception.

IV.3.1 Prétraitement d'image

Le prétraitement est une phase importante dans le processus globale d'identification, C'est une méthode simple qui augmente en général les performances du système. Elle permet souvent une première réduction des données et elle atténue les effets d'une différence de conditions lors des prises de vues. Pour notre travail nous avons besoin de trois phases de prétraitement : la décimation conversion de la couleur en l'espace de couleur HSV puis détection de visage.

IV.3.1.1 La composante V de l'espace de couleur TSV de l'image

Les systèmes d'identification de visages utilisent souvent la représentation en niveaux de gris de l'image de visage comme caractéristique d'entrée. Mais lorsque les images sont en couleurs, il est avantageux d'utiliser cette information supplémentaire pour améliorer les performances de ce système d'identification. Nous utilisons en général l'espace de couleur **TSV** (Teinte, Saturation, Valeur) ou **HSV** (Hue, Saturation, Value) ce sont les deux systèmes les plus utilisés dans la littérature [Mel09]. La quantification de la couleur de notre base de données **CASIA** est faite dans le système **RVB** (Rouge, Vert et Bleu), Le passage du système de couleurs **RVB** au système **TSV** se fait par une fonction de transfert non linéaire. La figure IV.6 présente les deux systèmes colorimétriques **RVB** et **TSV**.

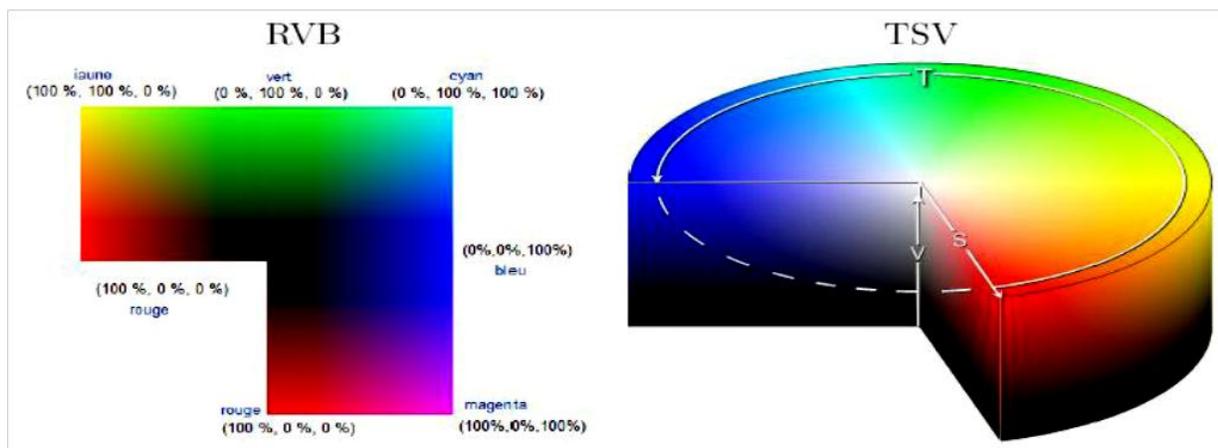


Fig IV.6: Différents systèmes colorimétriques : (a) plan du système de couleurs **RVB**, (b) cylindre du système de couleurs **TSV**.

Soient min et max , respectivement, les valeurs minimales et maximales des composantes R, V et B ($min = minimum(R, V, B)$ et $max = maximum(R, V, B)$) et soient x et y les coordonnées d'un pixel de l'image. Les transformations des espaces RVB vers TSV sont données par les équations suivantes.

$$T(x, y) = \begin{cases} \text{non défini, si } \min = \max \\ 60 * \frac{V-B}{\max-\min}, \text{ si } \max = R \text{ et } V \geq B \\ 60 * \frac{V-B}{\max-\min} + 360, \text{ si } \max = R \text{ et } V < B \\ 60 * \frac{V-B}{\max-\min} + 120, \text{ si } \max = V \\ 60 * \frac{V-B}{\max-\min} + 240, \text{ si } \max = B \end{cases} \quad (\text{Equ .IV.1})$$

$$S(x, y) \begin{cases} 0, \text{ si } \max = 0 \\ 1 - \frac{\min}{\max}, \text{ si non} \end{cases} \quad (\text{Equ .IV.2})$$

$$V(x, y) = \max \quad (\text{Equ .IV.3})$$

Pour plus de détails sur cette transformation, voir [Raf02].

Dans [Mel09] on a les espaces de couleurs TSV et on a trouvé que la composante V a donné le meilleur résultat, partant de ce résultat nous avons utilisé la composante V de l'espace de couleur TSV comme caractéristique de l'image.

La figure IV.7 présente la conversion de couleur et l'extraction de composante V de visage

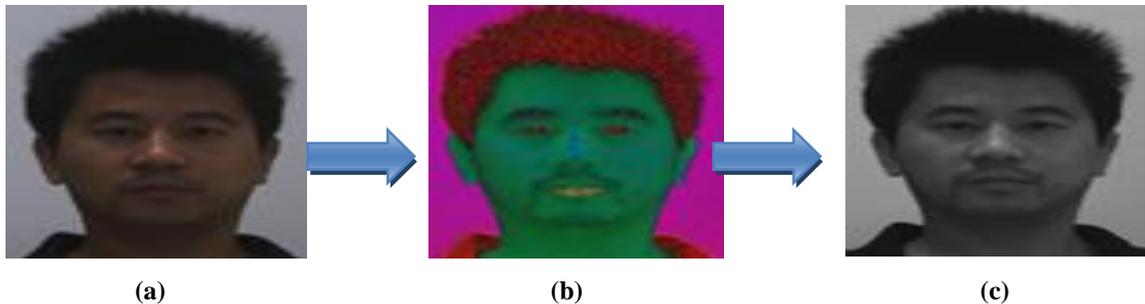


Fig IV.7 : (a) Image couleur, (b) Image en système TSV, (c) La composante V de TSV.

IV.3.1.2 Détection de visage et de points et régions d'intérêts

1) Détecteurs de visages et points d'intérêts

Avant d'entamer toute opération d'extraction de paramètres caractéristiques de notre image. Nous procédons tout d'abord à la détection de ce qui est significatif dans l'image du visage captée c'est-à-dire nous détectons le visage proprement dit. Car les cheveux, fond, la peau présentent peu d'intérêt pour l'analyse et l'identification. La détection du visage est la première phase de prétraitement et la figure nous montre le résultat de notre détection sur une image de visage neutre de la base CASIA 3D expression utilisée.

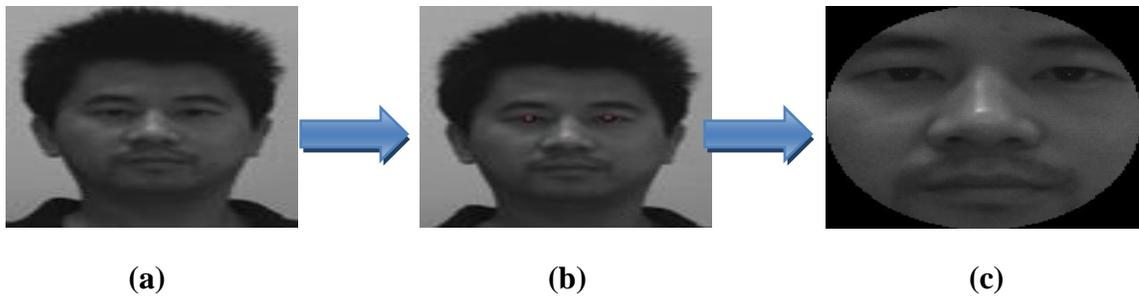


Fig IV.8: (a) La composante V de TSV,(b) détection manuelle d'œil,(c) détection de visage.

2) Détecteurs des régions d'intérêts

Les régions d'intérêts sont illustrées par la figure IV.9.

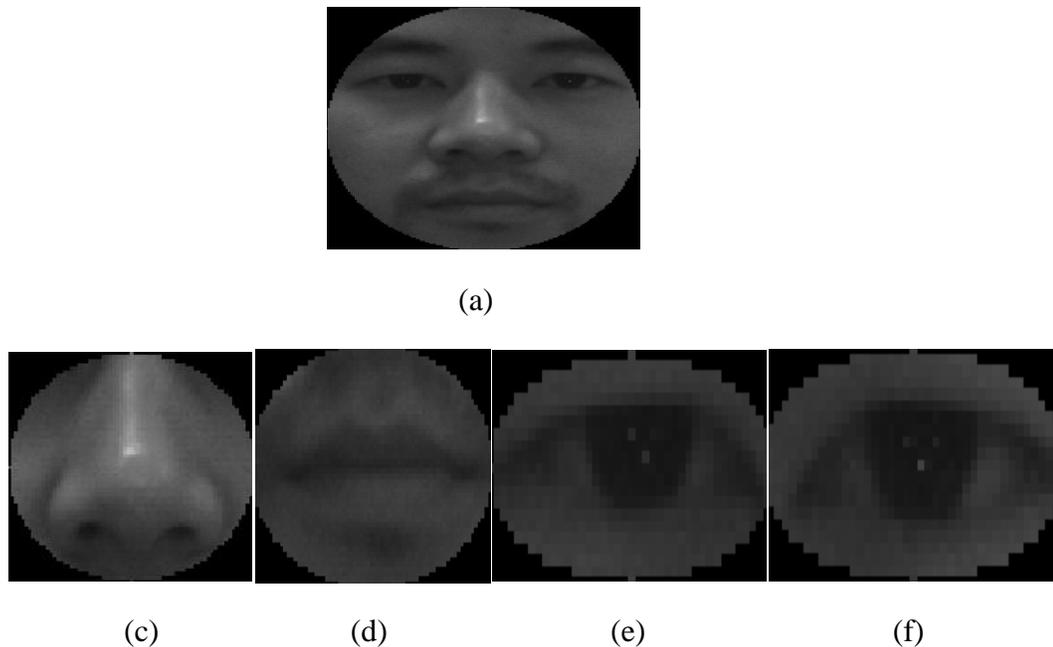


Fig IV.9 : (a) détection de visage,(c) détection de nez,(d) détection de bouche,
(e) détection d'œil droit,(f) détection d'œil gauche.

IV.3.2 Extraction des caractéristiques par filtre de Gabor

La méthode de reconnaissance que nous étudions dans cette partie repose principalement sur l'extraction des caractéristiques du visage par filtrage de Gabor 2D.

Il y a donc cinq paramètres qui contrôlent l'analyse par ondelettes, selon [Wal09,Moh08]. Notons que les étapes de cet algorithme sont les suivantes :

1) θ spécifie l'orientation du filtre. Nous utilisons dans ce cas huit orientations :

$$\theta = \{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}.$$

2) λ spécifie la longueur d'onde et donc la fréquence de la sinusoïde. Cette variable peut prendre cinq valeurs comprises entre 4 et 16 :

$$\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}.$$

3) ϕ spécifie la phase de la sinusoïde. Elle vaut 0 ou $\pi/2$ selon que l'on veut la partie réelle ou imaginaire.

4) σ spécifie la variance de la gaussienne (sa taille ou résolution). Elle est proportionnelle à la longueur d'onde de la sinusoïde. Dans notre cas $\sigma = \lambda$.

5) γ spécifie l'aspect de la gaussienne. Ici les gaussiennes sont de forme circulaire : $\gamma = 1$.

