

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique



Université De Batna
Faculté De Technologie



MEMOIRE

Présenté au

DEPARTEMENT D'ELECTRONIQUE

Pour l'obtention du diplôme de

Magister En Electronique

OPTION : Traitement du signal

Par

LEMMOUCHE Mansoura

Ingénieur, département d'Electronique - Université de Batna

Thème :

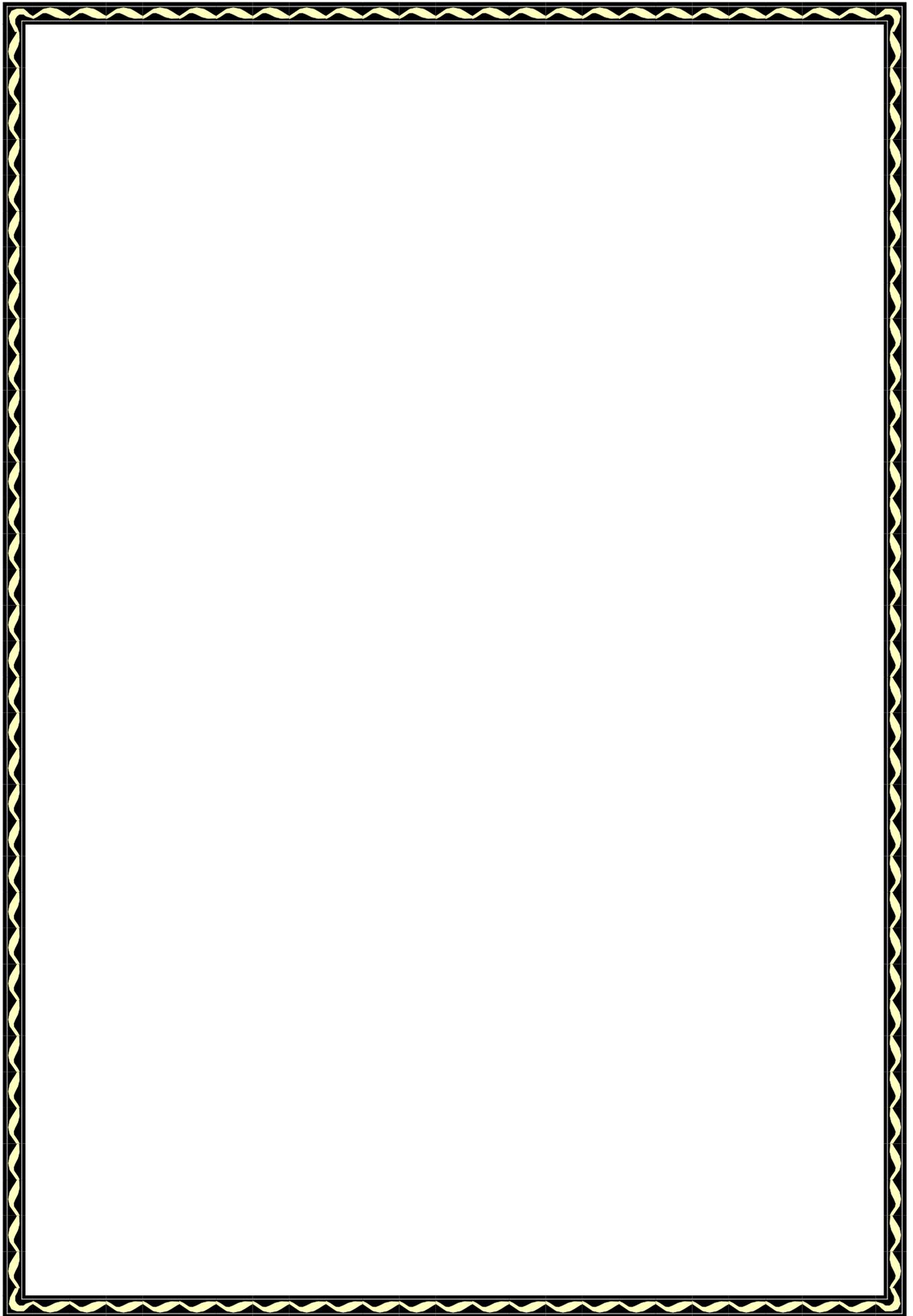
Identification des Visages Humains par
Réseaux de Neurones

Mémoire soutenu publiquement le : / /.

Devant le jury composé de :

<i>Dr D. MELAAB</i>	<i>Université de Batna</i>	<i>Président</i>
<i>Dr N. ATHAMENA</i>	<i>Université de Batna</i>	<i>Rapporteur</i>
<i>Dr A. DIBILOU</i>	<i>Université de Biskra</i>	<i>Examineur</i>
<i>Dr Y. ABDESSAMED</i>	<i>Université de Batna</i>	<i>Examineur</i>
<i>Dr M. S. KHIREDDINE</i>	<i>Université de Batna</i>	<i>Examineur</i>

2012/2013



Résumé : *Le travail proposé ici étudie l'identification (reconnaissance) de l'identité employant des images faciales. Plusieurs techniques de reconnaissance de visage ont été testées sur une base de données standard de visage ORL.*

La première méthode étudiée repose sur une comparaison simple des distances Euclidiennes entre les vecteurs représentant les images pour déterminer la personne la plus favorable (probables). Dans la deuxième méthode, nous avons utilisé une approche statistique, la PCA pour réduire la dimensionnalité des images. Durant les tests, nous avons constaté que ces méthodes produisent des résultats acceptables, mais n'atteignent pas une précision suffisante. Pour améliorer les performances du système d'identifications, nous avons étudié l'utilisation des réseaux de neurones de type MLP. Les essais ont montré aussi que l'algorithme de rétro propagation nécessite une très longue période d'entraînement. La discrimination représente l'étape finale de la reconnaissance. Les résultats que nous avons obtenus sont très encourageants.

Sommaire

CHAPITRE 1 : Introduction générale

1.1 Introduction.....	1
1.1.1 Acquisition de l'image	2
1.1.2 Prétraitement.....	3
1.1.3 Détection puis localisation.....	3
1.1.4 Extraction des paramètres et Classification	3
1.1.5 Apprentissage	3
1.1.6 Décision	4
1.2 Prétraitements sur les images	4
1.2.1 Introduction	4
1.2.2 Traitements des images	4
1.2.2.1 Définition d'une image	4
1.2.2.2 L'image numérique	4
2.2.3 Caractéristiques d'une image numérique	5
1.2.2.4 prétraitement et amélioration d'images	6
1.2.2.4.1 Binarisation	7
2.2.4.2 Egalisation de l'histogramme	7
1.2.2.4.3 Filtrage	8
1.3 Objectif du travail	8
1.4 Présentation du mémoire	11

CHAPITRE 2 : La Biométrie

2.1 Introduction	12
2.2 Les différentes techniques biométriques	13
2.2.1 La reconnaissance des empreintes digitales	13
2.2.2 La reconnaissance de visages	14
2.2.3 La reconnaissance de la main	15
2.2.4 La reconnaissance de l'iris	15
2.2.5 La reconnaissance de la rétine	16
2.2.6 La reconnaissance vocale	16
2.2.7 La reconnaissance de la dynamique de signature	17
2.2.8 La reconnaissance de la dynamique de la frappe au clavier	17

2.3 Détails des techniques les plus utilisées	17
2.3.1 Principe de fonctionnement	17
2.3.2 La reconnaissance des empreintes digitales	18
2.3.3 La reconnaissance des visages	21
2.3.4 La reconnaissance de la main	23
2.3.5 La reconnaissance de l'iris	24
2.4 Conclusion	26

CHAPITRE 3 Techniques de Reconnaissance de visages

3.1 Introduction	27
3.2 Motivation	27
3.3 Processus de reconnaissance des visages	28
3.4 Méthode de vérification ou identification de visage	29
3.4.1 Méthodes Géométriques	30
3.4.2 Méthodes Globales	31
3.4.2.1 Approche ACP	31
3.4.2.2 Approche Corrélacion	31
3.4.2.3 Approche DCT	32
3.4.2.4 Approche Neuronal	32
3.5 Méthode de reconnaissance par distance Euclidienne	33
3.6 Reconnaissance de visage par l'Eigen Face	34
3.7 Résultats de Simulation	39
3.8 Conclusion	46

CHAPITRE 4 : Reconnaissance de visage par Réseau de Neurons MLP

4.1 Introduction	47
4.2 Généralité sur les Réseaux de neurones Artificiels	47
4.2.1 Introduction et Historique	47
4.2.2 Du neurone biologique au modèle mathématique	49
4.2.3 Architecture des réseaux de neurones	53
4.2.3.1 Connectivité	54
4.2.4 Apprentissage	55
4.2.4.1 Types d'apprentissage	55
4.2.4.2 Application des réseaux de neurones selon le mode d'apprentissage	56
4.2.5 Les types de réseaux	56

4.2 .5.1 Le perceptron	57
4.2 .5.2 Perceptron multicouche	58
4.2 .5.2.1 Apprentissage du perceptron multicouche	59
4.2 .5.2.2 Réseau de rétro-propagation	59
4.2 .5.2.3 Algorithme de rétro propagation	59
4.2 .5.2.4 Les paramètres d'apprentissage	61
4.2 .5.3 Le réseau à fonction de base radiale	62
4.3 Reconnaissance de visage : approche Mono-Réseau MLP	63
4.3.1 Base utilisée	63
4.3.2 Structure du réseau neuronal	63
4.3.3 Reconnaissance pour un ensemble autorisée	64
4.4 Reconnaissance de visage par PCA+Réseau MLP	67
4.4.1 Processus d'identification	69
4.5 Conclusion	71

Conclusion générale

Conclusion générale.....	72
Bibliographie	75
Annexe	77

Liste des figures

Figure1.1 : Système de reconnaissance	2
Figure1.3: Histogramme des niveaux de gris d'une image	6
Figure 1.4: Utilité d'égalisation d'histogramme	8
Figure2.1: La méthode eigenface	22
Figure 2.2: http://vismod.media.mit.edu/vismod/demos/facerec/system.html	22
Figure 3.1: Processus de reconnaissance des visages	30
Figure 3.2: Phase d'apprentissage	33
Figure 3.3: Phase de test.	34
Figure 3.4: Exemple d'Eigen Faces	34
Figure 3.5: Organigramme de la phase d'apprentissage de ACP	37
Figure 3.6: Phase d'apprentissage	38
Figure 3.7: Phase de test.....	39
Figure 3.8: Base de donnée ORL	40
Figure 3.9: Variation du taux de reconnaissance en fonction de la dimension de l'image.	41
Figure 3.10: Variation du taux de reconnaissance en fonction du N ^{bre} d'Eigen Face	42
Figure 3.11: La base d'apprentissage contient 9 images.	43
Figure 3.12: La base de test contient 9 images.	43
Figure 3.13: La base d'apprentissage contient 25 images	44
Figure 3.14: La base de test contient 25 images	45
Figure 4. 1 : Représentation schématique d'un neurone biologique.	49
Figure 4.2 : Représentation d'un neurone mathématique.....	51
Figure 4.3 : La fonction sigmoïde	51
Figure 4.4 : La fonction gaussienne	52
Figure 4.5 : La fonction à seuil	53

Figure 4.6 : La fonction linéaire par parties	53
Figure 4.7: Réseau à connexions locales.	54
Figure 4.8 : Réseau à connexions totales	54
Figure 4.9 : Architecture du perceptron.	57
Figure 4.10 : Architecture du perceptron multicouche.	58
Figure. 4.11 : Architecture du réseau RBF	62
Figure 4.12: Modèle de reconnaissance mono réseau	64
Figure 4.13: variation de taux de reconnaissance en fonction de nombre de neurones dans la couche cachée.	66
Figure 4. 14: Variation de taux de reconnaissance avec le seuil	67
Figure 4.15: Organigramme générale d'apprentissage	69
Figure 4.16: Processus d'identification	70

Liste des Tableaux

Tableau 3.1 : Traits utilisés pour identifier un visage par la méthode des faces.	31
Tableau 3.2: Variation du taux de reconnaissance en fonction de la dimension de l'image.	40
Tableau 3.3: Variation de taux de reconnaissance en fonction de N ^{bre} d'Eigen Face.	41
Tableau 4.1 : Taux de reconnaissance en fonction de dimension de l'image, et nombre de neurones cachés.	65
Tableau 4.2: Taux de reconnaissance en fonction du seuil.	67
Tableau 4.7 : Taux de reconnaissance en fonction du seuil.	71
Tableau 5.1 : les meilleurs résultats de reconnaissance obtenus	73

Introduction Générale

1.1 Introduction :

La reconnaissance de formes est une partie essentielle de n'importe quel système d'analyse d'image de haut niveau. Le but d'un système typique de vision par ordinateur est d'analyser des images d'une scène donnée et d'identifier le contenu de la scène. De tels systèmes sont maintenant en service dans différents domaines, tels que la robotique, la reconnaissance militaire, la télédétection, le traitement de documents, et l'automatisation industrielle [1].

La reconnaissance automatique des formes est un domaine qui connaît depuis une dizaine d'années une forte évolution, en grande partie liée à l'augmentation de la puissance de calcul des ordinateurs et à la grande variété d'application de ces techniques [2].

La reconnaissance de formes est l'opération d'identification automatique de certaines structures: images, sons, spectres, ...etc. Dans ce contexte, le concept de forme doit être compris dans un sens élargi. La forme est une description quantitative ou structurale d'un objet ou d'une entité [3].

Le problème de la reconnaissance de formes se pose de plus en plus souvent dans la technique du traitement de l'information. Les formes qui possèdent certaines propriétés communes sont groupées en classes [3]. La reconnaissance de formes par des machines signifie que les formes soumises à ce traitement sont réparties dans les classes auxquelles elles appartiennent. Cette opération doit être exécutée automatiquement, avec un minimum d'intervention humaine [2]. Donc la reconnaissance de formes consiste à l'identification automatique d'une forme donnée et à la décision de l'appartenance de cet objet à une classe existante dans la base de données.

Les moyens classiques de vérification d'identité pour les contrôles d'accès, comme les passeports ou carte d'identité, les mots de passe ou les codes secrets, peuvent être facilement falsifiés. La solution apparaît pour remédier à ces inconvénients est d'utiliser la biométrie tout en basant la vérification sur les aspects propres de chaque individu. Cette technologie est maintenant envisagée pour un très grand nombre d'autres applications. Dans notre travail nous nous intéressons à la reconnaissance de visage.

La reconnaissance de visages humains est considérée actuellement comme l'une des tâches les plus importantes pour relever le défi en reconnaissance des formes. La facilité et la précision avec lesquelles nous pouvons identifier nos amis et nos ennemis, même dans des conditions défavorables, sont des capacités des plus étonnantes du système visuel humain.

Un système de reconnaissance de visage est un système d'identification et de vérification d'individus, qui permet de vérifier si une personne appartient à la base de donnée du système, et de l'identifier si c'est le cas. On peut représenter les systèmes de reconnaissance par le diagramme suivant [4]:

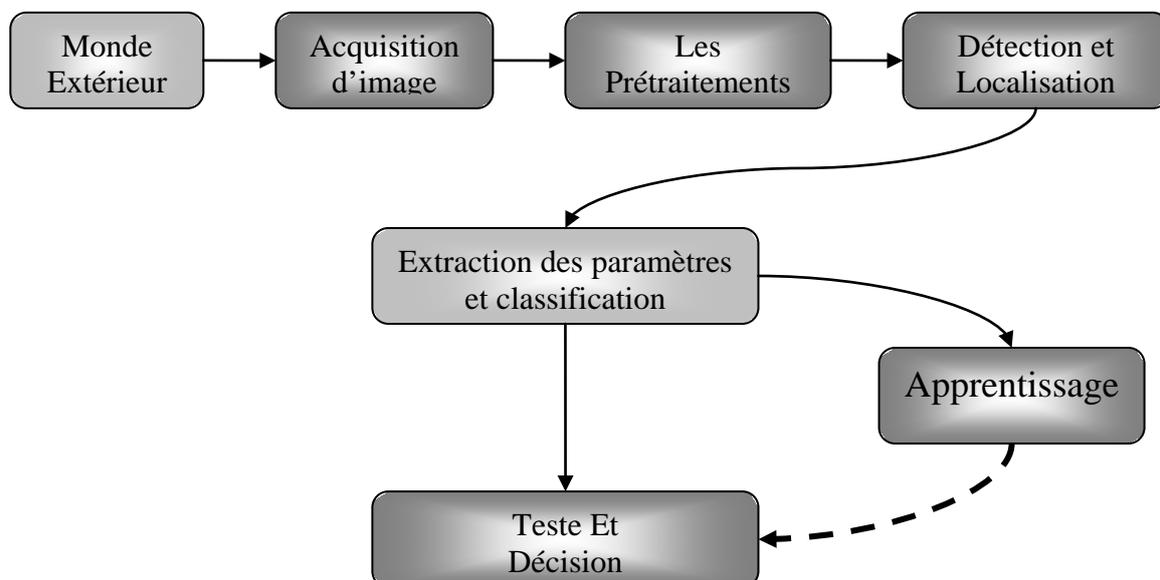


Figure1.1 Système de reconnaissance

Comme l'indique la figure précédente, un système de reconnaissance passe par plusieurs étapes qui sont:

1.1.1 Acquisition de l'image :

C'est l'opération qui permet d'extraire du monde réel une représentation bidimensionnelle pour des objets en 3D, cette opération peut être statique (Appareil Photo,

Scanner...etc.) ou dynamique (Caméra, WebCam), dans ce cas on aura une séquence vidéo. A ce niveau on aura une image brute.

1.1.2. Prétraitement :

Les données brutes issues des capteurs sont les représentations initiales des données, à partir desquelles des traitements permettent de construire celles qui seront utilisées pour la reconnaissance. L'image brute peut être affectée par différents facteurs causant ainsi sa détérioration, elle peut être bruitée, c'est à dire contenir des informations parasites à cause des dispositifs optiques ou électroniques. Pour pallier à ces problèmes, il existe plusieurs méthodes de traitement et d'amélioration des images, telle que: la normalisation, l'égalisation de l'histogramme, etc.

1.1.3. Détection puis localisation :

Les systèmes de reconnaissance de visages sont complexes. La difficulté réside notamment dans la partie détection automatique du visage, bien que nous développerons surtout la partie reconnaissance, il est intéressant de parler de l'étape de détection automatique du visage qui est très importante dans un système de reconnaissance. Ce qui rend la détection de visages dans une image très difficile, c'est surtout la complexité du décor, les variations de poses, les conditions de lumières généralement inconnues, etc. Il existe plusieurs méthodes qui peuvent être appliquées à la détection automatique des visages [8]. Il faut détecter la présence d'un visage dans l'image, ensuite le localiser en vue d'extraire les traits pour le caractériser et le différencier des autres. Le résultat de cette étape est l'obtention de la partie d'image à traiter.

1.1.4 Extraction des paramètres et Classification :

Dans cette étape on extrait de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase de Décision. Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes. Ces informations seront ensuite classées, en d'autre terme, affectées à la classe la plus proche, les individus ayant des similarités sont regroupés dans la même classe. Ces classes varient selon le type de décision.

1.1.5 Apprentissage :

L'apprentissage consiste à mémoriser les modèles calculés dans la phase analyse pour les individus connus. Un modèle est une représentation compacte des images qui permet de faciliter la phase de reconnaissance mais aussi de diminuer la quantité de données à stocker en quelque sorte l'apprentissage est la mémoire du système.

1.1.6 Décision :

La décision est la partie du système où on tranche sur l'appartenance d'un individu à l'ensemble des visages ou pas, et si oui quelle est son identité. Donc la décision c'est l'aboutissement du processus. On peut le valoriser par le taux de reconnaissance (fiabilité) qui est déterminé par le taux de justesse de la décision.

1.2. Prétraitements sur les images

1.2.1 Introduction :

L'utilisation des images aide considérablement l'homme à communiquer avec les siens. Le traitement numérique des images est l'ensemble des méthodes qui cherchent à rendre cette opération possible, et plus simple.

Les débuts du traitement d'images remontent au 18^{ème} siècle avec l'apparition de l'appareil photo puis l'introduction de l'électronique au 19^{ème} siècle.

Le traitement d'image trouve son application dans diverses disciplines, telles que : l'électronique, l'informatique, la neurophysiologie, les statistiques, etc.

1.2.2. Traitements des images :

1.2.2.1 Définition d'une image :

C'est la représentation d'une personne ou d'un objet par la peinture, la sculpture, la photographie, etc. C'est aussi un ensemble structuré d'informations qui après affichage sur écran, ont une signification pour l'œil humain. Elle peut être écrite sous forme de fonction $I(x, y)$ où I est une fonction d'intensité lumineuse ou de couleur aux coordonnées spatiales (x, y) . De cette façon l'image est exploitable par la machine, d'où la nécessité de sa numérisation.

1.2.2.2 L'image numérique :

Il est clair que les images manipulées par l'ordinateur sont numériques (série de bits). L'image numérique est l'image dont la surface est divisée en éléments de taille fixe appelés pixels, ayant comme caractéristique le niveau de gris ou de couleur. La numérisation d'une image est la conversion de celle-ci en une image numérique représentée par une matrice bidimensionnelle de valeurs numériques $f(x, y)$ qui sont les niveaux de gris des coordonnées réelles (x, y) .

1.2.2.3 Caractéristiques d'une image numérique :

Pixel :

C'est le plus petit point de l'image. Chaque pixel a une valeur numérique qui représente le niveau de gris ou de couleur selon la nature de l'image.

Dimension :

C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentant les intensités lumineuses (pixel).

Résolution :

C'est la clarté ou la finesse des détails atteinte par un moniteur ou une imprimante dans la production d'image, sur les moniteurs d'ordinateurs, la résolution est exprimée en nombre de pixel par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels affichable horizontalement ou verticalement sur un moniteur; plus grand est ce nombre, meilleure est la résolution.

Voisinage d'un pixel [5] :

Le voisinage d'un pixel est l'ensemble des pixels qui lui sont adjacents. Selon le codage de FREEMAN on distingue deux types de voisinages : 4-connexité et 8-connexité.

Niveaux de gris :

C'est la valeur numérique qui reflète l'intensité de la luminosité d'un point. Pour 256 niveaux de gris compris entre 0 et 255, chaque pixel sera codé sur 8 bits, et ses niveaux de gris seront obtenus après dégradation de la couleur noire, 0 représente le blanc et 255 représente le noir.

Image en couleur :

C'est une image où chaque pixel est codé dans l'espace de couleur RGB (rouge, vert, bleu). Donc c'est une représentation dans un espace tridimensionnel de la valeur d'intensité lumineuse du pixel. Ce dernier sera codé sur trois octets, un pour chacune des couleurs.

Bruit :

C'est un phénomène de brusques variations d'un pixel par rapport à ces voisins suivant un certain seuil. Il existe quatre sources de dégradation induisant le bruit, qui sont : le bruit lié au contexte d'acquisition, le bruit lié au capteur, le bruit lié à l'échantillonnage, et celui lié à la nature de la scène.

Histogramme:

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image. Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différentes, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant. Il permet de donner un grand nombre d'informations sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (couleurs) dans le cas d'une image trop claire ou d'une image trop foncée. Il peut être utilisé pour améliorer la qualité d'une image (rehaussement d'image) en introduisant quelques modifications, pour pouvoir extraire les informations utiles de celle-ci.

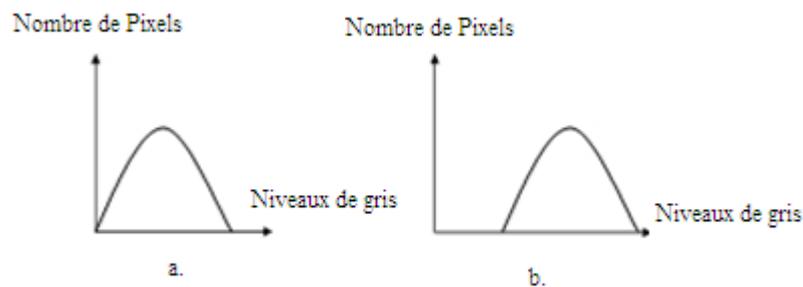


Figure1.3: Histogramme des niveaux de gris d'une image (a) image trop foncée, (b) image trop claire.

1.2.2.4 Prétraitement et amélioration d'images :

L'image brute est une structure simple dans son organisation mais complexe dans son contenu. Ceci est dû à plusieurs facteurs, d'une part la grande quantité d'informations qui y réside, d'autre part le nombre important de processus de l'utilisateur qui tentent à introduire des distorsions indésirables, donc à la dégrader. On peut citer comme élément de perturbation: l'éclairage qui peut être trop faible et rend l'image trop sombre, le bruit électronique inhérent au capteur qui engendre une imprécision dans la restitution des niveaux de gris et la qualité optique de l'objectif qui peut engendrer des aberrations.

L'amélioration d'images consiste en un ensemble de méthodes destinées à améliorer l'aspect visuel d'une image. Il existe différentes approches visant à améliorer une image:

- Les opérations basées sur l'examen de l'histogramme.
- Le filtrage, dont le but est de minimiser l'effet du bruit dans l'image.

1.2.2.4.1 Binarisation:

C'est la transformation de l'image à plusieurs niveaux de gris en une image binaire (à deux niveaux de gris). Elle peut être réalisée en choisissant un certain seuil, et tous les pixels ayant des valeurs inférieures à celui-ci deviennent noirs, tous ceux de valeurs supérieures deviennent blancs. Nous pouvons aussi effectuer la binarisation en choisissant un certain intervalle [Min, Max], et les pixels qui se trouvent à l'intérieur de ce dernier sont forcés à 1 et les autres à 0.

1.2.2.4.2 Egalisation de l'histogramme :

Dans une image naturelle qui a été quantifiée de manière linéaire, une majorité de pixels ont une valeur inférieure à la luminance moyenne. C'est pourquoi les détails dans les régions sombres sont difficilement perceptibles. Pour résoudre ce problème, on utilise alors une méthode appelée égalisation de l'histogramme. L'égalisation procède de la manière suivante: On part du niveau zéro et on calcule la valeur cumulée des niveaux suivants dans l'histogramme original jusqu'au moment où la somme est la plus près possible d'une certaine

valeur; en général, cette valeur est égal à $\frac{P \times P}{N_f}$ qui est le nombre de pixels moyen, avec

P est la dimension de l'image en pixels, et N_f est le nombre de niveaux de gris dans l'image résultat.

La figure 1.4 illustre le principe de l'égalisation d'histogramme.

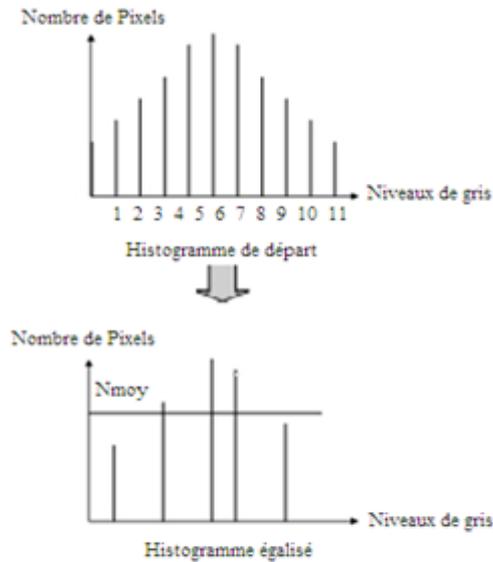


Figure 1.4: Utilité d'égalisation d'histogramme.

1.2.2.4.3 Filtrage:

Tout ce qui est indésirable dans une image est considéré comme un phénomène de brusque variation d'un pixel isolé par rapport à ses voisins. Il peut provenir de l'éclairage, des dispositifs optiques et électroniques du capteur ou de l'image elle-même. Pour lutter contre les effets du bruit, il est nécessaire de faire des opérations de filtrage sur chaque pixel en tenant compte de son voisinage.

Les opérations de filtrage, qui ont pour objet d'atténuer l'effet du bruit et faire disparaître les défauts et les perturbations continues dans une image, sont indispensables pour une bonne extraction de l'information utile. On distingue deux types de filtrage: le filtrage linéaire où la transformation d'un pixel est le fruit d'une combinaison linéaire des pixels voisins, et le filtrage non linéaire où les pixels voisins interviennent suivant une loi non linéaire.