IV.3.3 Réduction d'espace

IV.3.3.1 L'analyse en composantes principales(PCA)

Après l'extraction des caractéristiques des visages par le filtre de Gabor

Nous tentons de réduire par un petit programme en Matlab calculant la **PCA (Analyse en Composante Principale)**. Les étapes de cet algorithme sont les suivants :

1) *Prétraitement.*

2) *Création de la matrice des données X (L'ordonnement des vecteurs images des clients coté à coté).*

3) *Centrer la matrice des données \bar{X}*

4) *Calcul de la matrice de covariance des données $X\bar{X}^T$.*

5) *Calcul des valeurs et vecteurs propres de la matrice de covariance.*

6) *La matrice de transformation de la PCA est les m premiers vecteurs propres (U_{PCA}).*

IV.3.3.2 Le Modèle Discriminant Linéaire Amélioré de Fisher (EFM)

- 1) *Prétraitement.*
- 2) *Calcul la matrice de dispersion intra-classes S_w (équ 3.16).*
- 3) *Calcul de la matrice de dispersion inter-classes S_B (équ 3.17).*
- 4) *Calcul des valeurs (λ) et vecteurs (V) propres de la matrice S_w .*
- 5) *Calculer la nouvelle matrice de dispersion inter-classe $K_b = A^{-1/2} V^T S_b V A^{-1/2}$*
- 6) *Calcul des valeurs (λ_b) et vecteurs (V_b) propres de la matrice K_b .*
- 7) *Calcul de la matrice $U = V A^{-1/2} V_b$.*
- 8) *La matrice de transformation de l'EFM est les m premiers vecteurs de U (U_{EFM}).*

Finalement nous réalisons notre base de données indexée dans la phase enrôlement selon la figure IV.10 .

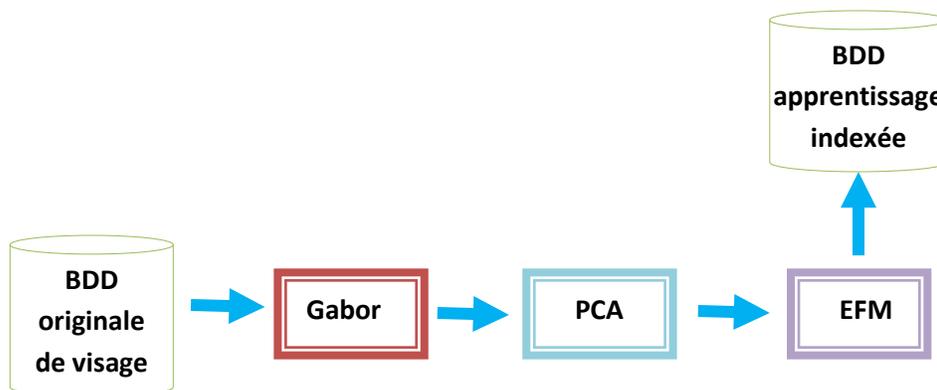


Fig IV.10 : Schéma de la phase enrôlement de notre modèle RV.

IV.3.4 Classification de scores avec les machines à vecteurs de support (SVM)

Les images originales des personnes de la base de données CASIA 3D expressions subissent la phase de prétraitement, puis des descripteurs sont extraits et ensuite projetés sur des espaces réduits et stockés dans une nouvelle base de données indexée. Celle-ci est répartie sous le protocole décrit précédemment. La base d'apprentissage nous servira dans le calcul de l'ACP et à l'entraînement de nos SVMs dans leur phase d'apprentissage. Une fois cette phase est effectuée

nous procédons à l'étape classification par SVM en utilisant la base d'évaluation et test de la BDD CASIA 3D expressions. Le schéma de ce processus est présenté dans la figure IV.11.

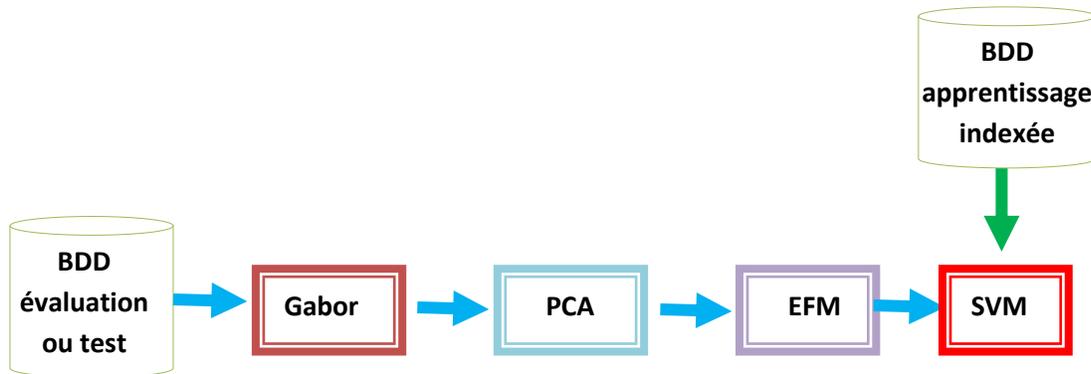


Fig IV.11: Schéma de principe de la classification par SVM.

Ainsi notre étude sur la classification des visages par SVM est menée sur deux phases :

- 1. phase d'apprentissage :** Pour notre protocole de la base de données CASIA 3D expressions, nous avons 123 images de personnes.

D'où : $N_c=123$ classe \Rightarrow 123 SVMs puisque le **Nombre de classes=nombre de SVMs**. Chaque personne représente 3 échantillons $\Rightarrow N_e=3$. Nous choisissons le numéro d'échantillons d'apprentissage égal 3. Alors $N_{e_{app}}=3$.

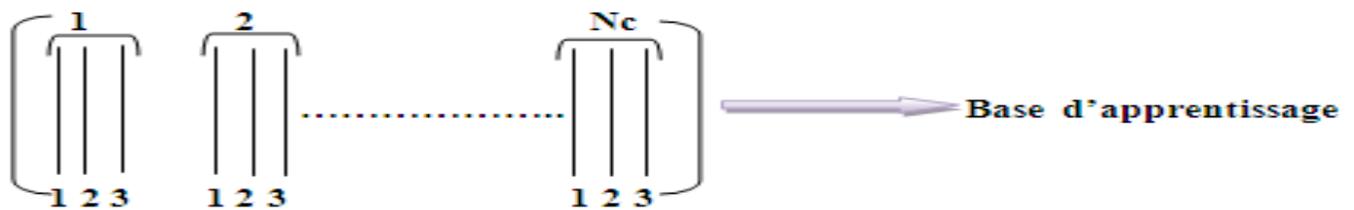


Fig IV.12: Schéma représentatif de la base d'apprentissage.

Alors que le schéma de la phase de test est comme suit :

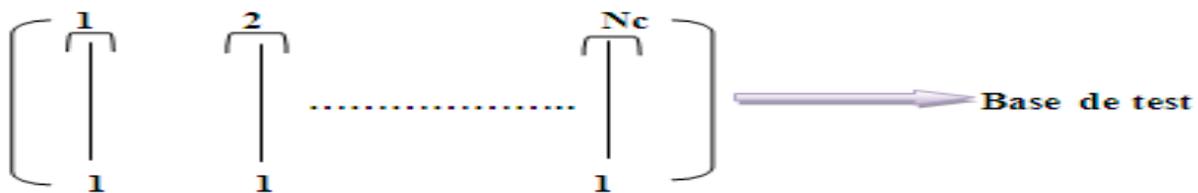


Fig IV.13 : Schéma représentatif de la base de test.

Donc la phase d'apprentissage par SVMs sera :

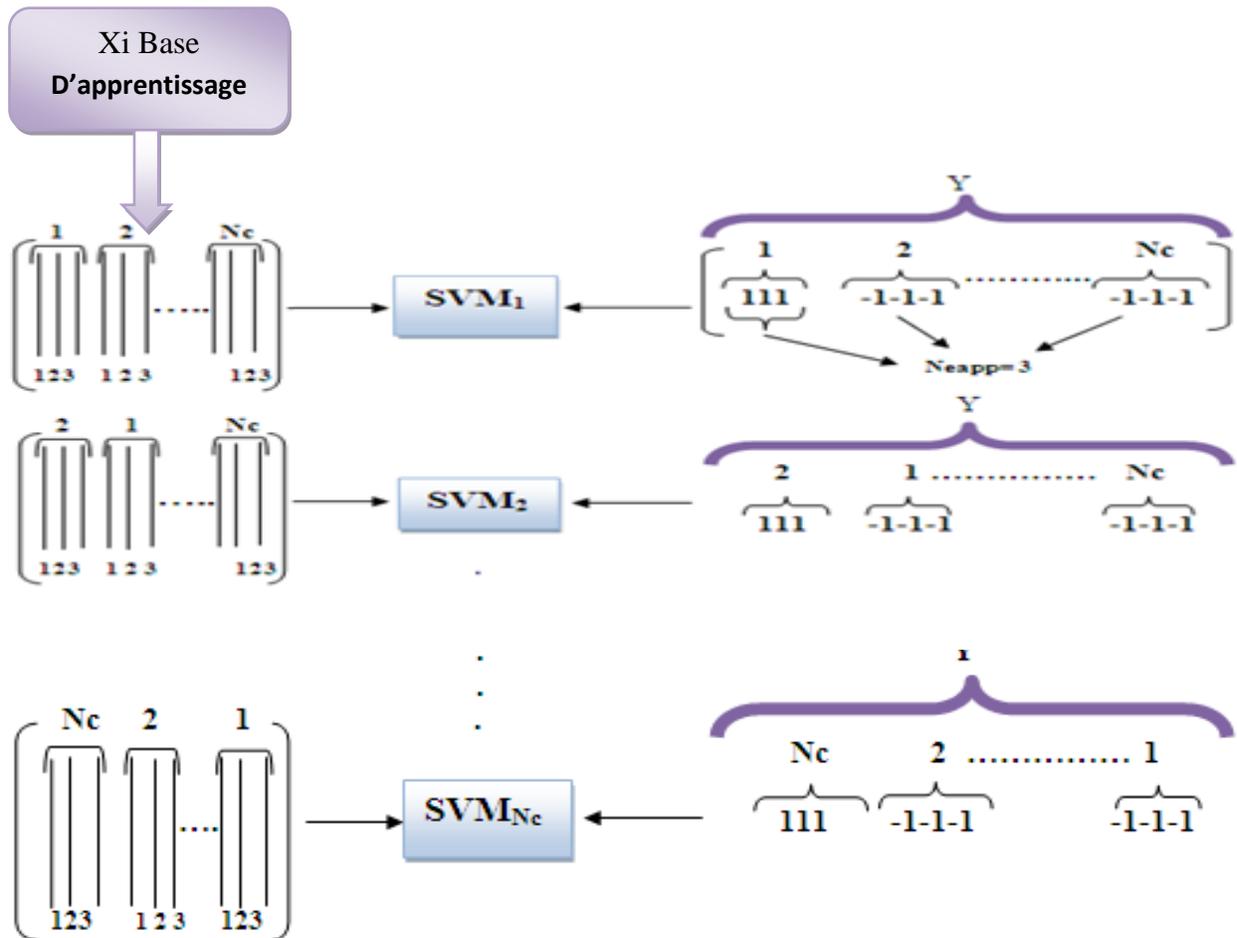


Fig IV.14 : Schéma de la phase d'apprentissage par SVMs.

D'après la figure **IV.14** nous constatons que 123 SVMs sont utilisée dans la phase apprentissage. Un SVM pour chaque classe (personne).

2. phase d'évaluation : Dans cette phase nous présentons à l'entrée des 123 SVMs le vecteur caractéristique réduit et nous sauvegardons l'indice de max des scores. Ainsi nous obtenons 2x123 scores/personnes d'où deux indices max/personne. Ces indices nous permettent de calculer notre taux de réussite TR.

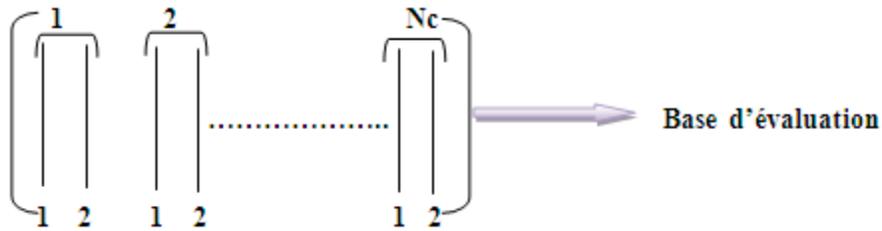


Fig IV.15 : Schéma représentatif de la base d'évaluation.

Le schéma de la phase d'évaluation est :

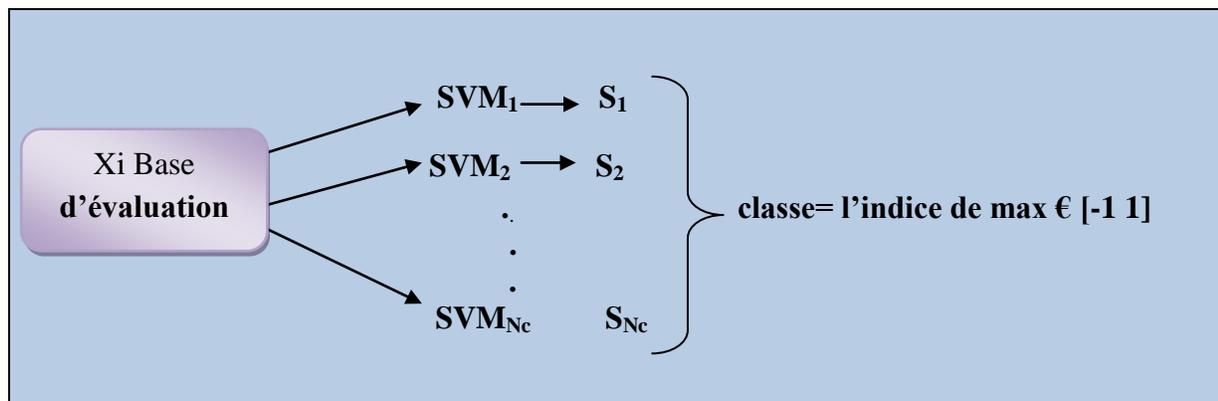


Fig IV.16 : Schéma de la phase d'évaluation par SVMs

IV.4 Résultats obtenus

Nous nous intéressons au calcul du taux de reconnaissance dans l'ensemble d'évaluation et l'ensemble de test pour $\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$ et $\theta = \{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$.

- **1^{er} expérience λ variable avec 8 orientations**

Dans les expériences menées nous fixons lambda (λ) et nous calculons le TR dans les 8 orientations pour divers nombre de paramètres caractéristiques puis, nous varions lambda et nous reprenons les mêmes expériences. Les meilleurs TR en fonction de N_p (Nombre de paramètres caractéristiques) ainsi obtenus sont présentés dans le tableau suivant.

| | | Nombre de caractéristique | | | | | | | | | | | |
|---------------------|------------------------|---------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| $\lambda=4$ | TR _{éva1} (%) | 84.14 | 90.24 | 93.08 | 92.68 | 93.08 | 93.49 | 93.90 | 94.30 | 94.71 | 94.71 | 94.71 | 94.71 |
| | TR _{tes1} (%) | 38.21 | 64.22 | 70.73 | 73.17 | 73.17 | 76.42 | 76.42 | 76.42 | 76.42 | 76.42 | 76.42 | 76.42 |
| $\lambda=4\sqrt{2}$ | TR _{éva2} (%) | 82.11 | 92.68 | 94.30 | 93.49 | 95.12 | 93.90 | 94.30 | 94.30 | 94.71 | 94.71 | 95.12 | 95.12 |
| | TR _{tes2} (%) | 48.78 | 62.60 | 68.29 | 69.91 | 70.73 | 72.35 | 73.98 | 73.98 | 74.79 | 75.60 | 76.42 | 75.60 |
| $\lambda=8$ | TR _{éva3} (%) | 83.33 | 92.27 | 94.30 | 95.12 | 95.52 | 95.52 | 95.52 | 95.12 | 94.71 | 95.12 | 95.12 | 95.12 |
| | TR _{tes3} (%) | 36.58 | 52.84 | 56.09 | 65.85 | 66.66 | 68.29 | 70.73 | 71.54 | 70.73 | 71.54 | 71.54 | 71.54 |
| $\lambda=8\sqrt{2}$ | TR _{éva4} (%) | 77.64 | 89.43 | 91.46 | 93.49 | 95.12 | 95.12 | 95.12 | 95.52 | 95.52 | 95.52 | 95.52 | 95.52 |
| | TR _{tes4} (%) | 39.02 | 51.21 | 60.16 | 67.47 | 65.85 | 69.91 | 73.17 | 71.54 | 72.35 | 73.17 | 72.35 | 72.35 |
| $\lambda=16$ | TR _{éva5} (%) | 80.48 | 88.21 | 92.68 | 92.68 | 92.68 | 93.08 | 93.08 | 93.49 | 93.49 | 93.49 | 93.90 | 93.90 |
| | TR _{tes5} (%) | 41.46 | 61.78 | 65.04 | 68.29 | 68.29 | 67.47 | 68.29 | 69.91 | 69.10 | 69.10 | 69.91 | 69.91 |

Tableau IV.1 : Valeurs de TR en fonction de Np pour 5 valeurs de λ et 8 orientations.

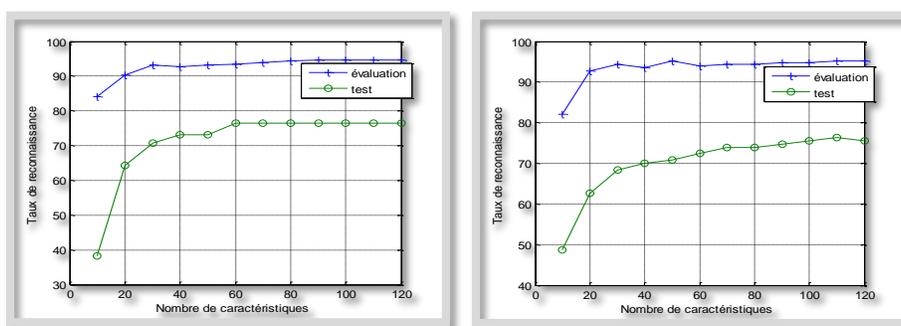


Fig IV.17: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=4$ et $\lambda=4\sqrt{2}$.

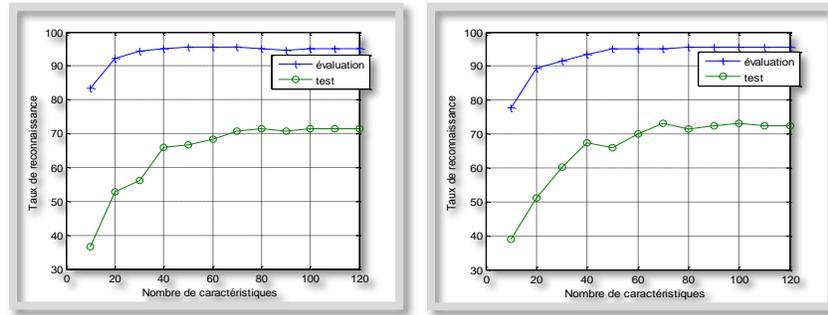


Fig IV.18: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=8$ et $\lambda=8\sqrt{2}$.

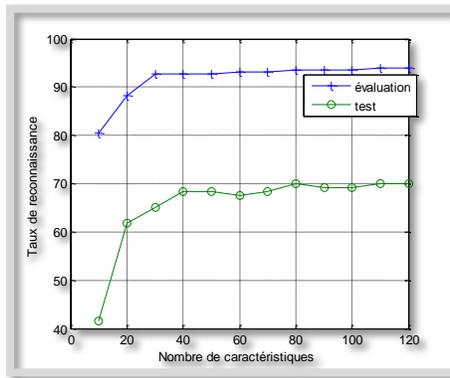


Fig IV.19: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=16$.

1) Comparaison des taux de reconnaissance dans l'ensemble d'évaluation pour 5 les résolutions et 8 orientations:

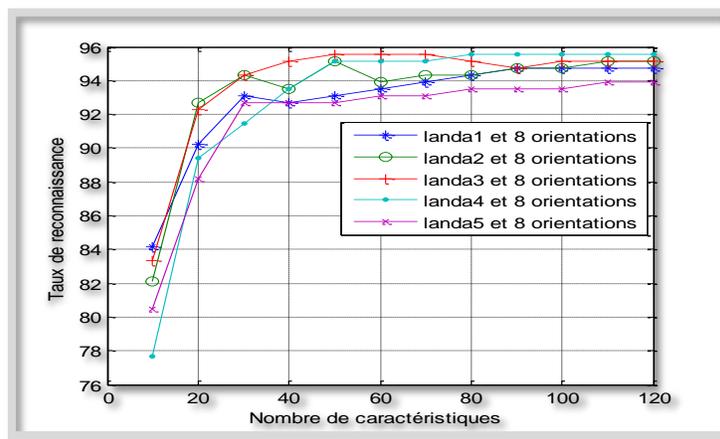


Fig IV.20 : Taux de reconnaissance de reconnaissance dans l'ensemble d'évaluation pour 5 résolutions et 8 orientations.

Nous remarquons que le meilleur TR dans cette base d'évaluation est obtenue pour $\lambda=8\sqrt{2}$ avec un TR = 95.52%.

2) Comparaison du taux de reconnaissance dans l'ensemble de test pour 5 résolutions et 8 orientations :

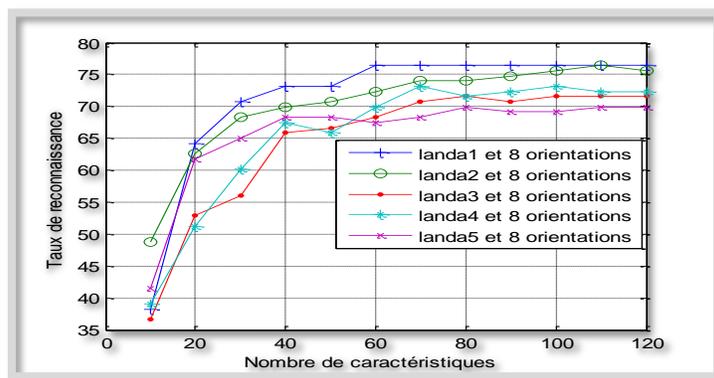


Fig IV.21 : Taux de reconnaissance dans l'ensemble de test pour 5 résolutions et 8 orientations.

Dans la base de test le meilleur taux de réussite est obtenu pour $\lambda=4$ et $\lambda=4\sqrt{2}$ avec un TR=76.42%.

Le résultat retenu dans cette première expérience

| Résolution (λ) | 4 | $4\sqrt{2}$ | 8 | $8\sqrt{2}$ | 16 |
|--------------------------|-------|-------------|-------|-------------|-------|
| TR % évaluation | 94.71 | 95.12 | 95.12 | 95.52 | 93.90 |
| TR% test | 76.42 | 76.42 | 76.42 | 73.17 | 69.91 |
| N p | 100 | 110 | 100 | 100 | 110 |

Tableau IV.2 : Tableau récapitulatif.

Dans ce tableau nous constatons que le meilleur obtenu est $TR_{\text{test}}= 76.42 \%$ avec résolution ($\lambda =4$) et 8 orientations (θ) et c'est un bon résultat concernant la reconnaissance de visages présentant des expressions et c'est un résultat qui reste à améliorer.