1.3 Objectif du travail

Le visage est une source riche d'informations sur le comportement humain. La capacité pour l'identification des visages est très importante dans notre vie sociale, particulièrement pour se rappeler et estimer les émotions des personnes que nous

rencontrons. Nous pouvons identifier des milliers de visages appris durant toute notre vie, et identifier les visages familiers d'un coup d'œil même après des années de séparation, même dans des conditions défavorables. Les humains peuvent identifier et interpréter des visages et des expressions faciales dans une scène avec peu ou pas d'effort. Cette compétence est tout à fait robuste, en dépit de grandes variations des caractéristiques faciales (modèle de cheveux, lunettes, ...etc.) et des états de scène autour d'un visage.

Au fil du temps, grâce à la puissance croissante de l'informatique, grâce au prix décroissant des ordinateurs, des applications d'informatique sont devenues populaires dans la vie quotidienne. De nos jours, ils existent des ordinateurs capables d'effectuer des tâches, qui faites par l'homme, demandent de l'intelligence.

Dans le domaine de la vision par ordinateur, l'identification automatique de visage est un sujet très intéressant. Dans les années dernières, il a attiré beaucoup l'attention des chercheurs. Le but de la reconnaissance des visages est de concevoir des systèmes informatiques capables d'égaliser les êtres humains dans ce domaine. Il y a beaucoup de secteurs dans lesquels l'identification de visage peut jouer un rôle important. Certains ne sont pas nécessairement des applications de sécurité élevée, mais l'identification de visage peut aider à surmonter un grand nombre de problèmes non résolus d'identification, en particulier dans les secteurs où l'identification instantanée de visage est nécessaire. [6, 7]

Le présent travail est consacré à un sous-domaine bien spécifié qui est la reconnaissance des visages. L'outil choisi à cette fin est les réseaux de neurones statiques, à savoir les réseaux MLP qui sont utilisés comme outil de classification et de décision. A titre de comparaison d'autres techniques classiques comme les faces propres et l'ACP seront implémentées.

Une caractéristique des réseaux de neurones est leur capacité à apprendre (par exemple à reconnaître une lettre, un son, un visage...). Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple (de la même manière qu'un enfant apprend à reconnaître un chien à partir d'exemples de chiens qu'il a vu). Ils ont donc une capacité à classer, généraliser, à mémoriser et aussi à oublier. Un réseau de neurones reconnaîtra d'autant plus facilement un objet qu'il l'aura "vu" souvent.

Par leur capacité d'apprentissage, les réseaux de neurones sont susceptibles d'améliorer le taux de reconnaissance et d'identification de visages. Plusieurs architectures

seront testées à savoir, des classifieurs mono-réseau avec comme entrées des données des images brutes, ou des ACP.

Dans notre travail nous avons utilisé un modèle paramétrique. Ce type de modèle consiste à caractériser une image suivant un vecteur d'éléments caractéristiques. Le choix d'un tel type de modèle est actuellement très fréquent à cause de sa simplicité de construction. La seule difficulté consiste à choisir pour une application donnée une liste de paramètres caractéristiques qui permettent une discrimination entre les différents objets (images).

Pour notre application, nous avons utilisé deux méthodes d'extraction de paramètres caractéristiques :

- Le filtrage bilinéaire apporté par la fonction MATLAB `imresize` permet de lisser l'image tout en conservant la morphologie dans son ensemble, se rapproche fortement d'une convolution par un masque moyennneur 3×3 .
- L'analyse en Composante Principale (ACP): est une technique standard dans l'analyse de données, elle est utilisée pour la réduction de dimensions ou d'une manière équivalente pour l'extraction de caractéristiques dans la représentation d'un signal.

Chaque image pour chaque méthode sera représenter par un ensemble de ces derniers dans une espace à n dimensions, appelé espace de représentation

Une fois le vecteur caractéristique des images est extrait, l'étape suivante consiste à le présenter comme entrée au réseau de neurones, le MLP; ce réseau est entraîné dans la phase d'apprentissage par la technique: algorithme de rétro propagation, Par la suite le taux de réussite est calculé pour chaque méthode.

Les expériences, menées dans ce mémoire, utilisent une base d'images, très populaire dans ce domaine de recherche, qui est l'ORL (Olivetti Research Laboratory). Cette base contient un ensemble de visage pris entre avril 1992 et 1994 au laboratoire d'Olivetti Research à Cambridge. La base de visages ORL (disponible gratuitement voir [8]) contient 40 personnes distinctes, chaque personne ayant 10 images différentes de son visage, ce qui donne 400 images de visages au total, avec 256 niveaux de gris et de taille 112×92 . Toute les images ont été prises sur un fond homogène foncé, et elles sont obtenues dans différentes situations, telles que temps différent, différents angles, différentes expressions (fermer les yeux,

ouvrir les yeux, sourire, surpris, heureux...) et différents détails (avec lunette, sans lunette, barbe, modèle de cheveux), et une inclinaison de plus ou moins 20 degrés. L'ensemble de ces images va être divisé en deux sous ensemble. Un ensemble représentant la base d'apprentissage pour entraîner le réseau de neurones; soit au total : 240 images. Le deuxième ensemble est l'ensemble des images de test (des images non utilisées dans la phase d'apprentissage), soit au total: 160 images. Cette division est menée d'une manière aléatoire grâce à une routine matlab fournie par Chang et Para [9],

1.4 Présentation du mémoire :

Le présent mémoire comporte quatre chapitres répartie comme suit:

1. L'introduction relative aux objectifs visés par ce travail a été déjà donnée dans le présent Chapitre.
2. Le deuxième chapitre présente les différents moyens d'authentification biométrique, leurs marchés, leurs domaines d'utilisation, leurs principes de fonctionnement ainsi que leurs performances.
3. Dans le troisième chapitre Nous étudierons de façon plus approfondie l'un des moyens d'authentification biométrique qui est la reconnaissance de visage, et nous présenterons les résultats expérimentales des méthodes de reconnaissance classiques, telles que la méthode de la distance Euclidienne et la méthode ACP.
4. La première partie du quatrième chapitre donne un aperçu général sur les réseaux de neurones MLP; les définitions et les concepts de base. La deuxième partie du chapitre présente les résultats des expériences entreprises avec des classifieurs mono-réseau, avec comme entrées les pixels des images ou les composantes principales. Les résultats obtenus, avec comme outil de décision le seuil.
5. La conclusion générale est consacrée à une évaluation globale des résultats obtenus par chaque méthode. Les avantages et les inconvénients de chaque technique sont évalués. Elle met aussi l'accent sur les difficultés rencontrées et les perspectives possibles dans ce champ de recherche très vaste.

La Biométrie

2.1 Introduction

Devant la croissance exponentielle des communications tant physiques que virtuelles et les risques que cela peut représenter, il est apparu nécessaire de contrôler l'identité des acteurs de ces échanges. La biométrie permet de vérifier que l'utilisateur est bien la personne qu'il prétend être. C'est une technologie qui utilise les caractéristiques physiques propres à chaque individu pour établir de façon aussi fiable que possible son identité. Jouissant actuellement d'un certain engouement dû, sans doute, aux différents gadgets d'identification que l'on a pu voir dans certaines productions cinématographiques, la biométrie tend à envahir notre quotidien. Devant cette déferlante, il était nécessaire de faire le point sur ce qu'est exactement la biométrie, quelles techniques existent vraiment et leur degré de fiabilité pour ensuite détailler les plus utilisées [10,11].

Qu'en est il de la biométrie.

De plus en plus, notre société éprouve le besoin de se contrôler. Que ce soit pour garantir la sécurité des gens dans les lieux publics ou pour éviter le détournement ou le vol d'informations sensibles. Il n'est pas rare qu'une personne ait à retenir plus d'une dizaine de codes d'accès tant dans son quotidien que dans son milieu professionnel et c'est souvent considéré comme frustrant par l'utilisateur. Il existe plusieurs façons de prouver son identité, mais deux grands modes se dégagent. L'identification par laquelle on cherchera à savoir si la personne est bien monsieur X en vérifiant par rapport à un code saisi que les données biométriques analysées sont bien en rapport. L'autre mode est l'authentification par laquelle on cherche simplement à savoir qui est la personne qui cherche à se faire reconnaître du système.

La biométrie, pour mettre en évidence l'identité d'une personne, utilise plusieurs technologies.

- Les analyses biologiques: analyses de sang, de salive, d'urine, d'odeurs ou encore d'ADN. Ces méthodes sont difficiles à mettre

en œuvre pour une utilisation courante. Je ne les décrirai donc pas ici.

- Les analyses comportementales: on analyse la dynamique de signature qui vérifie la pression du stylo sur le support ainsi que sa vitesse de déplacement et ses accélérations. On peut également analyser la façon d'utiliser un clavier, c'est-à-dire vitesse et force de frappe ou encore la démarche, le mouvement des hanches, des bras et des épaules.
- Les analyses morphologiques: les empreintes digitales, la forme de la main, les traits du visage, la forme de l'iris ou de la main, le dessin veineux de la rétine sont autant de caractéristiques qui existent chez toutes les personnes et qui pourtant sont propres à chacune d'entre elles.

Bien que jugée plus sûre et moins contraignante qu'un système à mots de passe, la biométrie doit cependant être capable de s'affranchir de l'évolution dans le temps des éléments qui permettent notre identification. Une personne n'a pas le même visage à dix-huit ans qu'à cinquante, son empreinte peut être altérée par une coupure ou une brûlure. L'œil peut subir des lésions. C'est pour cette raison que les systèmes biométriques sont souvent couplés entre eux [12,13].

2.2 Les différentes techniques biométriques[14],[15],[16],[17],[18],[19],[20],[21]

2.2.1 La reconnaissance des empreintes digitales.

La plus ancienne et la plus répandue des technologies d'identification. Elle représente plus de cinquante pour cent des applications. La donnée de base dans le cas des empreintes digitales est le dessin représenté par les crêtes et sillons de l'épiderme. Ce dessin est unique et différent pour chaque individu. En pratique, les informations fournies par ce dessin sont tellement nombreuses qu'il est quasiment impossible de toutes les utiliser. On préférera donc en extraire les caractéristiques principales telles que les bifurcations de crêtes, les "îles", les lignes qui disparaissent. Une empreinte complète contient environ une centaine de ces points caractéristiques encore appelés "minuties". Si l'on considère la zone réellement scannée, on peut extraire environ

quarante de ces minuties. La loi impose un minimum de douze pour identifier avec certitude une personne car il est statistiquement impossible de trouver deux individus présentant les mêmes douze points caractéristiques, même en considérant une population de plusieurs dizaines de millions de personnes.

La petite taille du lecteur en fait une technologie facilement intégrable pour la plupart des applications. Le traitement est rapide et fiable si le capteur est suffisamment évolué pour pouvoir détecter les moulages ou doigts coupés le plus souvent par une simple mesure de température. Cette technologie présente pourtant deux défauts. La propreté ou l'humidité de la main nuisent à la qualité de la mesure et peuvent générer de faux rejets et elle jouit auprès du grand public d'une image très "policière" et est de ce fait peu appréciée.

2.2.2 La reconnaissance de visages.

Cette technologie qui représente quinze pour cent des applications consiste à faire une photographie plus ou moins évoluée pour en extraire un ensemble de facteurs qui se veulent propres à chaque individu. Ces facteurs sont choisis pour leur forte invariabilité et concernent des zones du visage tel que le haut des joues, les coins de la bouche, l'écartement des yeux et la taille de la bouche par exemple. On évitera d'autre part les types de coiffures, les zones occupées par des cheveux en général ou toute zone sujette à modification durant la vie de la personne. Cette technique est capable de déjouer le port de lunettes, de barbe ou d'autres artifices. Seule une opération intervenant sur le cartilage est capable de la tromper de façon certaine.

Il existe plusieurs variantes de la technologie de reconnaissance du visage. La première est développée et supportée par le MIT et se nomme "Eigenface". Elle consiste à décomposer le visage en plusieurs images faites de nuances de gris, chacune mettant en évidence une caractéristique particulière. Une autre technique appelée "feature analysis" se base sur la précédente en y ajoutant des informations sur les distances entre les éléments et sur leurs positions. Elle se dit plus souple quant aux éventuelles modifications pouvant survenir : angle de prise de vue, inclinaison de la tête. Des techniques moins utilisées, basées sur des réseaux neuronaux sont actuellement testées. C'est une technique bien acceptée par le public car elle ne demande aucune action de la personne dont on veut vérifier l'identité. C'est une méthode peu intrusive.

2.2.3 La reconnaissance de la main.

Ce type de mesure biométrique est également très répandu, notamment aux Etats-Unis. Il représente dix pour cent des applications. Il s'appuie sur une image en trois dimensions de la main. Cela consiste à mesurer plusieurs caractéristiques de la main telles que la forme de la main, la longueur et la largeur des doigts, les formes des articulations, les longueurs entre les articulations. Jusqu'à quatre-vingt-dix caractéristiques différentes peuvent être mesurées. Cette technologie est peu sensible à l'état de la main: la saleté ou les coupures ne l'empêcheront pas de fonctionner. Elle est simple à mettre en œuvre. En revanche elle est sensible aux modifications de la forme de la main que peut provoquer le vieillissement ou encore un régime. La taille du capteur empêche son utilisation dans des applications bureautiques, dans une voiture ou encore sur un téléphone. De plus, avec cette technologie, il est assez difficile de différencier deux personnes de la même famille ou bien encore des jumeaux.

2.2.4 La reconnaissance de l'iris.

La technologie de reconnaissance de l'iris, qui représente six pour cent des applications, est très intéressante pour la biométrie, car elle permet de différencier même des jumeaux ou l'oeil gauche de l'oeil droit d'une même personne. L'iris est indépendant du code génétique de l'individu, et très difficilement falsifiable. En effet, l'iris présente une quasi-infinité de points caractéristiques qui ne varient pratiquement pas pendant la vie contrairement à sa couleur qui, elle, peut changer. Mais cela n'a aucune influence car les images d'iris obtenues par les capteurs sont en noir et blanc. Le seul problème de cette technique est lié à la mesure en elle-même, qui peut être source d'erreurs ou de problèmes. Ainsi, on peut quasiment dire que le nombre de problèmes rencontrés lors de cette mesure augmente proportionnellement avec la distance entre l'oeil et la caméra. D'autres problèmes se posent à cause des reflets. Il faut être très attentif à l'éclairage. Pour palier la détection de faux yeux ou de photographies, on fait appel à des caractéristiques dynamiques de l'oeil qui prouvent son authenticité, comme la réactivité de la pupille par exemple. Cette technique est assez délicate à mettre en œuvre et très intrusive.