• 2^{ème} expérience $\lambda=4$ avec 8 orientations

Le meilleur résultat de TR obtenu sur la base de test est **TR=76.42 %** pour la première résolution $\lambda =4$. Pour cela, nous essayons dans cette deuxième expérience de fixer $\lambda=4$ et changer les orientations $\{\theta =0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$. en appliquant ceci sur les deux BDD test et évaluation.

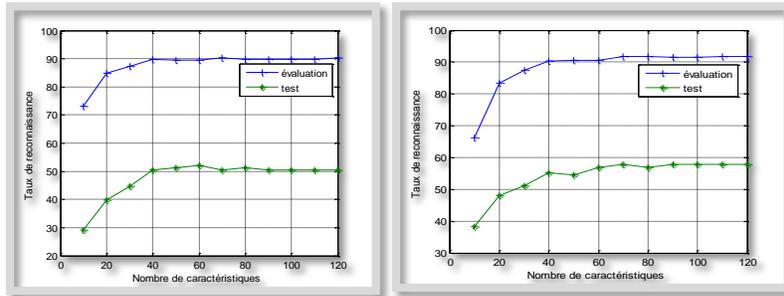


Fig IV.22: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=0$ et $\theta=\pi/8$.

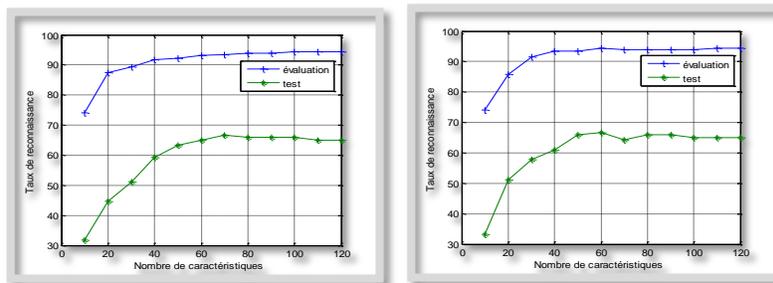


Fig IV.23: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=3\pi/8$ et $\theta=\pi/4$.

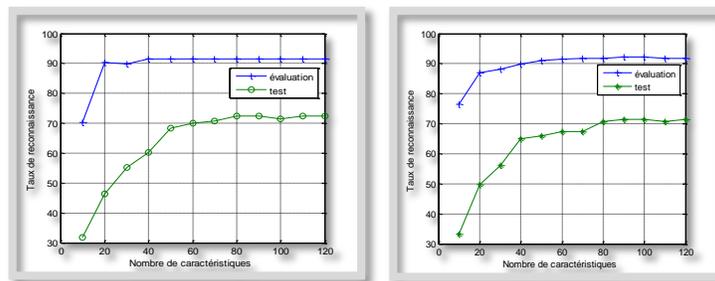


Fig IV.24 : Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=3\pi/8$ et $\theta= \pi/2$.

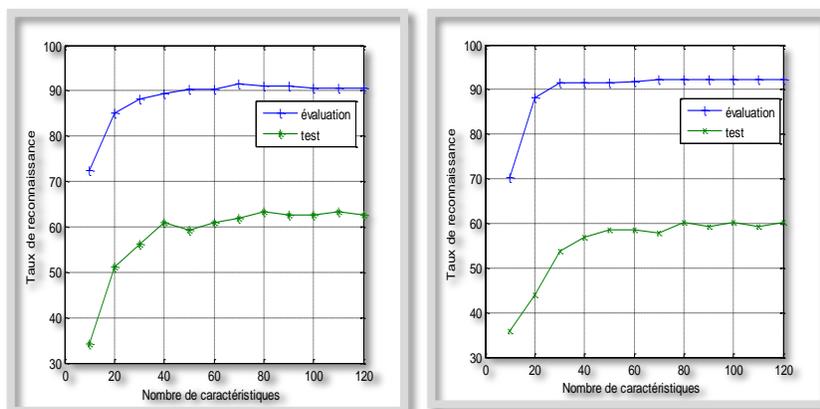


Fig IV.25: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour

$\lambda=4$ et $\theta=3\pi/4$ et $\theta= 7\pi/8$.

| | | Nombre de caractéristiques | | | | | | | | | | | |
|--------------------------------|------------------------|----------------------------|-------|-------|-------|-------|--------------|-------|--------------|--------------|--------------|-------|-------|
| | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| $\lambda=4$ et $\theta=0$ | TR _{éva1} (%) | 73.17 | 84.95 | 87.39 | 89.83 | 89.43 | 89.43 | 90.24 | 89.83 | 89.83 | 89.83 | 89.83 | 90.24 |
| | TR _{tes1} (%) | 29.26 | 39.83 | 44.71 | 50.40 | 51.21 | 52.03 | 50.40 | 51.21 | 50.40 | 50.40 | 50.40 | 50.40 |
| $\lambda=4$ et $\theta=\pi/8$ | TR _{éva2} (%) | 66.26 | 83.33 | 87.39 | 90.24 | 90.65 | 90.65 | 91.86 | 91.86 | 91.46 | 91.46 | 91.86 | 91.86 |
| | TR _{tes2} (%) | 38.21 | 47.96 | 51.21 | 55.28 | 54.47 | 56.91 | 57.72 | 56.91 | 57.72 | 57.72 | 57.72 | 57.72 |
| $\lambda=4$ et $\theta=\pi/4$ | TR _{éva3} (%) | 73.98 | 87.39 | 89.43 | 91.86 | 92.27 | 93.08 | 93.49 | 93.90 | 93.90 | 94.30 | 94.30 | 94.30 |
| | TR _{tes3} (%) | 31.70 | 44.71 | 51.21 | 59.34 | 63.41 | 65.04 | 66.66 | 65.85 | 65.85 | 65.85 | 65.04 | 65.04 |
| $\lambda=4$ et $\theta=3\pi/8$ | TR _{éva4} (%) | 73.98 | 85.77 | 91.46 | 93.49 | 93.49 | 94.30 | 93.90 | 93.90 | 93.90 | 93.90 | 94.30 | 94.30 |
| | TR _{tes4} (%) | 33.33 | 51.21 | 57.72 | 60.97 | 65.85 | 66.66 | 64.22 | 65.85 | 65.85 | 65.04 | 65.04 | 65.04 |
| $\lambda=4$ et | TR _{éva5} (%) | 70.32 | 90.24 | 89.83 | 91.46 | 91.46 | 91.46 | 91.46 | 91.46 | 91.46 | 91.46 | 91.46 | 91.46 |

| | | | | | | | | | | | | | |
|------------------------------------|------------------|-------|-------|-------|-------|-------|-------|-------|--------------|--------------|-------|-------|-------|
| $\theta = \pi/2$ | $TR_{tes5} (\%)$ | 31.70 | 46.34 | 55.28 | 60.61 | 68.29 | 69.91 | 70.73 | 72.35 | 72.35 | 71.54 | 72.35 | 72.35 |
| $\lambda = 4$ et $\theta = 5\pi/8$ | $TR_{éva6} (\%)$ | 76.42 | 86.99 | 88.21 | 89.83 | 91.05 | 91.46 | 91.86 | 91.86 | 92.27 | 92.27 | 91.86 | 91.86 |
| | $TR_{tes6} (\%)$ | 33.33 | 49.59 | 56.09 | 65.04 | 65.85 | 67.47 | 67.47 | 70.73 | 71.54 | 71.54 | 70.73 | 71.54 |
| $\lambda = 4$ et $\theta = 3\pi/4$ | $TR_{éva7} (\%)$ | 72.35 | 84.95 | 88.21 | 89.43 | 90.24 | 90.24 | 91.46 | 91.05 | 91.05 | 90.65 | 90.65 | 90.65 |
| | $TR_{tes7} (\%)$ | 34.14 | 51.21 | 56.09 | 60.97 | 59.34 | 60.97 | 61.78 | 63.41 | 62.60 | 62.60 | 63.41 | 62.60 |
| $\lambda = 4$ et $\theta = 3\pi/4$ | $TR_{éva8} (\%)$ | 70.32 | 88.21 | 91.46 | 91.46 | 91.46 | 91.86 | 92.27 | 92.27 | 92.27 | 92.27 | 92.27 | 92.27 |
| | $TR_{tes8} (\%)$ | 35.77 | 43.90 | 53.65 | 56.91 | 58.53 | 58.53 | 57.72 | 60.16 | 59.34 | 60.16 | 59.34 | 60.16 |

Tableau IV.3 : Valeurs de TR en fonction de Np pour 5 valeurs de $\lambda = 4$ avec θ orientation.

1) Comparaison du taux de reconnaissance dans l'ensemble d'évaluation pour $\lambda = 4$ dans les 8 orientations:

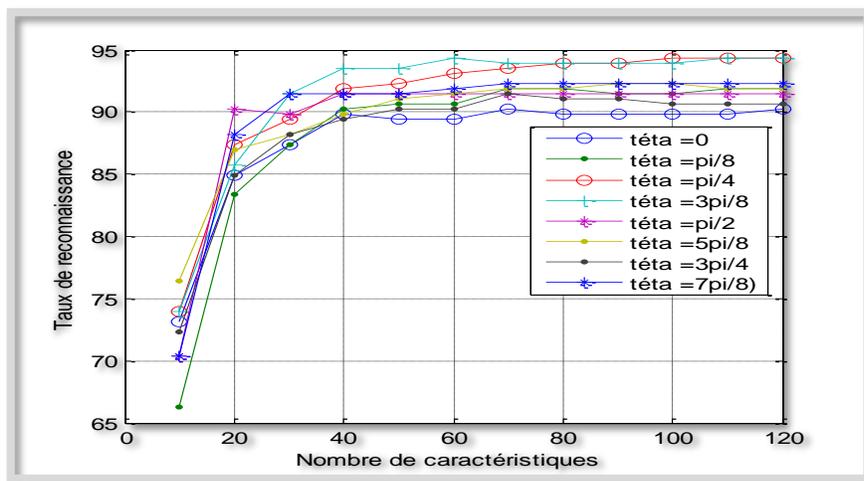


Fig IV.26 : Comparaison du taux de reconnaissance dans l'ensemble d'évaluation pour $\lambda = 4$ dans les 8 orientations.

Nous pouvons dire que le meilleur résultat est obtenu pour les orientations $\theta = \frac{3\pi}{8}$ et $\theta = \pi/4$.

2) Comparaison du taux de reconnaissance dans l'ensemble de **test** pour $\lambda = 4$ dans les 8 orientations:

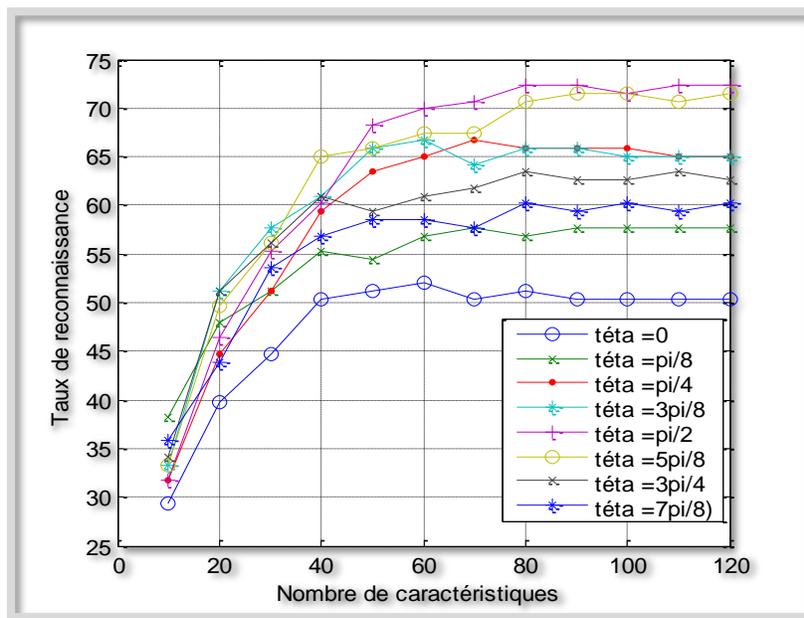


Fig IV.27 : Comparaisant taux de reconnaissance dans l'ensemble de test pour $\lambda = 4$ dans 8 orientations.

Dans ce cas l'orientation $\theta = \pi/2$ et $5\pi/8$ s'avère la plus favorable.

| Orientations des filtres (θ) | 0 | $\pi/8$ | $\pi/4$ | $3\pi/8$ | $\pi/2$ | $5\pi/8$ | $3\pi/4$ | $7\pi/8$ |
|---------------------------------------|-------|---------|--------------|--------------|--------------|--------------|----------|----------|
| TR % évaluation | 89.43 | 91.46 | 94.30 | 94.30 | 91.46 | 92.27 | 91.05 | 92.27 |
| TR% test | 52.03 | 57.72 | 65.85 | 66.66 | 72.35 | 71.54 | 63.41 | 60.16 |
| Nombre de caractéristiques | 60 | 90 | 100 | 60 | 80 | 90 | 80 | 80 |

Tableau IV.4: Résultats du système d'identification pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test.

3) Fusion par concaténation des filtres de Gabor associés aux orientations optimales :

D'après les résultats obtenus précédemment nous retenons quatre orientations $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$, $\lambda = 4$ et pour diverses N_p . Ces descripteurs ainsi obtenus sont concaténés et de nouveaux TR calculés en fonction de N_p .

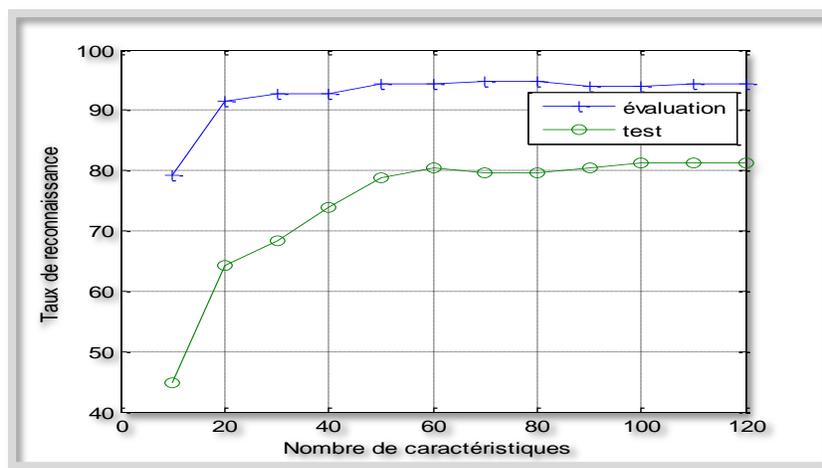


Fig IV.28 : Valeurs de TR pour $\lambda = 4$ et $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$ dans l'ensemble d'évaluation et test.

Les résultats de cette nouvelle expériences sont représentés dans le tableau suivant :

| | Nombre de caractéristiques | | | | | | | | | | | |
|-----------------|----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|-------|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| TR % évaluation | 79.26 | 91.46 | 92.68 | 92.68 | 94.30 | 94.30 | 94.71 | 94.71 | 93.90 | 93.90 | 94.30 | 94.30 |
| TR% test | 44.71 | 64.22 | 68.29 | 73.98 | 78.86 | 80.48 | 79.67 | 79.67 | 80.48 | 81.30 | 81.30 | 81.30 |

Tableau IV.5 : Valeurs de TR pour $\lambda = 4$ et $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$ dans l'ensemble d'évaluation et test.

Le meilleur résultat est obtenu avec un $\text{TR}_{\text{test}} = 81.30 \%$ avec un $\text{TR}_{\text{éval}} = 94.30 \%$. Nous pouvons affirmer que c'est un bon résultat relativement à ce que l'on espère avoir dans les applications 3D expressions.

IV.5 Conclusion

Dans ce chapitre nous avons présenté les différentes étapes de l'implémentation de notre modèle. Nous avons effectué plusieurs expériences en se basant sur les variations de l'orientation et de la résolution du filtre de Gabor d'une part et du nombre de paramètres de la PCA d'autre part. Car nous jugeons que la phase de filtrage est fort importante dans le système de reconnaissance de visages. La phase de classification est assurée par des SVMs dont l'apprentissage est primordial pour cela nous avons établi un protocole pour notre BDD CASIA 3D et assuré une bonne phase d'entraînement. Nos expériences ainsi menées nous conduisent à des résultats encourageants surtout que le domaine de l'identification des visages en présence d'expressions est très délicate d'autant plus que la base CASIA 3D est une collecte de visages Asiatiques où la ressemblance est très accentuée. D'après les résultats obtenus et présentés dans ce chapitre nous remarquons que le meilleur TR dans la base d'évaluation est obtenu pour $\lambda=8\sqrt{2}$ avec un TR = 95.52%. Dans la base de test le meilleur taux de réussite est obtenu pour $\lambda=4$ et $\lambda=4\sqrt{2}$ avec un TR=76.42%. Les orientations présentant les meilleurs TR sont $\theta=3\pi/8$ et $\theta = \pi/4$ dans l'ensemble d'évaluation et $\theta= \pi/2$ et $5\pi/8$ dans l'ensemble de test (pour $\lambda =4$). Une fois l'étude faite sur les différentes orientations et résolutions du filtre un choix est établi. Les meilleurs filtres sont concaténés et ce nouveau descripteur est soumis à l'étape de réduction d'espace. Et finalement, le meilleur résultat dans ce dernier cas est obtenu avec un $\mathbf{TR}_{\text{test}} = 81.30$ % avec un $\mathbf{TR}_{\text{éval}} = 94.30$ % . Nous pouvons affirmer que c'est un bon résultat relativement à ce que l'on espère avoir dans les applications 3D expressions.

Conclusion

Générale

Conclusion Générale

Le but visé de concevoir un système d'identification d'identité qui soit facile et peu coûteux dans l'implémentation et qui utilise une modalité biométrique particulière.

Nous avons présenté des méthodes. Nous avons également montré l'intérêt de ces représentations pour la reconnaissance d'expressions faciales. Le pourcentage global de bonne reconnaissance se situe aux alentours de 81.30% pour l'ensemble de test.

Un état de l'art des différentes approches récentes dans le domaine est établi. Les technologies utilisées dans les systèmes biométriques pour l'identification de personnes sont étudiées et par le biais de cette étude et de l'état de l'art un modèle est conçu et puis implémenté. Cette étude nous a permis de constater que la reconnaissance de visage suscite de plus en plus l'intérêt de la communauté scientifique, car elle présente plusieurs challenges et verrous technologiques. Nous avons mis en évidence les différentes difficultés inhérentes à la reconnaissance automatique de visages, notamment l'invariance à l'illumination. Enfin, nous avons aussi donné un aperçu sur les techniques de mesure de performances. Les techniques utilisées dans la reconnaissance des expressions faciale sont détaillées dans le chapitre III. Ce chapitre nous a permis de tirer un certain nombre d'enseignements du processus de reconnaissance d'expressions faciales. Après le développement des principales difficultés inhérentes à la reconnaissance d'expressions faciales, nous avons présenté les principales phases du système de reconnaissance d'expressions faciales. Un certain nombre des méthodes utilisées dans les cinq étapes d'un système de reconnaissance des expressions faciales sont présentées. Nous nous sommes limités aux méthodes proches du modèle de notre approche. Toutes ces méthodes ont des avantages et des inconvénients selon la complexité, le besoin de données d'apprentissage ou les paramètres à optimiser. Le choix d'une méthode va donc dépendre de l'application visée selon le compromis entre performance, simplicité, confort et quantité de données d'apprentissage disponible. Toutes ces méthodes sont implémentées puis comparées sur la base de données **CASIA** dans le Chapitre IV. Dans ce chapitre nous avons présenté les résultats obtenus en effectuant une étude des différentes étapes de notre système, des expériences sont menées et les résultats sont

encourageants. Nous avons présenté les différentes étapes de l'implémentation de notre modèle. Nous avons aussi effectué plusieurs expériences en se basant sur les variations de l'orientation et de la résolution du filtre de Gabor d'une part et du nombre de paramètres de la PCA d'autre part. Car nous jugeons que la phase de filtrage est fort importante dans le système de reconnaissance de visages. La phase de classification est assurée par des SVMs dont l'apprentissage est primordial pour cela nous avons établi un protocole pour notre BDD CASIA 3D et assuré une bonne phase d'entraînement. Nos expériences ainsi menées nous conduisent à des résultats encourageants surtout que le domaine de l'identification des visages en présence d'expressions est très délicate d'autant plus que la base CASIA 3D est une collecte de visages Asiatiques où la ressemblance est très accentuée. D'après les résultats obtenus et présentés dans ce chapitre nous remarquons que le meilleur TR dans la base d'évaluation est obtenu pour $\lambda=8\sqrt{2}$ avec un TR = 95.52%. Dans la base de test le meilleur taux de réussite est obtenu pour $\lambda=4$ et $\lambda=4\sqrt{2}$ avec un TR=76.42%. Les orientations présentant les meilleurs TR sont $\theta=3\pi/8$ et $\theta = \pi/4$ dans l'ensemble d'évaluation et $\theta= \pi/2$ et $5\pi/8$ dans l'ensemble de test (pour $\lambda =4$). Une fois l'étude faite sur les différentes orientations et résolutions du filtre un choix est établi. Les meilleurs filtres sont concaténés et ce nouveau descripteur est soumis à l'étape de réduction d'espace. Et finalement, le meilleur résultat dans ce dernier cas est obtenu avec un $TR_{\text{test}} = 81.30 \%$ avec un $TR_{\text{éval}} = 94.30 \%$. Nous pouvons affirmer que c'est un bon résultat relativement à ce que l'on espère avoir dans les applications 3D expressions.

Ce travail ainsi présenté peut être amélioré par :

- l'amélioration de méthodes de discrimination,
- une bonne détection,
- une fusion multi algorithmes ou multi biométrie,
- et aussi une meilleure classification.

Nous conseillons aussi l'application de cette approche sur d'autres bases de données universelles.

Dédicace

A mes chers parents qui ont toujours été là pour moi. Je vous suis reconnaissante pour les sacrifices consentis pour mon bien être. Votre soutien et votre éducation ont fait de moi ce que je suis aujourd'hui, puissé-je vous rendre fiers et heureux.

A ma grand-mère.

A mes sœurs et, mes frères et à toute ma famille pour leurs soutiens inconditionnels.

A mon oncle ouamane foudhile.

A mes amies qui m'ont apporté soutien et consolation dans les moments de besoin.

A walid qui partageait ma vie.

ET A ma mère, ma mère et ma mère pour son amour.

Je dédie ce mémoire.