2.2.5 La reconnaissance de la rétine.

Cette mesure biométrique est plus ancienne que celle utilisant l'iris, mais elle est très peu utilisée. Elle a été moins bien acceptée par le public et les utilisateurs, sans doute à cause de son caractère trop contraignant : la mesure doit s'effectuer à très faible distance du capteur, à peine quelques centimètres. Le capteur effectue ensuite un balayage de la rétine. Il est physiquement impossible d'effectuer une mesure rétinienne à plus de trente centimètres sur un sujet mobile, comme on peut le voir dans certains films. Cette méthode requiert des sujets coopératifs et entraînés. Pourtant cette technique semble être tout aussi fiable que celle de l'iris. Elle se base sur le fait que le schéma et le dessin formés par les vaisseaux sanguins de la rétine sont uniques pour chaque personne, différents entre jumeaux et assez stables tout au long de la vie. La mesure peut ainsi fournir jusqu'à quatre cents points caractéristiques. C'est une mesure très fiable, très difficile à falsifier. La capture de l'image est très intrusive et mal acceptée par le public.

2.2.6 La reconnaissance vocale.

La mesure biométrique de la voix traite des données qui proviennent à la fois de facteurs physiologiques dépendants de l'âge, du sexe, de la tonalité, de l'accent et de facteurs comportementaux comme la vitesse et le rythme. Ils ne sont en général pas imitables. C'est la seule technique qui permette à l'heure actuelle de reconnaître une personne à distance. Cependant cette technique est très facilement falsifiable, en utilisant un enregistrement. Cette technique nécessite en plus une excellente qualité d'enregistrement ce qui n'est pas possible dans les lieux publics par exemple. Le peu de différences qui existe entre deux voix rend également cette technique peu fiable.

2.2.7 La reconnaissance de la dynamique de signature.

La reconnaissance de signature est la seule technologie, classée comme "comportementale", qui soit véritablement aboutie. Cette technologie permet d'identifier un utilisateur en le faisant signer avec un stylo électronique. Les produits les plus évolués sont capables d'évaluer, en plus du tracé de la signature, la façon dont le stylo est manipulé, sa vitesse, ses accélérations, la pression et l'inclinaison avec laquelle il a été utilisé. Cependant, le tracé est particulièrement instable. Certaines personnes ont une signature très erratique qui peut changer sensiblement au cours du temps. Chaque personne a un style d'écriture unique. On peut donc définir, à partir de

la signature, un modèle qui pourra être employé pour effectuer une identification. De plus, la signature est utilisée dans beaucoup de pays comme élément juridique ou administratif. Elle permet de justifier de la bonne foi d'une personne ou de la confondre devant des documents signés. Cette méthode peu intrusive est bien acceptée par le grand public.

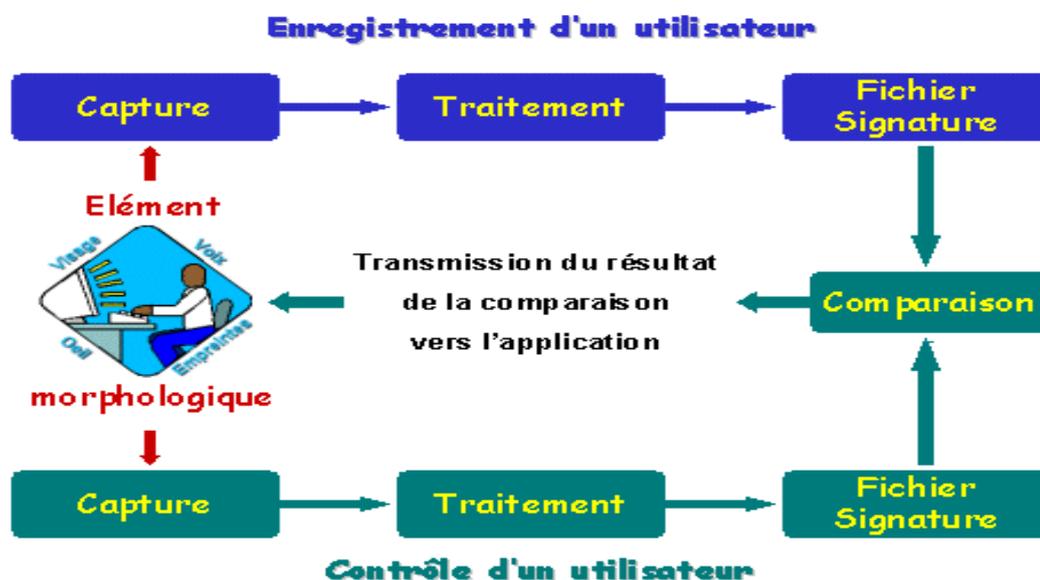
2.2.8 La reconnaissance de la dynamique de la frappe au clavier.

La reconnaissance de la dynamique de frappe au clavier est une autre technologie "comportementale". Elle est l'équivalent de l'étude graphologique d'une personne. Les durées entre frappes, la fréquence des erreurs et la durée de la frappe elle-même sont étudiées de façon statistique. En revanche, cette technologie peu intrusive est tributaire de l'état physique et psychique de la personne qui utilise le clavier. La fatigue, le stress sont autant de facteurs qui feront varier la qualité de la frappe.

2.3 Détails des techniques les plus utilisées

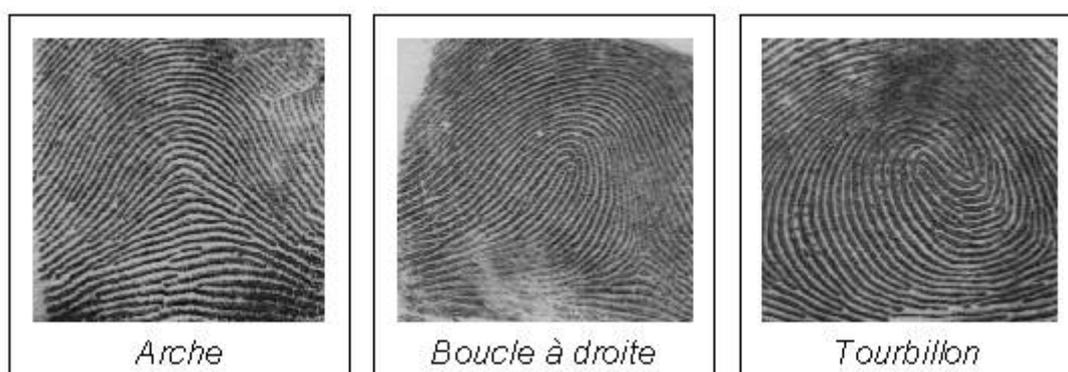
2.3.1 Principe de fonctionnement.

Dans un premier temps, il faut créer un modèle qui servira de gabarit aux futures identifications de la personne. Ensuite, chaque fois que la personne voudra pénétrer dans un lieu sécurisé ou accéder à des données protégées il lui faudra se soumettre à un test permettant de l'identifier par comparaison avec le gabarit mémorisé. Ce principe sera le même quelle que soit la technologie utilisée. Le schéma ci-dessous illustre ces différentes séquences.

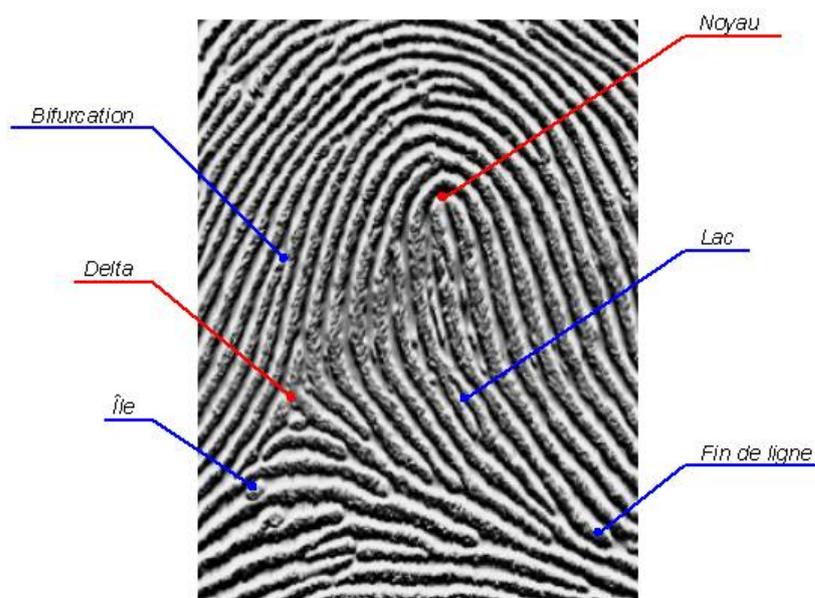


2.3.2 La reconnaissance des empreintes digitales.

L'empreinte digitale est le modèle du relief de la peau des doigts. Ce relief, formé pendant la période fœtale est classé d'après le système de "Henry" selon la topographie générale de l'empreinte. On distingue trois grandes familles. Les empreintes dites en "arches", en "boucles" qu'elles soient à droite ou à gauche et les empreintes en "tourbillon".

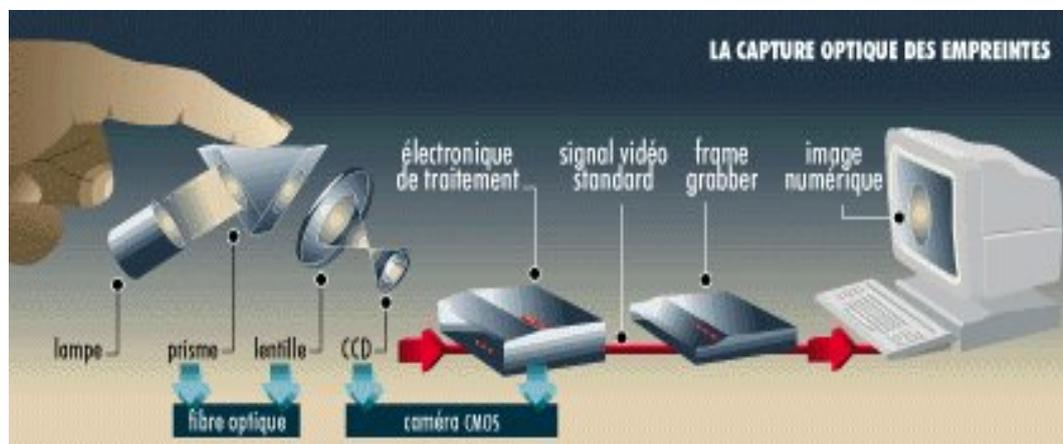


Sur ces empreintes, pour les différencier, on recherchera les accidents encore appelés minuties.



Le principe du traitement est le suivant.

Tout d'abord un système d'acquisition permet d'enregistrer l'image du doigt d'une personne voulant s'identifier. Cette image est ensuite transmise à un système de traitement qui va l'analyser et la comparer à celles déjà enregistrées.



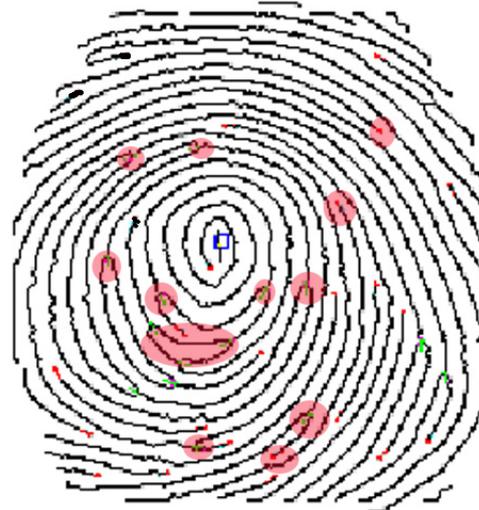
L'image capturée est envoyée au système de traitement.



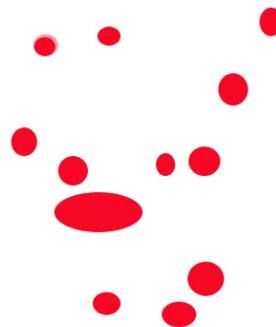
Afin d'y être dans un premier temps binarisée pour en augmenter le contraste, en diminuer le volume et ainsi accroître la rapidité du traitement.



Ensuite le système extrait les minuties qu'il considère comme fiables. C'est-à-dire les minuties qui ne risquent pas d'être issues de défauts lors de la capture de l'image.



Cette extraction permet d'établir un gabarit de l'empreinte qui sera ensuite comparé à celui déjà mémorisé permettant d'identifier ou non la personne qui a soumis son empreinte au test.



2.3.3 La reconnaissance des visages.

Rien n'est plus naturel qu'utiliser le visage pour identifier une personne. Les images faciales sont probablement la caractéristique biométrique la plus communément employée par l'homme pour effectuer une identification personnelle. L'utilisation d'une caméra permet de capter la forme du visage d'un individu et d'en dégager certaines particularités. Selon le système utilisé, l'individu doit être positionné devant l'appareil ou peut être en mouvement à une certaine distance. Les données biométriques qui sont obtenues sont par la suite comparées au fichier référence. La méthode la plus utilisée est la méthode Eigenface. Elle utilise une représentation des éléments caractéristiques d'une image de visage à partir d'images modèles en niveau de gris. Des variantes de Eigenface sont fréquemment utilisées

comme base pour d'autres méthodes de reconnaissance comme c'est le cas dans la méthode illustrée ci-dessous (figure 2.1).

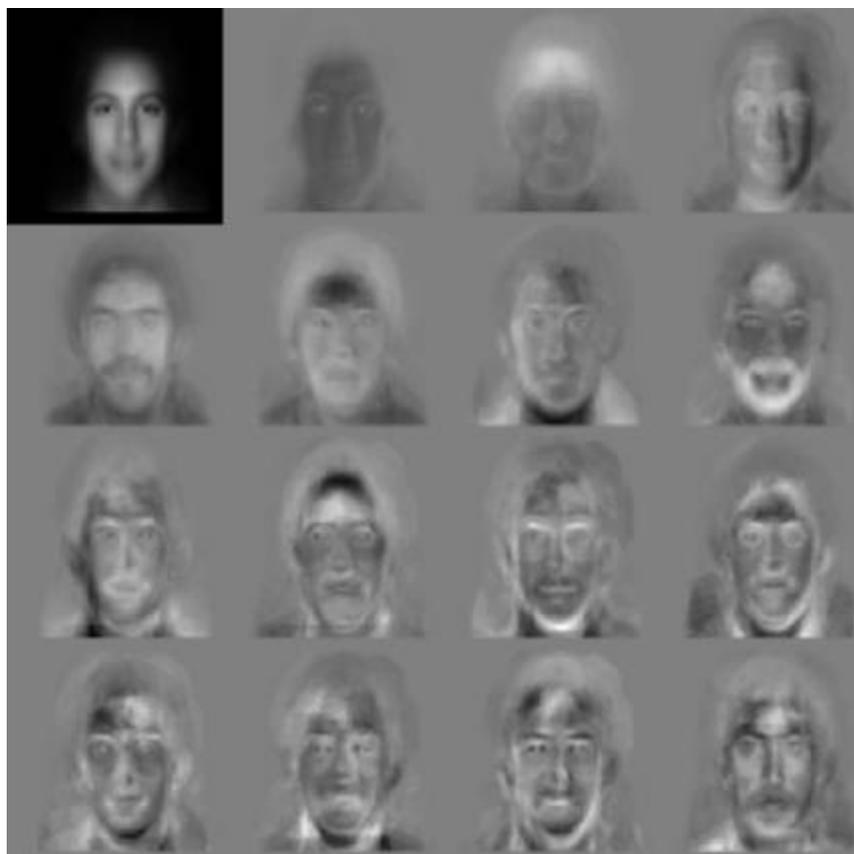


Figure2.1: La méthode eigenface

Une méthode de reconnaissance automatique des visages basée sur Eigenface.

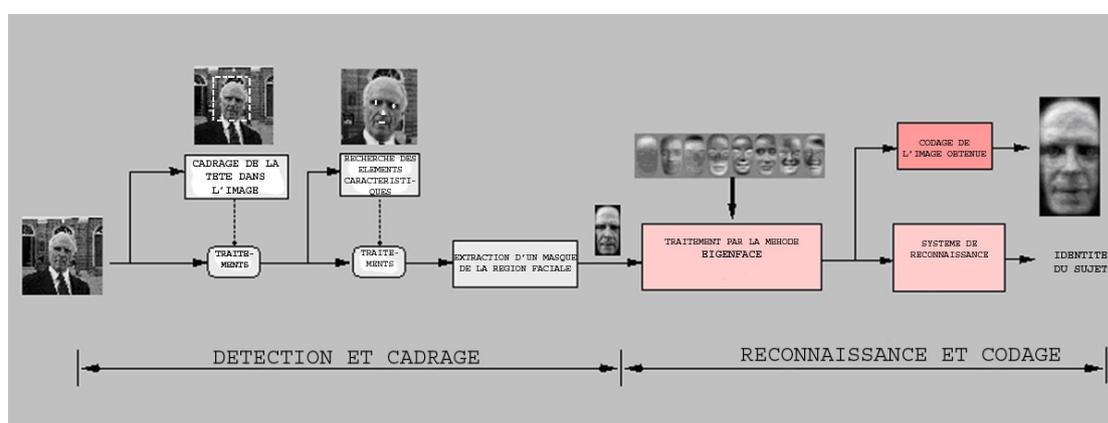


Figure 2.2: <http://vismod.media.mit.edu/vismod/demos/facerec/system.html>

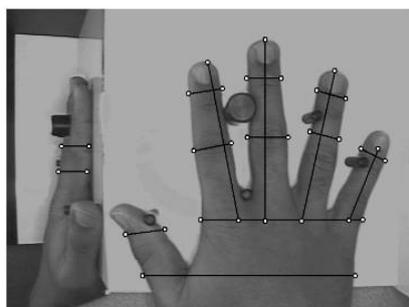
2.3.4 La reconnaissance de la main.

La biométrie par la forme de la main est une technologie populaire qui est largement employée pour le contrôle d'accès physique ou le pointage horaire. Elle est très simple et bon marché. L'exactitude d'un système biométrique basé sur la forme de la main est tout à fait raisonnable. Les éléments pris en compte ne reposent que sur la géométrie de la main et non sur l'empreinte palmaire. Le système prend une photo de la main et examine quatre vingt dix caractéristiques, y compris la forme tridimensionnelle de la main, la longueur et la largeur des doigts ainsi que la forme des articulations.

Pour utiliser la géométrie de la main, l'utilisateur place sa main sur une platine possédant des guides pour positionner les doigts.



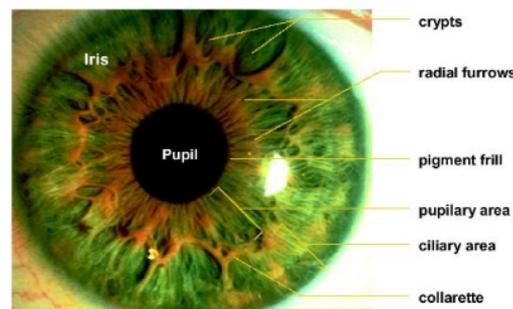
En général, une caméra infrarouge prend l'image de la main sous deux angles différents pour obtenir les trois dimensions comme l'illustre la photographie ci dessous. Les mesures qui sont faites ne prennent que les dimensions de certaines parties de la main et non la pigmentation de la peau ou les lignes situées au niveau des articulations. Le gabarit qui en résulte est d'une taille très faible, une dizaine d'octets, et permet de faire de l'authentification.



2.3.5 La reconnaissance de l'iris.

L'utilisation de l'iris pour l'identification d'une personne a été proposée à l'origine en 1936 par l'ophtalmologue Frank Burch. Dans les années 1980, l'idée réapparut à travers certaines productions cinématographiques. En 1987, les ophtalmologues Aran Safir et Léonard Flom ont fait breveter cette idée et en 1989 ils ont demandé à John Daugman d'essayer de créer des algorithmes pour l'identification par l'iris. Ces algorithmes basés sur la méthode des ondes de Gabo ont été brevetés par Daugman en 1994. Ils sont à la base de tous les systèmes d'identification par l'iris.

Ce sont là encore les particularités des iris de chacun qui sont utilisées.

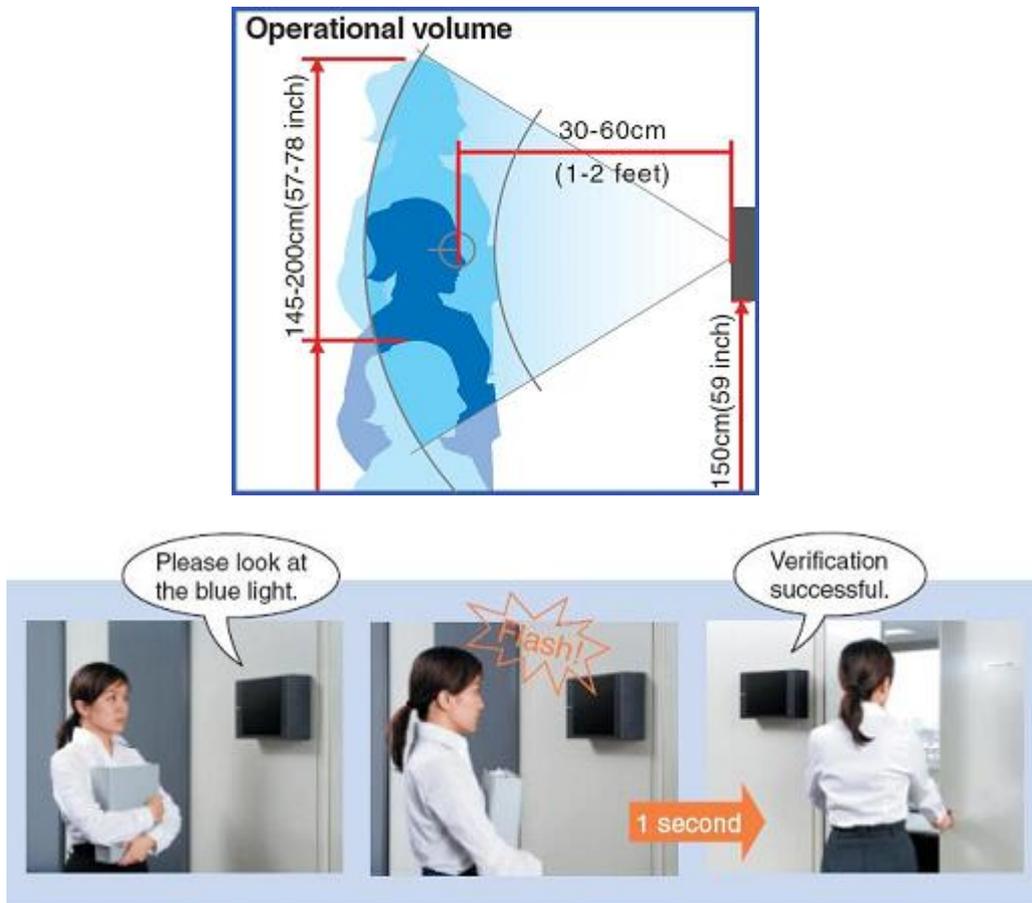


L'identification par l'iris utilise plus de paramètres que les autres méthodes d'identification et la fiabilité résultante est suffisante pour ne plus faire de l'identification mais de l'authentification. La probabilité de trouver deux iris suffisamment identiques pour être confondus est de 1 sur 10 puissance 72 selon les estimations de Daugmann.

La première étape, difficile, est la capture de l'image de l'iris. L'œil est en effet un organe très sensible à la lumière et à la fatigue, deux facteurs qui peuvent faire varier sa taille et sa netteté. De plus, il est souvent obscurci par les cils, les paupières, les lentilles, les réflexions de la lumière ou les mouvements incontrôlés de la personne. Le système de prise d'image doit être rapide, précis et ne doit pas utiliser une lumière qui se refléterait sur l'œil. Le système de numérisation doit permettre d'obtenir au moins soixante-dix pixels de rayon sur l'iris, et dans la plupart des cas, le

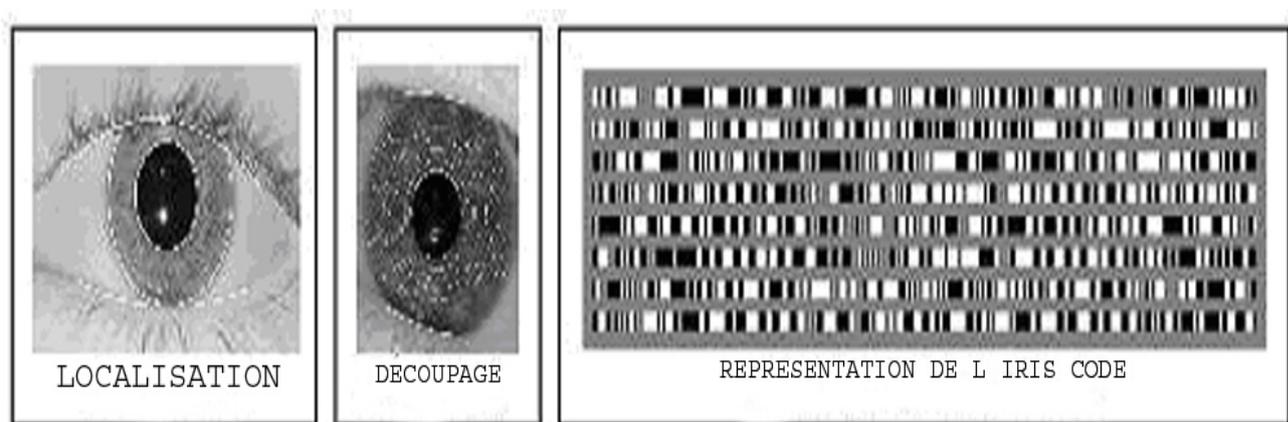
rayon est compris entre cent et cent quarante pixels. Une caméra CCD monochrome (640 x 480) est employée avec une source de lumière de longueur d'onde comprise entre sept cents et neuf cents nanomètres, invisible pour les humains.

Les figures ci-dessous montrent la position que doit avoir le capteur par rapport à la personne et la séquence d'acquisition.



La méthode employée pour la caractérisation d'un iris est celle brevetée par John Daugman. Après la numérisation de l'image de l'oeil, le logiciel va déterminer le centre de la pupille et la zone où se trouve l'iris. Puis sur deux secteurs angulaires, le logiciel découpe des bandes de taille régulière et génère un gabarit à partir de l'analyse locale de la texture de l'iris, comme l'illustre la figure ci-dessous.

Une fois le code de l'iris obtenu il est comparé à la base de données et l'identification est effectuée.



2.4 Conclusion

On peut constater que la biométrie est une véritable alternative aux mots de passe et autres identifiants. Elle permet de vérifier que l'utilisateur est bien la personne qu'il prétend être. Cette technologie est en pleine croissance et tend à s'associer à d'autres technologies comme la carte à puce. La fabrication des produits d'authentification est en pleine augmentation, due en l'occurrence à la nécessité croissante du besoin de sécurité de chacun aussi bien dans le cadre professionnel que privé. La diminution des coûts de ces technologies tend à les démocratiser aidée en cela par un phénomène de mode très sensible actuellement. En revanche, ces systèmes ne sont pas si fiables que les médias voudraient nous le faire croire et, souvent, pour passer de l'identification à l'authentification, il faut mettre en place des systèmes plus compliqués, comme la reconnaissance rétinienne ou encore couplés plusieurs systèmes entre eux. L'avenir de la biométrie sera certainement dans les reconnaissances comportementales. Souvent peu intrusives, elles sont bien admises par le public.