Bibliographie :

- [Adi97] Y. Adini, Y. Moses, S. Ullman, Face recognition: The problem of compensating for changes in illumination direction. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 721–732, 1997.
- [Adj07] Adjout Mohamed et Benaissa Abdelhak « Fusion de la DCT-PCA et la DCT-LDA appliquée à la reconnaissance de visages » Institut National de formation en Informatique (I.N.I)Oued-smar Alger 2007.
- [Aho04] T. Ahonen, A. Hadid, and M. Pietikainen. *Face Recognition with Local Binary Patterns*. 2004.
- [Ami12] Amit Konar, Aruna Chakraborty, Anisha Halder, Rajshree Mandal Ramadoss J.anarthanan- Type-2 Intervalle modèle flou pour Emotion Reconnaissance de Visage expressif ; 2012.
- [Ano09] Anour Mellak h .Reconnaissance des visages en conditions dégradées . THÈSE pour l'obtention du grade de Docteur de l'Institut National des Télécommunications Spécialité : Traitement d'image et de signal .soutenue 07 Avril 2009
- [Arc06] S. Arca, P. Campadelli, and R. Lanza. "A Face Recognition System Based On Automatically Determined Facial Fiducial Points". *Pattern Recognition*, Vol. 39, No. 3, pp. 432–443, 2006.
- [Bai09] Y. Bai, L. Guo, L. Jin, and Q. Huang. A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition. In *International Conference on Image Processing*, 2009.
- [Bar03] M. S. Bartlett, G. Littlewort, I. Fasel, J. R. Movellan, Real Time Face Detection and Facial Expression Re-cognition: Development and Applications to Human Computer Interaction, IEEE workshop on Computer Vision and Pattern Recognition for Human Computer Interaction (in conjunction with IEEE Intl. Conf. On CVPR) , Madison, U.S.A., June, 2003
- [Ben08] Bengherabi M., Mezai L., Harizi F., Cheriet M., Guessoum A, «Face recognition based on 2DPCA, DIAPCA and DIA2DPCA in DCT domain », *The 5th International Multi-Conference on Systems, Signals and Devices, IEEE SSD*, Amman, Jordan, July 20-23, 2008.

- [**Bla97**] M. J. Black and Y. Yacoob, Recognizing Facial Ex-pressions in Image Sequences Using Local Parametri- zed Models of Image Motion, *International Journal of Computer Vision*, Vol. 25, Number 1, pp. 23–48, 1997.
- [**Bos07**] A. Bosch, A. Zisserman, and X. Munoz. Representing shape witha spatial pyramid kernel. In 6th ACM International Conference onImage and Video Retrieval, pages 401–408, 2007.
- [**Bre97**] C. Bregler, M. Covel, and M. Slaney, Video Rewrite:Driving Visual Speech with Audio, *Siggraph proceedings*, pp. 353–360, 1997.
- [**Bro01**] G. Brown, S. Yamada, and T. Sejnowski. “Independent Component Analysis at the Neural Cocktail Party”. *Trends in Neuroscience*, Vol. 24, pp. 54–63, 2001.
- [**Cap04**] Caplier, Hammal, Z A. : Eye and Eyebrow Parametric Models for AutomaticSegmentation. *IEEE SSIAT*, Lake Tahoe, Nevada (2004)
- [**Cha09**] Anis CHAARI "*Nouvelle approche d'identification dans les bases de données biométriques basée sur une classification non supervisée*", 2009 au laboratoire IBISC.
- [**Che95**] D.T. Chen, A. State, and D. Banks, Interactive Shape Metamorphosis, *Siggraph proceedings* , pp. 43–44,1995.
- [**Che08**] Cheng Zhong, Zhenan Sun, Tieniu Tan, Learning Efficient Codes for 3D Face recognition,international conference on image processing(ICIP) 2008.
- [**Chu02**] X. HOU, S. LI, H. ZHANG, Q. CHENG, Direct appearance models,In *Intl. Conference on Computer Vision and Pattern Reocognition*,p. 828–833, 2001.
- [**Coh98**] J. Cohn, A. Zlochow, and J. J. Lien and T. Kanade,Feature-point Tracking by Optical Flow Discriminates Subtle Differences in Facial Expression, *Proceedings of the 3rd IEEE International Conference on Automa-tic Face and Gesture Recognition* , pp. 396–401, 1998.
- [**Coo01**] T. Cootes, G. Edwards, and C. Taylor. “Active Appearance Models”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 681–685, June 2001.
- [**Dja06**]D◊jamelSaigaa, Contribution à l'authentificationd'individus par reconnaissance de visages, thèse de Doctorat d'état en automatique, Université Mohamed khiderBiskraAlgérie, Novembre 2006.
- [**Dha11**] A. Dhall, A. Asthana, R. Goecke, and T. Gedeon. Emotion recognitionusing PHOG and LPQ features. In *IEEE Automatic Face andGesture Recognition*

Conference FG2011, Workshop on Facial Expression Recognition and Analysis Challenge FERA, 2011

[Don99] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski. Classifying facial actions. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 21 :974–989, 1999.

[Dra00] B. Draper, K. Baek, M.S. Bartlett and R. Beveridge, Recognizing Faces with PCA and ICA, *Computer Vision and Image Understanding* .

[Dra03] B. DRAPER, K. BAEK, M.S. BARTLETT, R. BEVERIDGE, Recognizing Faces with PCA and ICA, *Computer Vision and Image Understanding*, 91(1/2), p. 115–137, July 2003

[Edw98] G.J. Edwards, T.F. Cootes, and C.J. Taylor, Face Recognition Using Active Appearance Models, *Proceedings of the European Conference of Computer Vision*, pp. 581–695, 1998.

[Ekm99] P. Ekman, Facial Expressions, *The Handbook of Cognition and Emotion*, T. Dalgleish and M. Power, John Wiley & Sons Ltd. 1999.

[Ess97] I. Essa and A. Pentland, Coding, Analysis Interpretation Recognition of Facial Expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence* , Vol. 19, Number 7, pp. 757–763, 1997.

[Eva04] Y. Tian. Evaluation of face resolution for expression analysis. In *Computer Vision and Pattern Recognition Workshop*, 2004.

[Ezz02] T. Ezzat, G. Geiger and T. Poggio, Trainable Videorealistic Speech Animation, *ACM Transactions on Graphics (TOG)*, Vol. 21, Number 3, pp. 388–398, 2002.

[Fra04] Franck Davoine, Bouchra Abboud et Van Môm Dang-Analyse de visages et d'expressions faciales par modèle actif d'apparence Université de Technologie de Compiègne, France 9 février 2004.

[Fuk91] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, second edition, 1991.

[Gok01] S.B. GOKTURK, J.-Y. BOUGUET, R. GRZESZCZUK, A data-driven model for monocular face tracking, In *International Conference on Computer Vision*, Vancouver, Canada, July 2001.

[Gre01] H. Greenspan, J. Goldberger, I. Eshet. Mixture model for face color modeling and segmentation, *Pattern Recognition Lett.* 22 (14) 1525–1536, 2001.

- [**Gue11**] Gueddouh Soumia «*L'apprentissage des SVMs incrémentales*» Université Mohamed Khider Biskra, Alger 2011.
- [**Guo00**] G. Guo, S. Li, and K. Chan. "Face Recognition by Support Vector Machines". In : *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 196–201, 2000.
- [**Ham06**] Hammal.Z – Segmentation des traits du visage, analyse et reconnaissance d'expressions faciales par le Modèle de Croyance Transférable. PhD Thesis in Cognitive Science, Université de Joseph Fourier de Grenoble Juin 2006.
- [**Han06**] Hong-Viet LUONG- Reconnaissance multimodale de gestes de communication non verbale, Institut National Polytechnique de Grenoble Laboratoire des Images et des Signaux Novembre 2006.
- [**Hay05**] S. Haykin and Z. Chen. "The Cocktail Party Problem". *Neural Computation*, Vol. 17, No. 9, pp. 1875–1902, 2005.
- [**Hei01**] B. Heisele, P. Ho and T. Poggio, Face Recognition with Support Vector Machines: Global Versus Component-based Approach, , International Conference on Computer Vision (ICCV'01) , Vancouver, Canada, Vol. 2, pp. 688–694, 2001.
- [**Hei08**] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. In International conference on Image and Signal Processing, 2008.
- [**Hon98**] H. Hong, H. Neven and C. von der Malsburg, On-line Facial Expression Recognition based on Personalized Gallery, Intl. Conference on Automatic Face and Gesture Recognition, IEEE Comp. Soc, pp. 354–359, 1998.
- [**Jia11**] B. Jiang, M. F. Valstar, and M. Pantic. Action unit detection using sparse appearance descriptors in space-time video volumes. In IEEE Conference on Automatic Face and Gesture Recognition, 2011.
- [**Kan00**] T. Kanade, J.F. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In Fourth IEEE International Conference on Automatic face and Gesture Recognition (FG'00), pages 46–53, 2000.
- [**Kan02**] H. Kang, T.F. Cootes and C.J. Taylor, Face Expression Detection and Synthesis using Statistical Models of Appearance, Measuring Behavior, pp. 126–128, 2002.
- [**Kan73**] T. Kanade. "Picture Processing System by Computer Complex and Recognition of Human Faces". In : *Doctoral dissertation, Kyoto University*, November 1973.

- [**Kot08**] I. Kotsia, S. Zafeiriou, and I. Pitas. Texture and shape information fusion for facial expression and facial action unit recognition. *Pattern Recognition*, 41 :833–851, 2008.
- [**Lap07**] LAPIERRE Lauréline-Construction d'une échelle d'évaluation des capacités de reconnaissance des expressions faciales émotionnelles, Université Paul Sabatier Juin 2007.
- [**Lor09**] Lorène ALLANO La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles Thèse présentée pour l'obtention du grade de Docteur de *L'institut National Des Télécommunications* Soutenue le 12 Janvier 2009.
- [**Lou10**] LOUIBA Fadia et HADJ ALI Ryma "Système de contrôle d'accès physique basé sur le visage" .pp.19 ; 2010.
- [**Lia06**] S. Liao, W. Fan, A. C. S. Chung, and D. Yeung. Facial expression recognition using advanced local binary patterns, tsallis entropies and global appearance features. In *IEEE International Conference on Image Processing*, 2006.
- [**Lig11**] Ligang Zhang, Dian Tjondronegoro, and Vinod Chandran- *Géométrie vs Apparence pour discriminer entre les émotions qui se posent et spontanée*. Queensland University of Technology Australia 2011.
- [**Lio99**] M. J. Lyons, J. Budynek, and S. Akamatsu. Automatic classification of single facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 :1357–1362, 1999.
- [**Lit05**] G. Littlewort, M. S. Bartlett, I. Fasel, J. Susskind, and J. Movellan. Dynamics of facial expression extracted automatically from video. *Image and Vision Computing*, 24 :615–625, 2006
- [**Lyo99**] M.J. Lyons and J. Budynek and S. Akamatsu, Automatic Classification of Single Facial Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, pp. 1357–1362, 1999.
- [**Mor09**] Nicolas MORIZET. "Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris". 2009 .
- [**Mik99**] S. Mika, G. Ratsch, J. Weston, B. Schölkopf, and K.-R. Müller. "Fisher Discriminant Analysis With Kernels". In : *Neural Networks for Signal Processing IX*, pp. 41–48, 1999.
- [**Min00**] T. P. Minka. Automatic choice of dimensionality of pca. Technical Report TR 514, MIT Media Laboratory Perceptual computing section, 2000.