Techniques de Reconnaissance de visages

3.1 Introduction:

Depuis quelques années, on observe un besoin croissant pour des systèmes automatiques d'identification de personnes. Rien n'est plus naturel que d'utiliser le visage pour identifier une personne. Les images faciales sont probablement la caractéristique biométrique la plus communément employée par l'homme pour effectuer une identification personnelle [22]. La reconnaissance et l'identification de visages joue un rôle fondamental dans nos interactions sociales, elle est basée sur notre capacité de reconnaître les personnes, elle ne présente pas de difficultés énormes pour un être humain, mais elle constitue pour tout système informatique une situation extrêmement délicate. Un être humain moyen atteint un taux de reconnaissance de 99% peut-on atteindre ce taux ou le dépasser avec un système informatique [23], [24]

La reconnaissance des visages humains a pris de l'ampleur vers la fin des années soixante dix pour devenir un domaine de recherche très actif [25]. S'investir dans le domaine de la reconnaissance de visages est sans doute motivé par la multiplicité et la variété des champs d'application (contrôle des frontières, surveillance des lieux publics tels que les banques, les aéroports, les centres commerciaux, etc.) [26].

L'identification d'une personne, qui peut être réalisée à partir d'une image de l'individu, plus particulièrement de son visage, est une tâche cognitive complexe, [27]. La vision numérique vise ainsi l'acquisition, le traitement et l'interprétation de ces images pour réaliser la reconnaissance des personnes. Ces systèmes sont particulièrement intéressants car ils permettent la surveillance silencieuse d'un endroit, sans requérir la coopération des individus. [27]

3.2 Motivation [28]

La reconnaissance des visages est l'une des techniques de biométrie la plus utilisée, ceci est dû à ses caractéristiques avantageuses dont on peut citer:

-Disponibilité des équipements d'acquisition et leur simplicité.

-Passivité du système: un système de reconnaissance des visages ne nécessite aucune coopération de l'individu, du genre mettre le doigt ou la main sur un dispositif spécifique ou parler dans un microphone. En effet, la personne n'a qu'à rester ou marcher devant une caméra pour qu'elle puisse être identifiée par le système. En plus, cette technique est très efficace pour la situation non standard, c'est le cas où on ne peut avoir la coopération de l'individu à identifier, par exemple lors d'une arrestation des criminels [29].

Certes que la reconnaissance des visages n'es pas la plus fiable comparée aux autres techniques de biométrie, mais elle peut être ainsi si on utilise des approches plus efficaces en plus du bon choix des caractéristiques d'identification représentant le visage en question.

3.3 Processus de reconnaissance des visages [7]:

Tout processus automatique de reconnaissance des visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage est une entité dynamique qui change constamment sous l'influence de plusieurs facteurs. Un système de reconnaissance des visages a pour objectif de prédire la classe d'appartenance d'un visage inconnue. La figure 3.1, illustre la démarche générale adoptée pour réaliser de tels systèmes.

Dans le monde physique, il y a trois paramètres à considérer: l'éclairage, la variation de la posture et l'échelle. La variation de l'un de ces trois paramètres peut construire une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents. Le codage consiste en l'acquisition d'image et sa digitalisation.

La première étape dite de Prétraitement consiste à éliminer le bruit par des techniques de traitement et de restauration d'images et procéder à une localisation de la tête dans l'image, cette opération est très complexe, beaucoup supposent que l'image contient un visage et que l'arrière plan est neutre. La restauration d'images ou l'élimination du bruit consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image. Dans l'étape Analyse (appelée aussi indexation,

représentation, modélisation ou extraction de caractéristiques), il faut extraire de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase Décision. Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes.

L'apprentissage consiste à mémoriser les représentations calculées dans la phase Analyse pour les individus connus dans une base de signature.

La décision: pour estimer la différence entre deux images, il faut introduire une mesure de similarité.

3.4 Méthodes de vérification ou identification de visages

Au début des années 1970, la reconnaissance par le visage était principalement basée sur des attributs faciaux mesurables comme l'écartement des yeux, des sourcils, des lèvres, la position du menton, la forme, etc. Depuis les années 1990, les différentes technologies utilisées exploitent toutes les découvertes effectuées dans le domaine du traitement d'image, et plus récemment encore les réseaux de neurones [22].

Actuellement ils existent de nombreuses méthodes qui permettent de reconnaître et d'identifier une personne dans une image. On peut diviser ces méthodes en deux catégories, les méthodes géométriques et les méthodes globales.

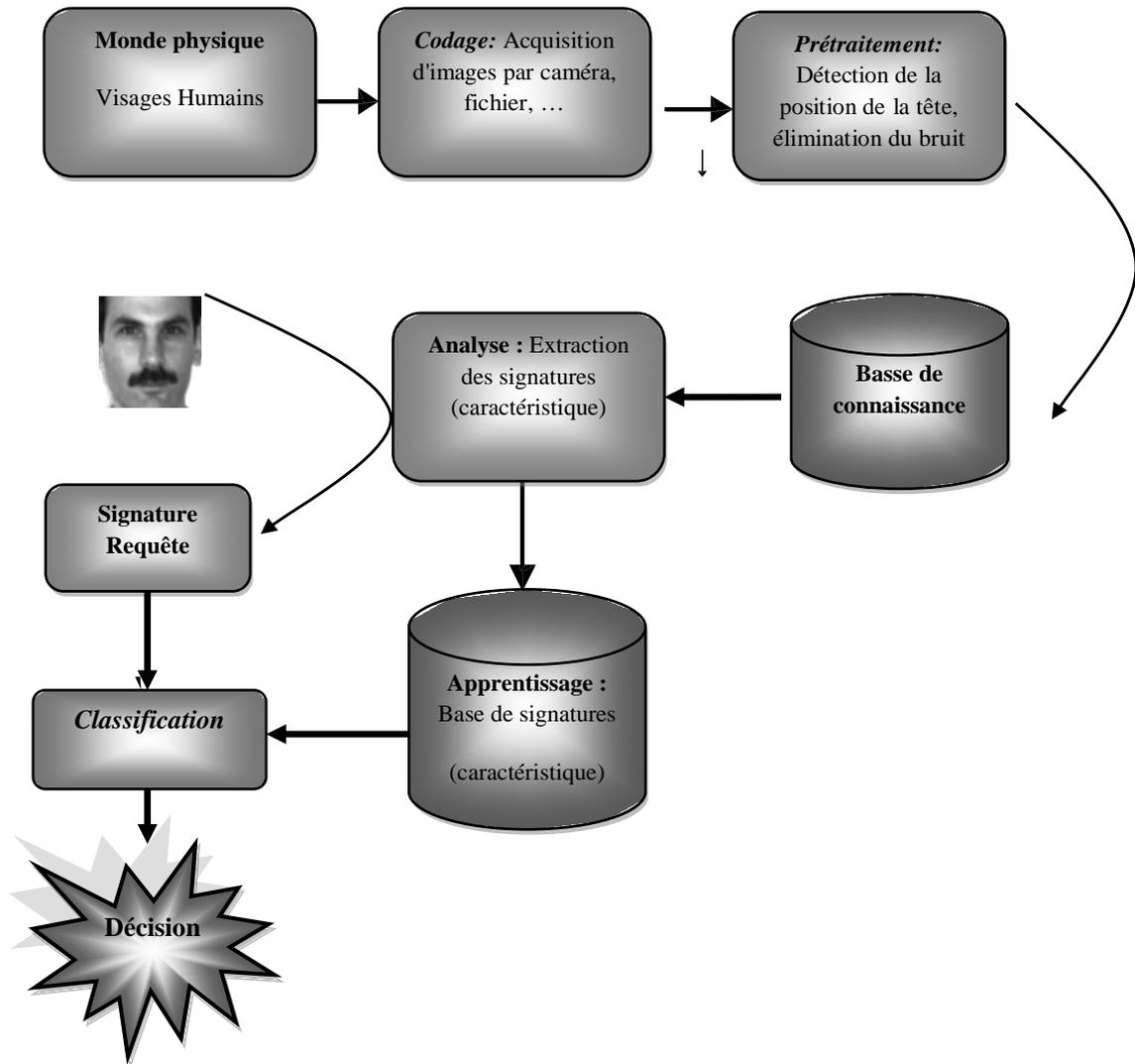


Figure 3.1: Processus de reconnaissance des visages

3.4.1 Méthodes Géométriques [30]:

Plusieurs travaux pour identifier les visages avec des traits tirés de face ont été développés par Kanade en 1973 [31]. Les traits utilisés mesurent les différents aspects des cheveux, sourcils, yeux, bouche, etc. Le système calcule la distance entre le visage inconnu et les visages de la base, le visage ayant la plus petite distance est celui qui correspond à l'inconnu. Les résultats ont montré que seuls 6 à 7 traits sont suffisants pour identifier la plupart des visages. Le tableau 3.1 suivant récapitule quelques traits utilisés pour l'identification par la méthode géométrique (traits de face) [32].

Catégorie	Traits
Yeux	Forme, couleurs, distance entre les yeux
Bouche	Gabarit, largeur, longueur, superficie de la bouche ouverte
Sourcils	Séparation, épaisseur.
Cheveux	Intensité, forme, couleur.
Lèvres	Largeur, forme, couleur
Joues	Intensité.
Nez	Longueur, largeur
Distances	D (yeux, centre du nez), D (mention, bouche)...
Rapports	D (centre de la face, bouche) / D (mention, bouche) ...

Tableau 3.1 : Traits utilisés pour identifier un visage par la méthode des faces.

3.4.2 Méthodes Globales [7] [26]:

Cette classe regroupe les méthodes qui mettent en valeur les propriétés globales de la forme. Le visage est traité comme un tout. Parmi les approches les plus importantes réunies au sein de cette classe, on trouve:

3.4.2.1 Approche ACP (ou les visages propres)

Son but est de capturer la variation dans une collection d'images de visages et d'utiliser cette information pour coder et comparer les visages (en termes mathématique: trouver les vecteurs propres de la matrice de covariance de l'ensemble des images de visages). Le nombre possible de visages propres peut être approximé en utilisant seulement les meilleurs visages propres qui correspondent aux plus grandes valeurs propres. Cette approche rencontre le problème du coût des calculs élevé et celui de la détermination du nombre de visages propres utiles.

3.4.2.2 Approche Corrélation:

La technique de corrélation est basée sur une comparaison simple entre une image test et les visages d'apprentissage. Celui d'entre eux se trouvant à la plus faible distance du visage test sera sélectionné comme premier choix.

Plusieurs métriques peuvent être utilisées afin d'évaluer cette valeur comme par

exemples les distances L_1 (city-block) et L_2 (euclidienne), la cross-corrélation, la distance de Mahalanobis, etc. Ce processus de décision est communément appelé algorithme du K plus proche voisin.

Malgré sa grande simplicité, cette méthode n'offre cependant pas d'avantages particulièrement intéressants. En effet, elle n'utilise pas des informations de plus haut niveau, comme la variation d'éclairage et les changements physiques (p. ex.: barbe).

3.4.2.3 Approche DCT:

L'utilisation de la transformée de cosinus discrète (Discrete Cosine Transform ou DCT) à des fins de reconnaissance de visage est assez récente. Similaire aux Faces propres d'un point de vue mathématique, elle est par contre beaucoup plus rapide, tant en phase d'apprentissage qu'en phase de reconnaissance.

Cela étant dit, chaque image de visage est représentée par un vecteur composé des premiers coefficients de la transformée DCT. Lorsqu'un visage est présenté au module, sa transformée est calculée et un certain nombre de coefficients est réalisée à l'aide de la distance L_1 ou avec d'autres métriques pertinentes.

3.4.2.4 Approche Neuronal:

Cette technique envisagée utilise des réseaux de neurones comme engin d'apprentissage et de reconnaissance. Pour débiter, une image brute (ou prétraitée) de dimensions fixes constitue habituellement la source d'entrées des réseaux. Les dimensions doivent être établies au préalable car le nombre de neurones sur la couche d'entrée en dépend.

Cela étant dit, plus les dimensions de l'image sont élevées, plus la complexité et le temps d'apprentissage augmentent.

Certains auteurs ont par ailleurs utilisé des variantes de la technique de base en modifiant les données d'entrée. Les coefficients de projections d'images dans un espace des visages (Eigen Faces) peuvent par exemple être utilisée comme source d'information. Cette méthode peut évidemment être étendue aux coefficients de DCT, de Fourier, etc.

3.5 Méthode de reconnaissance par distance Euclidienne :

C'est la méthode la plus classique, elle consiste simplement à mesurer la distance euclidienne entre l'image de test et toutes les images présentes dans la base d'apprentissage afin d'en déterminer la plus faible.

- Phase d'apprentissage :

La phase d'apprentissage correspondrait à un enroulement réel de personnes qui seraient enregistrées dans une base de données.

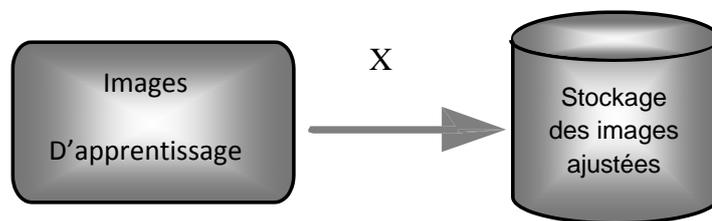


Figure 3.2: phase d'apprentissage

- Phase de test (Reconnaissance):

Lorsqu'une nouvelle image de la base de test est présentée, elle est comparée avec toutes les images de la phase d'apprentissage. La comparaison est faite par un calcul de la distance Euclidienne entre l'image (vecteur) de test et les images (vecteurs) de la base d'apprentissage. Il semble logique que plus la distance entre deux images est petite plus ces deux images se ressemblent. Ainsi, le résultat de la reconnaissance est l'image de la base d'apprentissage qui ressemble le plus à la nouvelle image présenté

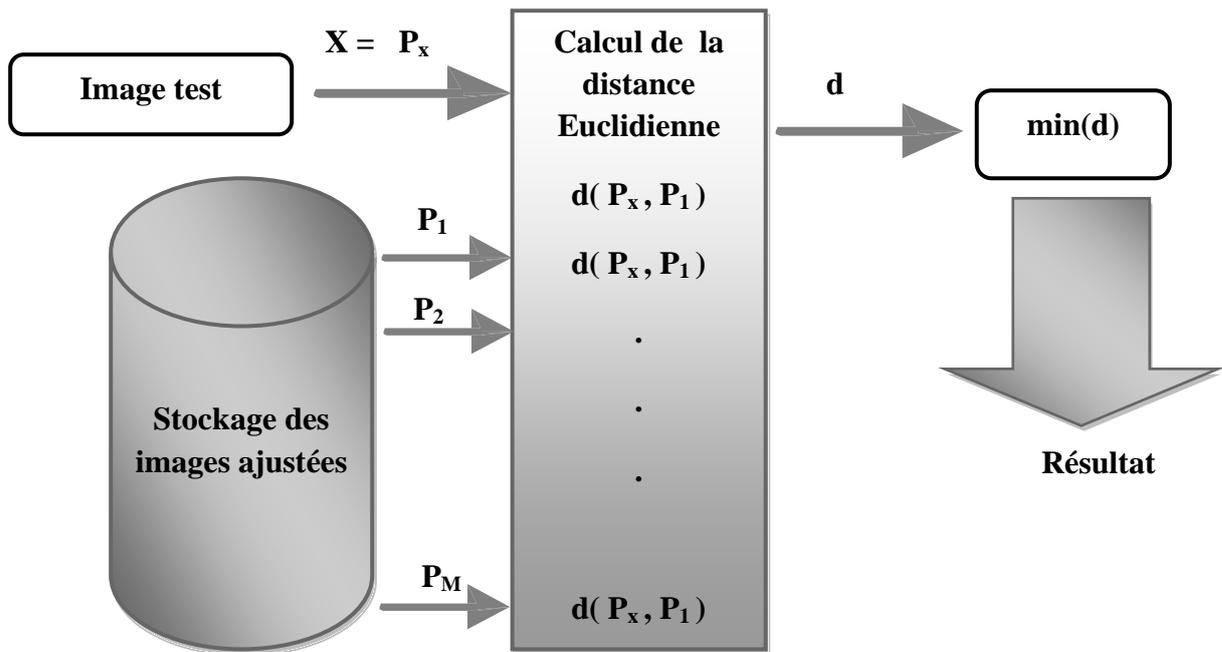


Figure 3.3: phase de test.

3.6 Reconnaissance de visage par L'Eigen Faces (EF) : [27]

L'utilisation des méthodes statistiques appliquées à la modélisation et à la reconnaissance de visage est largement répandue. Kirby et Sirovich ont d'ailleurs utilisé la transformée K-L (c.-à-d. : Karhunen-Loève) afin de coder des visages et ainsi réduire la dimensionnalité de leur représentation. En 1991, Turk et Pentland introduisirent le concept d'Eigen faces à des fins de reconnaissance. Basée sur une analyse en composantes principales (ACP), la méthode des EF repose sur une utilisation des premiers vecteurs propres comme visages propres, d'où le terme Eigen faces.



Figure 3.4: Exemple d'Eigen Faces

La base formée par ces vecteurs génère alors un espace utilisé pour représenter les images des visages. Les personnes se voient donc attribuer un vecteur

d'appartenance pour chacune de leur image.

Cela étant dit, la reconnaissance est réalisée en comparant les coefficients de projection d'un visage test avec ceux appartenant aux visages d'entraînement. La méthode est relativement rapide en phase de reconnaissance et peut également bénéficier de plusieurs optimisations algorithmiques.

Cette méthode est tout d'abord basée sur une analyse en composantes principales (ACP) Appliquée sur l'ensemble des visages d'une banque d'entraînement. Elle consiste essentiellement à effectuer une réduction de dimensionnalité en codant les visages dans une nouvelle base formée par les premiers vecteurs propres (c.-à-d.: Eigen Faces) provenant du calcul de l'ACP.

Les Eigen Faces associés aux plus fortes valeurs propres représentent donc, dans l'espace des images, les directions dans lesquelles les variations sont les plus marquées. C'est ainsi que les premiers visages propres représentent habituellement les différences d'éclairage ainsi que les personnes portant des lunettes ou une barbe.

- Phase d'apprentissage :

La phase d'apprentissage correspondrait à un enroulement réel de personnes qui seraient enregistrées dans une base de données. Cette phase consiste donc à récolter une grande quantité d'images de visage afin de se constituer une base de données de départ. Dans un premier temps, on construit une matrice contenant N images de la base d'apprentissage, puis l'image moyenne est calculée. On réajuste ensuite les données par rapport à la moyenne, pour pouvoir suivre, de manière simple, le comportement des valeurs d'écart type, de variance et de covariance (rappelons nous que toutes ces formules font intervenir une soustraction par rapport à la moyenne). On applique alors un algorithme de reconnaissance globale à cette matrice réajustée. La chose à retenir est que ces algorithmes fournissent en sortie ce que l'on appelle une matrice de projection W qui va nous être très utile dans la seconde partie de la phase d'apprentissage. Cette phase consiste en la projection des images apprises sur un espace vectoriel dont les vecteurs sont les éléments de notre matrice de projection W . Toutes ces projections sont finalement stockées dans une grande base de données [33].

La phase d'apprentissage (ou de modélisation) des Eigen Faces se déroule comme suit:

1. Un visage moyen ψ est calculé à partir des N images d'entraînement I_i de dimensions $L \times H$

$$\psi = \frac{1}{N} \sum_{i=1}^N I_i$$

2- Ce visage moyen est soustrait des images d'apprentissage (on élimine donc les ressemblances pour se concentrer sur les différences), ce qui génère les vecteurs de différences associés à chacune des images :

$$\phi_i = I_i - \psi$$

3- La matrice de covariance C est construite (approche inter pixels) :

$$C = \frac{1}{N} \sum_{i=1}^N \phi_i \phi_i^T = AA^T$$

Où chacune des colonnes de A représente un vecteur de différences, soit :

$$A = [\phi_1 \phi_2 \dots \phi_N]$$

4- Etant donné les dimensions élevées de C ($LH \times LH$), une approche inter-images est privilégiée. Le calcul se limite donc à une matrice $L = A^T A$ dépendant du nombre d'images dans la banque d'apprentissage (c.-à-d. : dimension $N \times N$) ;

5- Calcul des valeurs et vecteurs propres de la matrice L ;

6- Le visage propre u_i associé à la $i^{\text{ème}}$ valeur propre est formé en utilisant les vecteurs propres v_i de la matrice L :

$$u_i = \sum_{k=1}^N v_{ik} \phi_k$$

7- Les M premiers vecteurs propres (EF) (c.-à-d. : associés aux plus fortes valeurs propres) sont conservés. Ils définissent ainsi l'espace des visages (face space) ;

8- Les images originales sont projetées dans l'espace des visages pour former une suite de coefficients d'appartenance, ce qui donne pour une l'image I_i :

$$\omega_k = u_k^T \phi_k \quad \text{où } k=1, \dots, M$$

9- Ces coefficients forment alors un vecteur représentant l'image I_i :

$$\Omega_i = [\omega_1, \omega_2, \dots, \omega_M]$$

Une fois l'apprentissage complété, les différentes représentations d'un individu peuvent être regroupées afin de former une classe. Ceci peut être réalisé notamment en calculant une moyenne des différents vecteurs Ω_i correspondants à la personne. Il est également possible de considérer les représentations individuellement et les utiliser directement avec l'algorithme du K-ppv.

Les figures 3.5 et 3.6 montrent la phase d'apprentissage.

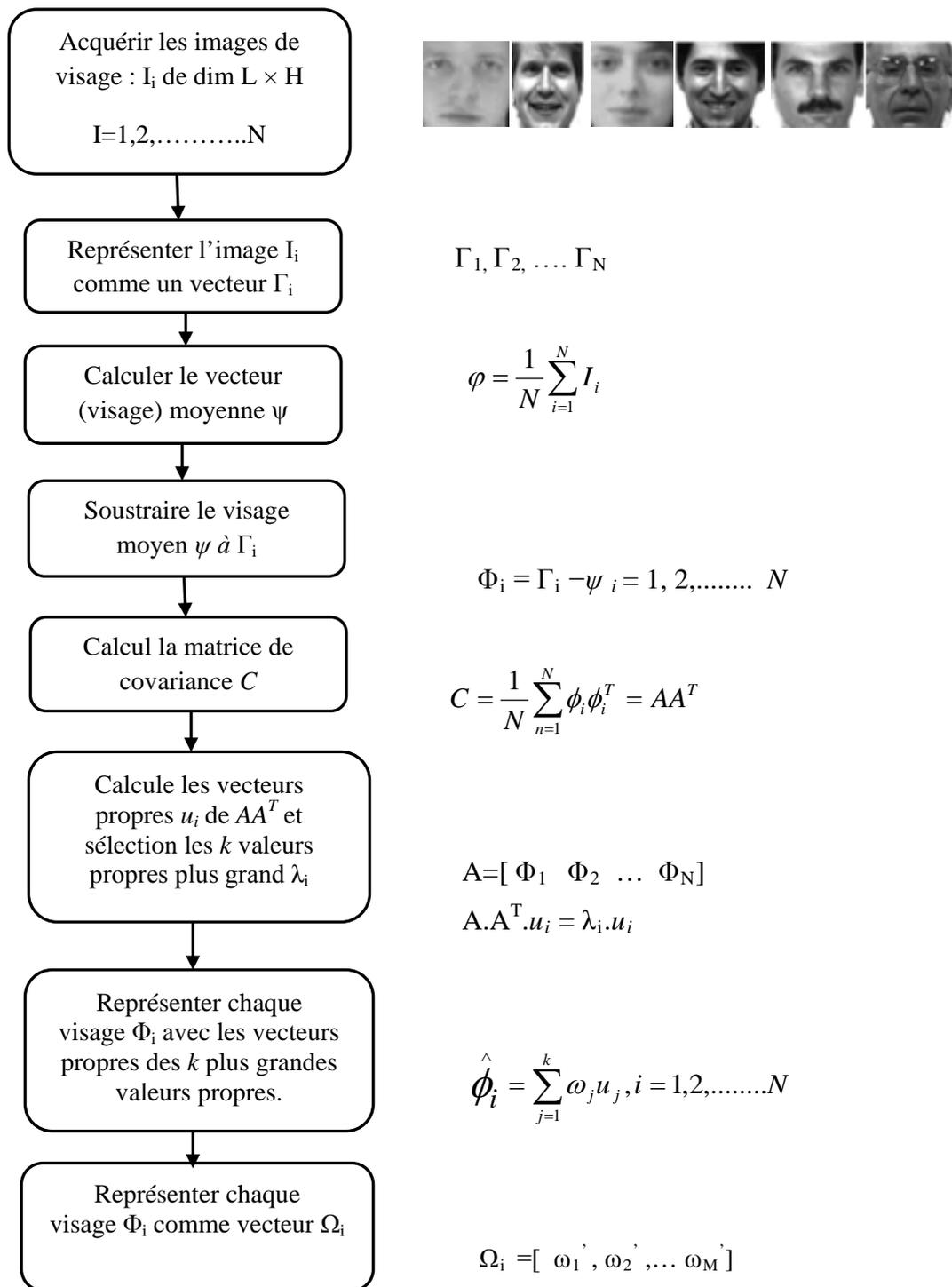


Figure 3.5: Organigramme de la phase d'apprentissage de ACP

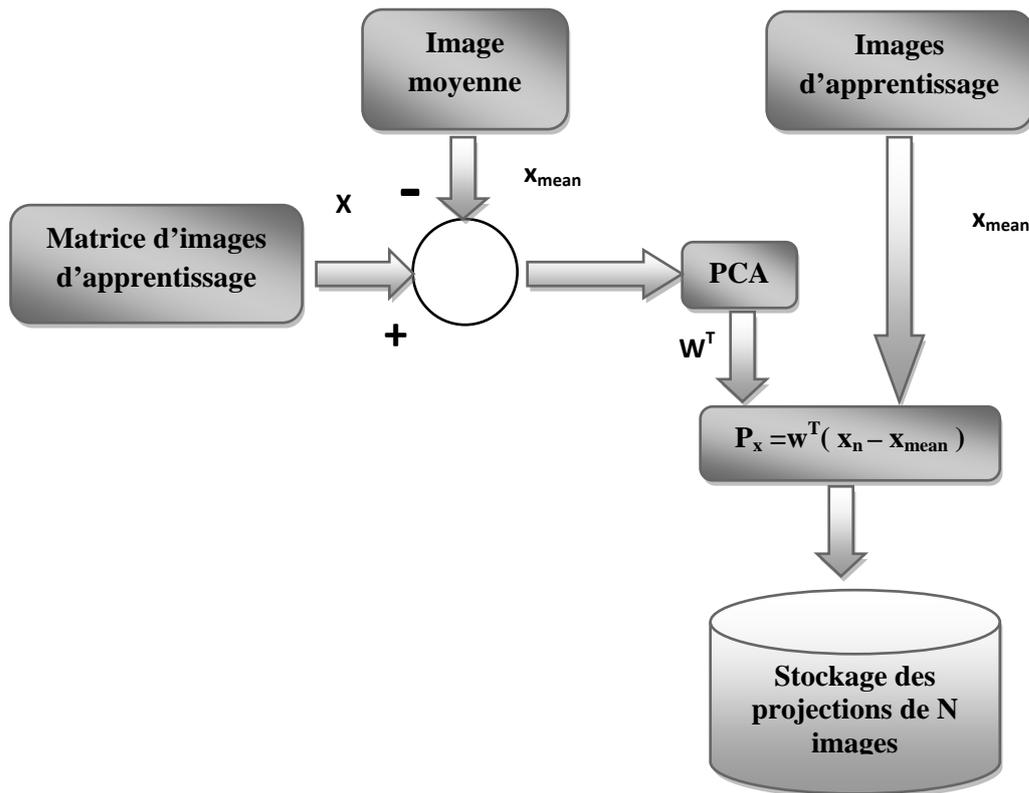


Figure 3.6: Phase d'apprentissage

- Phase de test (Reconnaissance):

Lorsqu'une nouvelle image de la base de test est mise devant le système, on la soustrait par rapport à la moyenne, on la projette ensuite sur l'espace vectoriel relatif à la matrice de projection W afin de la comparer avec toutes les projections issues de la phase d'apprentissage et qui étaient stockées dans la base de données. Par le terme comparer, il faut comprendre effectuer un calcul de distance entre les projections vectorielles. Il semble logique que plus la distance entre deux projections est petite, plus ces deux projections se ressemblent. Ainsi le résultat de la reconnaissance est l'image de la base d'apprentissage qui ressemble le plus à la nouvelle image présentée au système. [33]

L'algorithme de reconnaissance est:

Lorsqu'un visage est présenté au système, la procédure d'identification consiste à :

- 1- Projeter l'images d'entrée I dans l'espace des visages, ce qui engendre un coefficient d'appartenance ω_k à un Eigen face u_k en utilisant l'équation :

$$\omega_k = u_k^T \phi_k .$$

- 2- Les coefficients d'appartenance forment alors un vecteur de représentation de taille M ;
- 3- Ce dernier est comparé avec ceux obtenus lors de la phase d'apprentissage à l'aide de l'algorithme K-ppv en appliquant une métrique de distance particulière. Dans ce cas on utilise la distance Euclidienne.