- [**Moh05**] 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.
- [**Moh08**] MohamadHoseynSigari, MahmoodFathy. Best Wavelength Selection for Gabor Wavelet using GA in EBGM Algorithm. Proceedings of the International MultiConference of Engineers and Computer Scientists Hong Kong, pp: 19-21 March, 2008.
- [**Mar98**] A.M Martínez, R. Banavente. The AR face database. Tech. Report 24 CVC Barcelone, Espagne, June 1998.
- [**Oja96**] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distribution. *Pattern Recognition*, 29 :51–59, 1996.
- [**Oua11**] Ouamane Abdelmalik « Etude de la fusion de modalités pour l'authentification en biométrie » Université Mohamed Khider Biskra, Alger 2011.
- [**Pad98**] C. Padgett, G. Cottrell and R. Adolphs, Categorical Perception in Facial Emotion Classification, *Siggraph proceedings*, pp. 75–84 1998.
- [**Pat05**] M.F. Valstar, I. Patras, and M. Pantic. Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data. In *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pages 76–84, 2005.
- [**Pen96**] P. Penev and J. Atick, *Local Feature Analysis: A general Statistical Theory for Object Representation*, 1996.
- [**Per05**] V. Perlibakas. "Face recognition using Principal Component Analysis and Log-Gabor Filters". March 2005.
- [**Pet09**] Dijana Petrovska-Delacrétaz, Gérard Chollet, Bernadette Dorizzi « *Guide to Biometric Reference Systems and Performance Evaluation* »; ISBN: 978-1-84800-291-3; 2009.
- [**Phi03**] P. J. Phillips, P. J. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone. Face recognition vendor test 2002: Evaluation report. NISTIR 6965. Available online at <http://www.frvt.org>, 2003.
- [**Pig96**] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, D.H. Salesin, Synthesizing Realistic Facial Expressions from Photographs, *Proceedings of Cognitive Science Conference*, 1996.

- [**Pla81**] S. M. Platt and N. I. Badler. Animating facial expressions. In Proceedings of the 8th annual conference on Computer graphics and interactive techniques, pages 245–252, 1981.
- [**Rad05**] M. Pantic, M. F. Valstar, R. Rademaker, and L. Maat. Web-based database for facial expression analysis. In IEEE International Conference on Multimedia and Expo, 2005.
- [**Raf02**] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing (2nd Edition). Prentice Hall, January 2002.
- [**Riz12**] Rizwan Ahmed Khan, Alexandre Meyer, Hubert Konik, Saida Bouakaz - *Une méthode de reconnaissance des expressions du visage basée sur la perception*. Université de Lyon, France. Université Jean Monnet, Laboratoire Hubert Curien 2012.
- [**Sou08**] Souhila Guerfi Ababsa - *Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D* - Docteur de l'Université Evry Val d'Essonne Spécialité-03 octobre 2008.
- [**Sag01**] B. Sagi, S. Nemat-Nasser, R. Kerr, R. Hayek, D. C.C., and R. Hecht-Nielsen. "A Biologically Motivated Solution to the Cocktail Party Problem". *Neural Computation*, Vol. 13, No. 7, pp. 1575–1602, 2001.
- [**Tin05**] Y. Tian, T. Kanade, and J. F. Cohn. Handbook of Face Recognition. Springer, 2005 (Chapter 11. Facial Expression Analysis).
- [**Vap95**] Vladimir N. Vapnik. The nature of statistical learning theory. Springer Verlag New York, Inc., New York, NY, USA, 1995.
- [**Vis04**] M. Visani, C. Garcia et C. Laurent. Comparing Robustness of Two-Dimensional PCA and Eigenfaces for Face Recognition. Dans *Proc. of the Int. Conf. on Image Analysis and Recognition ICIAR'04*, Springer LNCS 3211, A. Campilho, M. Kamel (eds), pages 717–724, Porto, Portugal, Octobre 2004.
- [**Wal09**] Walid Hizem . Capteur Intelligent pour la Reconnaissance de Visage. Thèse Doctorat délivré conjointement par L'Institut National des Télécommunications et l'Université Pierre et Marie Curie - Paris 6 Spécialité : Electronique/Informatique , Soutenue le 2009.
- [**Xio08**] Xiaoxing Li, Greg Mori and Hao Zhang - Expression-invariants de reconnaissance de visage avec de classification Expression- Simon Fraser University, Canada
- [**Bla03**] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *Pattern Analysis and Machine Intelligence*, IEEE Transactions, 25(9):1–11, Sept. 2003.

[**Yan02**] M. H. Yang, Face Recognition Using Kernel Methods, *Advances in Neural Information Processing Systems 14 (NIPS 14)* , pp. 215–220, 2002.

[**Zuo05**] Zuo W., Wang K., Zhang D., « Assembled Matrix Distance Metric for 2DPCA-based Face and Palmprint Recognition», *Proceeding of 4th International Conference on Machine Learning and Cybernetics*, Guangzhou, 18-21 August 2005.

[**Zha07**] G. Zhao and M. Pietikäinen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 29 :915–928, 2007.

Remerciements

Je remercie tout d'abord « Allah » le Tout Puissant, de m'avoir donné le courage et la patience, et de m'avoir guidé là où je suis arrivée.

Je remercie particulièrement Mme : BELAHCENE BENATIA Mébarka qui a été patiente avec nous du début jusqu'à la fin, sans hésiter à nous venir en aide à tout moment. Son sérieux, sa générosité et sa sympathie nous ont particulièrement touchés.

Je remercie aussi Mr OUAMANE Abdelmalik pour ses encouragements et pour son aide qui m'était précieuse.

Nous tenons aussi à remercier les membres du jury pour leur temps accordé à l'étude de notre mémoire.

Mes remerciements et ma gratitude vont aux professeurs et enseignants ainsi qu'à tous les étudiants et au personnel côtoyés tout au long de notre cursus universitaire.

Je remercie également tous nos amis qui nous ont soutenus.

Je remercie tous ceux qui ont contribué de près ou de loin et par une quelconque forme de contribution pour la réalisation de ce projet.

Liste des tableaux

Tableau I.1 : Avantages et inconvénients des technologies biométriques.

Tableau II.1 : Tableau de définition de distances.

Tableau II.2 : États symboliques associés à chaque expression.

Tableau II.3 : Comparaison des performances des algorithmes proposés sur les bases de visages.

Tableau IV.1 : Valeurs de TR en fonction de N_p pour 5 valeurs de λ et 8 orientations.

Tableau IV.2 : Tableau récapitulatif.

Tableau IV.3 : Valeurs de TR en fonction de N_p pour 5 valeurs de $\lambda=4$ avec θ orientation.

Tableau IV.4: Résultats du système d'identification pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test.

Tableau IV.5 : Valeurs de TR pour $\lambda =4$ et $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$ dans l'ensemble d'évaluation et test.

Liste des figures

Fig I.1 : Les principales caractéristiques biométriques pouvant servir comme moyen d'identification

Fig I.2 : Revenus de la vente de technologies biométriques de 2009 à 2014 selon IBG

Fig I.3 : Diagrammes des processus d'enrôlement, de vérification et d'identification

Fig I.4: Scores de compatibilité pour différentes technologies biométriques dans un système MRTD.

Fig I.5: Principe de fonctionnement de base d'un système de reconnaissance faciale.

Fig I.6 : Classification des algorithmes principaux utilisés en reconnaissance faciale.

Fig I.7 : Exemple de variation d'éclairage.

Fig I.8 : Exemples de variation de poses.

Fig I.9 : Exemples de variation d'expressions.

Fig I.10 : La distribution du taux de vraisemblances.

Fig 1.11: La courbe ROC.

Fig II.1 : Mouvements faciaux globaux.

Fig II.2: Relation entre expression faciale et émotion.

Fig II.3: Méthodes d'analyse automatique d'un système de reconnaissance d'expressions faciales.

Fig II.4 : Les étapes dans le processus de reconnaissance des expressions faciales.

Fig II.5 : Extraction des contours et définition des 5 distances.

Fig II.6 : Masque pour le visage.

Fig II.7 : Masques de montage. (a), (c): manuellement marqueurs placés (points blancs), (b), (d): équipés masques.

Fig II.8 : Cadre du système d'évaluation.

Fig II.9 : (a) 68 points de référence pour la formation et (b) les résultats de détection ASM .

Fig II.10 : Caractéristiques Faciales.

Fig III.1 : Les étapes dans le processus de reconnaissance des expressions faciales.

Fig III.2 : Orientation de l'ondelette.

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

Fig III.4 : Comparaison entre les projections de deux classes de points ("class 1" et "class 2") sur les axes principaux construits par ACP et par LDA.

Fig III.5 : Principe de la technique SVM (hyperplan optimal, vecteurs de supports, marge maximale).

Fig IV.1 : Capture des scènes de la base de données CASIA .

Fig IV.2: Variations d'éclairage de la base de données CASIA visage 3D.

Fig IV.3: Pose des variations de la base de données CASIA visage 3D.

Fig IV.4: Variations d'expression de la base de données CASIA:(a)visage neutre, (b) sourire, (c) joie, (d) colère, (e) surprise, (f) prés des yeux.

Fig IV.5 : Schéma de l'architecture structurelle du système de conception.

Fig IV.6: Différents systèmes colorimétriques : (a) plan du système de couleurs **RVB**, (b) cylindre du système de couleurs **TSV**.

Fig IV.7 : (a) Image couler, (b) Image en système TSV, (c) La composons V de TSV.

Fig IV.8: (a) La composante V de TSV,(b) détection manuelle d'œil,(c) détection de visage

Fig IV.9 : (a) détection de visage,(c) détection de nez,(d) détection de bouche,(e) détection d'œil droit,(f) détection d'œil gauche.

Fig IV.10 : Classification de

Fig IV.11:

Fig IV.12: Schéma représentatif de la base d'apprentissage.

Fig IV.13 : Schéma représentatif de la base de test.

Fig IV.14 : Schéma de la phase d'apprentissage par SVMs.

Fig IV.15 : Schéma représentatif de la base d'évaluation.

Fig IV.16 : Schéma de la phase d'évaluation par SVMs

Fig IV.17: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=4$ et $\lambda=4\sqrt{2}$.

Fig IV.18: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=8$ et $\lambda=8\sqrt{2}$.

Fig IV.19: Valeurs de TR dans l'ensemble d'évaluation et de test pour les 8 orientations $\lambda=16$.

Fig IV.20 : Taux de reconnaissance de reconnaissance dans l'ensemble d'évaluation pour 5 résolutions et 8 orientations.

Fig IV.21 : Taux de reconnaissance dans l'ensemble de test pour 5 résolutions et 8 orientations.

Fig IV.22: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=0$ et $\theta=\pi/8$.

Fig IV.23: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=\pi/4$ et $\theta=3\pi/8$.

Fig IV.24 : Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=\pi/2$ et $\theta=5\pi/8$.

Fig IV.25: Taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda=4$ et $\theta=3\pi/4$ et $\theta=7\pi/8$.

Fig IV.26 : Comparaison du taux de reconnaissance dans l'ensemble d'évaluation pour $\lambda=4$ dans les 8 orientations.

Fig IV.27 : Comparaison du taux de reconnaissance dans l'ensemble de test pour $\lambda=4$ dans les 8 orientations

Fig IV.28 Valeurs de TR pour $\lambda = 4$ et $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$. dans l'ensemble d'évaluation et test.

Liste des abréviations

ACP: Analyse en Composantes Principales.

BDD: Base De Données

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

CK: Cohn-Kana

DLDA: LDA directe.

EFM: Enhanced Fisher Model.

FAP : Paramètres D'animation Faciale.

GT2FS : General Type-2 Fuzzy sets.

IBG: International Biometric Group.

IT2FS: Interval Type-2 Fuzzy Face-Space.

JAFFE: Japanese Female Facial Expression.

LBP: Local Binary Pattern.

LDA : Linear Discriminant Analysis.

LE: Laplacian Eigenmaps.

mRMS : minimal Redundancy Maximal relevance.

NVIE : Naturel D'expression Visible Et Infrarouge Du Visage.

RF: Random Forest.

TR: Taux de Reconnaissance.

SVM: Machine Vecteur de Support.

SIFT : Scale-Invariant Feature Transform.

Résumé

Notre objectif est l'étude des techniques d'identification de visages en présence d'expression dans un premier temps. Puis la mise en œuvre d'un système de reconnaissance en choisissant celui qui nous présente les meilleures performances au sens biométrique. Un état de l'art des approches récentes dans le domaine est établi. Ce qui nous a guidé dans notre choix des techniques utilisées dans la conception de notre système de reconnaissance d'expression faciale. Notre approche est basée sur un prétraitement qui nous assure la préparation de l'image pour la soumettre à l'analyse. Cette dernière est basée sur une étape d'extraction de paramètres qui sont les régions d'intérêts du visage qui sont fusionnées. Puis projetées sur un espace réduit réalisé par Le modèle discriminant Linéaire amélioré de Fisher l'EFM. Ensuite ces descripteurs sont soumis à un classifieur connu pour son efficacité c'est le SVM. Et finalement nous avons tenté la concaténation des meilleurs filtres pour améliorer le taux de réussite. Les résultats obtenus sont encourageants et prometteurs. Nos résultats montrent une précision de taux de reconnaissance automatique des cinq expressions faciales de 81.30% sur la base de données d'expression faciale CASIA 3D.

Mots clés : Expressions Faciales, PCA, EFM, SVM, CASIA 3D.

ملخص

هدفنا هو استكشاف تقنيات للتعرف على الوجوه في وجود تعبير في البداية. ثم تنفيذ نظام التعرف من خلال اختيار واحد وأفضل في إطار التحقق من الهوية. و لهذا الصدد يتم استخدام المناهج الحديثة في المجال البيومتري. تسترشد ما لنا في اختيارنا من التقنيات المستخدمة في تصميم نظامنا من التعرف على تعبير الوجه. ويستند نهجنا على المعالجة التي تضمن التحضير لتقديم صورة للتحليل. ويستند هذا الأخير على خطوة من استخراج المعلومات في المناطق المهمة من الوجه التي تم دمجها. من المتوقع بعد ذلك على مساحة صغيرة أنشأتها نموذج التمايز الخطي EFM في التحسين. ثم تقدم هذه المواصفات إلى المصنف المعروف بفعاليتيه و هو SVM. و في الأخير حاولنا بالتسلسل المنطقي للوصول إلى أفضل المرشحات لتحسين معدل النجاح. النتائج التي تم الحصول عليها مشجعة وواعدة. نتائجنّا تظهر معدل دقة من الاعتراف التلقائي من التعابير الخمس للوجه بدقة 81.30% على قاعدة البيانات CASIA D3 للوجه.

Sommaire

Sommaire

| | |
|--|---------------------------------|
| INTRODUCTION GÉNÉRALE | 01 |
| Chapitre I | Reconnaissance de visage |
| I.1 Introduction | 03 |
| I.2 La reconnaissance biométrique..... | 04 |
| I.2.1 Les différentes techniques biométriques..... | 05 |
| I.2.2 Le marché mondial de la biométrie..... | 06 |
| I.3 Le mode de reconnaissance | 07 |
| I.3.1 Le mode de vérification ou d'authentification..... | 07 |
| I.3.2 Le mode d'identification | 07 |
| I.4 Pourquoi choisir le visage ?..... | 08 |
| I.5 Reconnaissance du visage | 09 |
| I.5.1 La détection du visage | 10 |
| I.5.2 L'extraction des caractéristiques | 11 |
| I.5.3 La comparaison des caractéristiques (classification)..... | 11 |
| I.6 Les méthodes de reconnaissance faciales..... | 11 |
| I.6.1 Méthodes globales | 12 |
| I.6.2 Méthodes locales..... | 14 |
| I.6.3 Méthodes hybrides..... | 15 |
| I.7 Principales difficultés de la reconnaissance de visage..... | 16 |
| I.7.1 Changement d'illumination..... | 16 |
| I.7.2 Variation de pose..... | 16 |
| I.7.3 Expressions faciales..... | 17 |
| I.7.4 Présence ou absence des composants structurels..... | 17 |
| I.7.5 Occultations partielles..... | 17 |
| I.8 Performances d'un système de reconnaissances de visage..... | 18 |
| I.8.1 Performance d'un système d'identification..... | 18 |
| I.8.2 Performance d'un système de vérification..... | 18 |
| I.9 Nouvelle technologie améliorant la reconnaissance faciale..... | 20 |

| | |
|---|----|
| I.9.1 L'utilisation de la technologie infrarouge..... | 20 |
| I.9.2 L'utilisation de la technologie 3D..... | 20 |
| I.9.3 Avantage de la technologie 3D | 20 |
| I.10 Conclusion..... | 21 |

Chapitre II Etat de l'art de reconnaissance d'expression faciale

| | |
|---|----|
| II.1 Introduction | 22 |
| II.2 Qu'est-ce qu'une expression faciale?..... | 22 |
| II.3 Le développement émotionnel..... | 23 |
| II.3.1 Comment étudie-t-on les expressions émotionnelles chez l'humain ?..... | 23 |
| II.4 Système de reconnaissance d'expression faciale existant..... | 25 |
| II.4.1 Méthode basée sur la segmentation..... | 25 |
| II.4.1.1 Segmentation des traits du visage..... | 26 |
| II.4.1.2 Extraction de données caractéristiques..... | 26 |
| II.4.1.3 Classification par le Modèle de Croyance Transférable..... | 27 |
| II.4.2 Méthode par masque pour le visage..... | 27 |
| II.4.3 Méthode par le modèle d'apparence (AAM)..... | 28 |
| II.4.4 Autres méthodes pour la reconnaissance d'expressions faciales..... | 29 |
| II.5 État de l'art des techniques récentes de reconnaissance d'expression faciale.... | 33 |
| II.6 Tableau comparatif des travaux récents sur la REF..... | 36 |
| II.7 Conclusion..... | 37 |

Chapitre III Etude du modèle de reconnaissance de visage avec expression

| | |
|--|----|
| III.1 Introduction..... | 38 |
| III.2 Présentation du système d'analyse..... | 38 |
| III.2.1 Prétraitement..... | 39 |
| III.2.2 Extraction de caractéristiques des visages avec expressions..... | 39 |
| III.2.2.1 Les ondelettes de Gabor..... | 40 |
| III.2.3 Les algorithmes de réduction l'espace de données..... | 42 |
| III.2.3.1 Analyse en Composantes Principales (PCA)..... | 43 |
| III.2.3.2 Analyse Discriminante Linéaire(LDA)..... | 46 |
| III.2.3.3 Le modèle discriminant Linéaire amélioré de Fisher (EFM)..... | 50 |
| III.2.4 Comparaisons..... | 52 |

| | |
|--|-----------|
| III.2.4.1 Comparaisons entre deux vecteurs..... | 52 |
| III.2.4.2 Comparaisons entre deux matrices..... | 53 |
| III.2.4.3 Machine à Vecteurs de Support (SVM)..... | 54 |
| III.2.5 La Décision..... | 57 |
| III.3 Conclusion..... | 57 |
| Chapitre IV Implémentation et résultats | |
| IV.1 Introduction..... | 58 |
| IV.2 La base de données CASIA (3D expressions)..... | 58 |
| IV.2.1 Description de la BDD CASIA 3D expression..... | 58 |
| IV.2.2 Protocole de la BDD CASIA 3D expressions..... | 60 |
| IV.3 Architecture structurelle du système de conception..... | 61 |
| IV.3.1 Prétraitement d'image..... | 62 |
| IV.3.1.1 La composante V de l'espace de couleur TSV de l'image..... | 62 |
| IV.3.1.2 Détection de visage et de points et régions d'intérêts..... | 63 |
| IV.3.2 Extraction des caractéristiques par filtre de Gabor..... | 64 |
| IV.3.3 Réduction d'espace..... | 65 |
| IV.3.3.1 L'analyse en composantes principales(PCA)..... | 65 |
| IV.3.3.2 Le Modèle Discriminant Linéaire Amélioré de Fisher (EFM)..... | 66 |
| IV.3.4 Classification de scores avec (SVM)..... | 66 |
| IV.4 Résultats obtenus..... | 69 |
| IV .5 Conclusion..... | 78 |
| CONCLUSION GÉNÉRALE..... | 79 |
| Bibliographie | |

Introduction
générale

Introduction Générale

De nos jours, beaucoup de technologies biométriques ont été développées, et un bon nombre d'entre elles est utilisé dans une variété d'applications. Parmi ces techniques on retrouve les systèmes d'identification automatique des visages. Cependant, ce système exige que celles-ci soient comparées avec un très grand nombre de visages dans une base de données.

Nous identifions quelqu'un en regardant son visage, puisque chaque personne a généralement des caractéristiques uniques et distinctives dans le visage. Par ailleurs, le visage humain est un grand dispositif de communication, parce que le visage peut évoquer diverses expressions faciales selon les émotions internes. Donc, on peut lire l'état émotionnel de quelqu'un de son expression faciale et d'y répondre de façon appropriée, bien que l'étude sur l'expression du visage a une histoire assez longue. Depuis longtemps Charles Darwin, s'est intéressé à l'analyse automatique des expressions du visage humain. De nos jours et à l'aide de l'ordinateur c'est une tendance récente. Il est connu qu'il y'a six expressions de base (ou prototypique). Les expressions faciales pour les êtres humains à travers les diverses ethnies et cultures sont : le bonheur, la surprise, la tristesse, la peur, la colère et le dégoût.

Notre étude est également basée sur cette hypothèse et traite ces cas. Les êtres humains sont capables de reconnaître les expressions faciales du visage rendu sur l'écran (ou papier) ainsi que du véritable visage. En fait, les chercheurs ont mis au point une série de modèles pour le visage et leurs implémentations pour l'utilisation dans diverses applications d'interaction homme-ordinateur ou pour animer un avatar dans le cyberspace. A titre d'exemple, au début des années 2000 tentative d'unification de la reconnaissance des expressions faciales et l'animation faciale envisagé tels Avatar visage virtuel en tant que future communication des médias. En fait, une telle Avatar a un certain avantage par rapport à un vrai visage, car il peut être anonyme, sympathique, drôle et animé en temps réel de base. La présente étude présentera un pipeline multi-étape:

- Dans la première étape de la reconnaissance du visage on biométrie.
- Dans la deuxième étape état de l'art sur la reconnaissance d'expression faciale.

- Dans la troisième étape étude du modèle de reconnaissance de visages avec expressions.
- Dans la dernière étape la présentation de la base de données CASIA 3D expressions, implémentation et résultats.

Ainsi notre manuscrit est organisé autour de quatre chapitres, de la manière suivante :

- Le premier chapitre expose les technologies utilisées dans les systèmes biométriques pour l'identification de personnes. Nous avons mis en évidence les différentes difficultés inhérentes à la reconnaissance automatique de visages, notamment l'invariance à l'illumination.
- Nous développons dans le deuxième chapitre les principales difficultés inhérentes à la reconnaissance d'expressions faciales, Nous présentons aussi les plus importantes techniques utilisées dans le système de reconnaissance d'expressions faciales et un état de l'art des techniques proposées jusqu'à ce jour.
- Le troisième chapitre est consacré à l'étude de notre modèle d'identification de reconnaissance des visages avec expressions. Nous étudions les méthodes utilisées dans chaque étape du modèle.
- Dans le quatrième chapitre nous détaillons l'implémentation du modèle et toutes les expériences et essais menés pour le choix du modèle optimal retenu. Les résultats sont aussi présentés dans ce chapitre.

Nous terminerons enfin par une conclusion générale dans laquelle nous résumerons le travail réalisé en faisant état de l'analyse globale des résultats.