Le schéma suivant montre la phase de reconnaissance

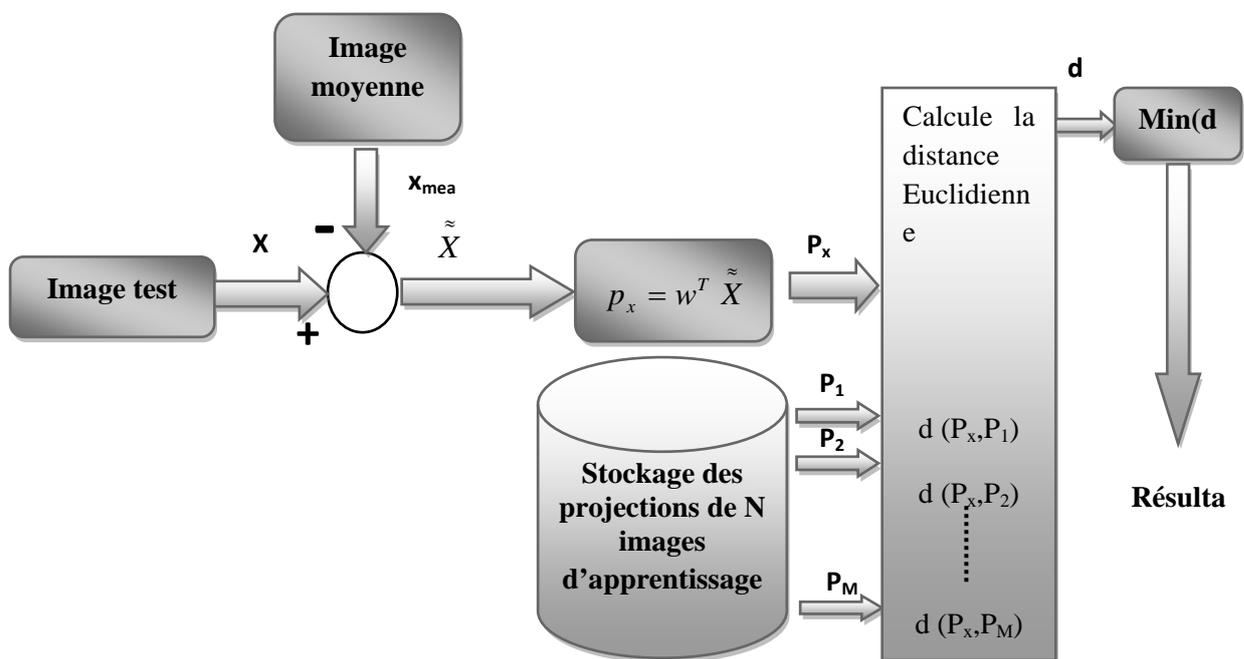


Figure 3.7: phase de test.

3.7 Résultats de simulation :

❖ Base d'image contient 40 personnes :

Pour tester les deux méthodes précédentes: en utilise une base d'image appelée ORL []. Cette base contient 40 personnes, où chaque personne est représentée par un ensemble de 10 images différentes de taille 48×48 .

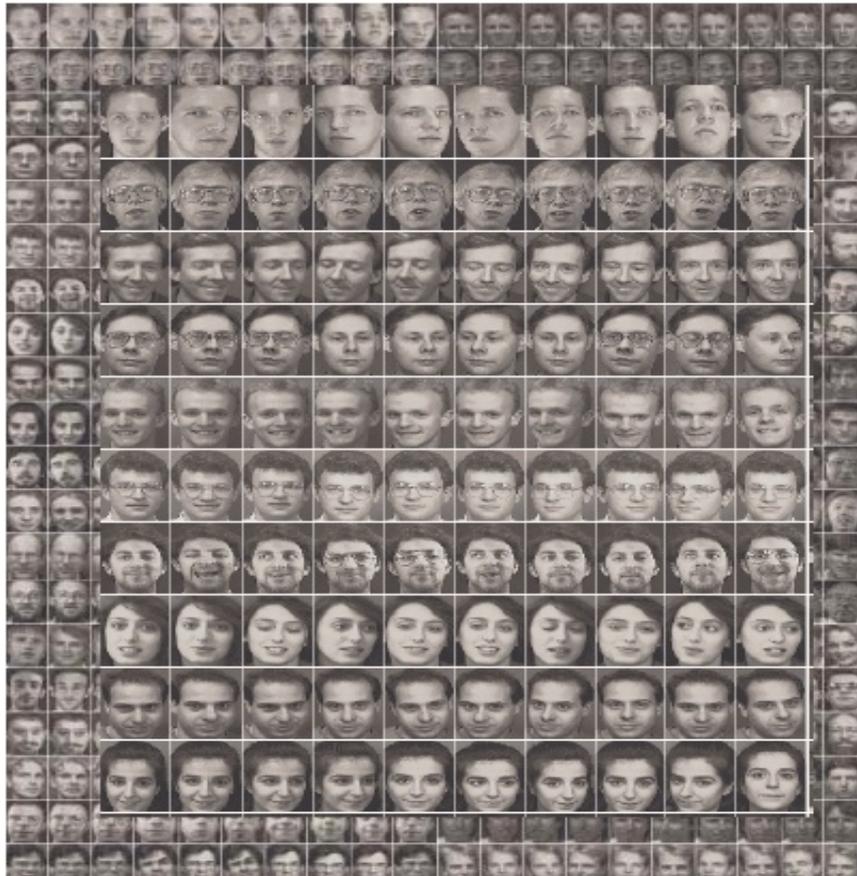


Figure 3.8: Base de donnée ORL

- Un ensemble représentant la base d'apprentissage contient 240 images répartie sur 40 personnes (c à d 6 exemples de chaque personne).
- Un ensemble représentant la base de test contient 160 images (4 images de chaque personne) et sera utilisé pour tester et comparer les différentes classes.

Méthode de la distance Euclidienne :

On calcule le taux de reconnaissance pour différentes dimensions d'image:

dimension	10×10	15×15	20×20	25×25	30×30
Taux de reconnaissance en %	91.8750	92.50	92.50	91.8750	90

Tableau 3.2: Variation du taux de reconnaissance en fonction de la dimension de l'image.

On obtient un taux de reconnaissance de 91.8750% avec les dimensions 10×10,25×25 et 92.50 % avec les dimensions 15×15,20 × 20 et 90 % avec dimension 30× 30

La variation du taux de reconnaissance avec la dimension d'image est représentée sur la figure suivante.

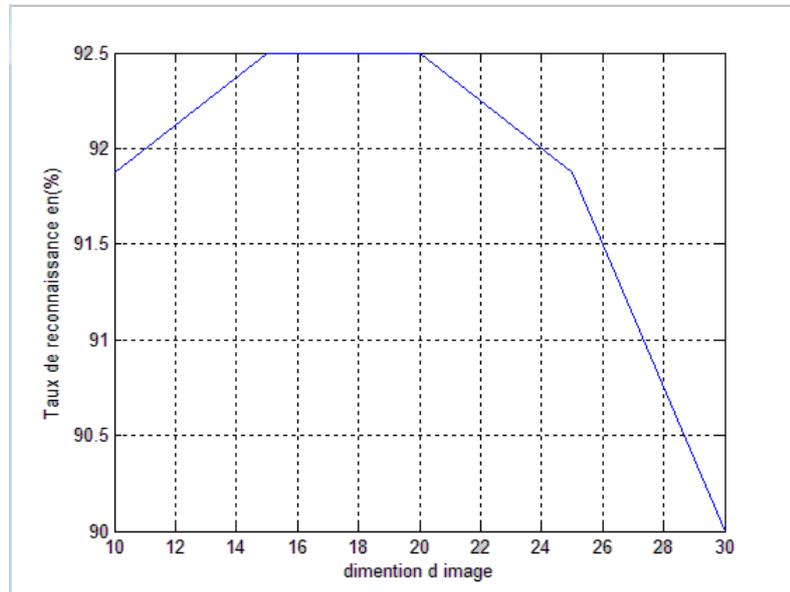


Figure 3.9: Variation du taux de reconnaissance en fonction de la dimension de l'image.

On remarque que le taux de reconnaissance ne s'améliore donc pas lorsque en augmente la dimension de l'image.

Méthode PCA:

On calcule le taux de reconnaissance pour différents nombres de vecteurs propres (Eigen_Face)

N ^{bre} d'Eigenface	10	20	30	40	50	60	70	80	90	180	200	225
T.R	86.2500	90.6250	90.6250	91	91.8750	91.8750	91.8750	91.8750	92.5000	92.5000	92.5000	92.5000

Tableau 3.3: Variation de taux de reconnaissance en fonction de N^{bre} d'Eigen Face.

On obtient un taux de reconnaissance égale 86.2500 % avec 10 Eigen_Face, le taux de reconnaissance augmente lorsque on augmente le nombre d'Eigen_Face jusqu'à 90 Eigen Face. D'après ces tests, le taux de reconnaissance prend une valeur fixe égale à 92.5000 % quelque soit le nombre des Eigen_Face.

La figure 3.10 illustre la variation du taux de reconnaissance avec le nombre d'Eigen_Face.

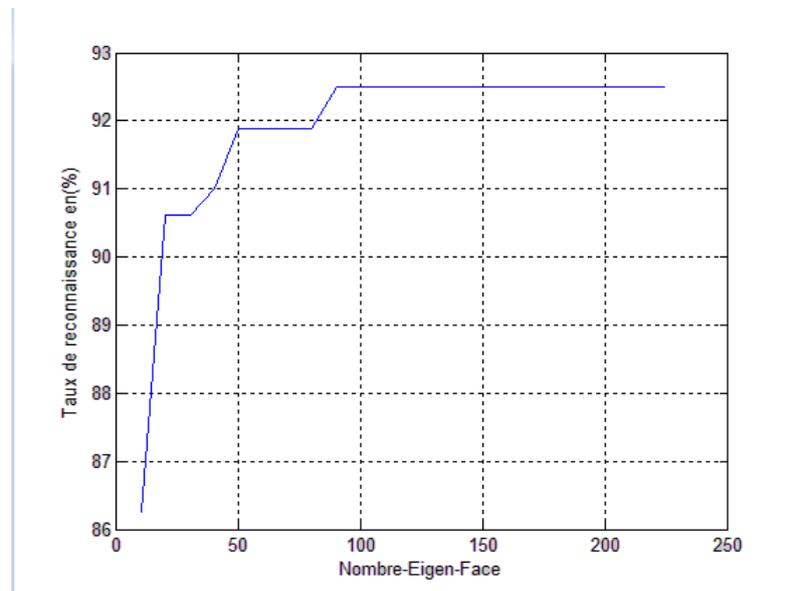


Figure 3.10: Variation du taux de reconnaissance en fonction du N^{bre} d'Eigen Face.

❖ **Base d'image contient 3 personnes :**

Pour tester les deux méthodes précédentes: en utilise une base d'image. Cette base contient 3 personnes, où chaque personne est représentée par un ensemble de 6 images différentes de taille 48×48 .

- Un ensemble représentant la base d'apprentissage contient 9 images répartie sur 3 personnes (c à d 3 exemples de chaque personne).
- Un ensemble représentant la base de test contient 9 images (3 images de chaque personne) et sera utilisé pour tester et comparer les différentes classes.

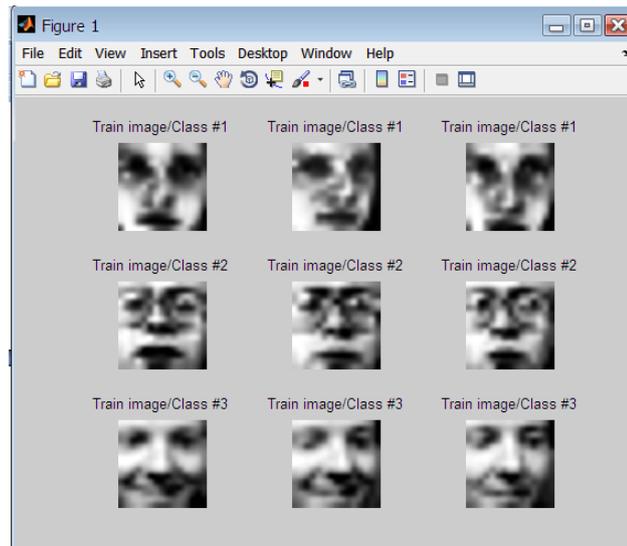


Figure 3.11: La base d'apprentissage contient 9 images.

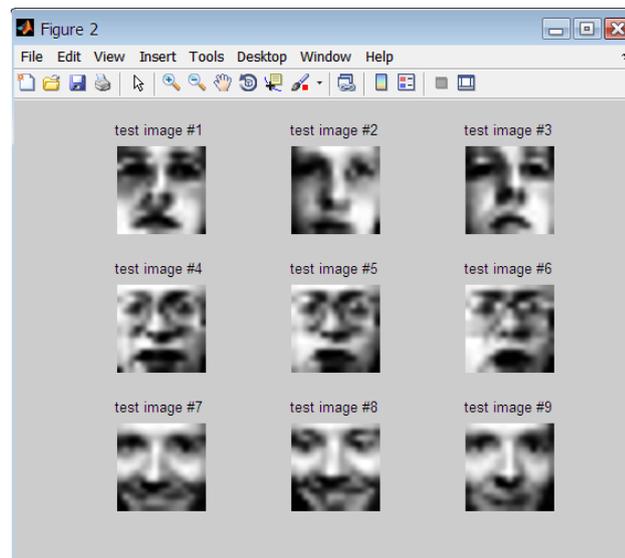


Figure 3.12: La base de test contient 9 images.

Méthode de la distance Euclidienne et Méthode PCA:

On calcule le taux de reconnaissance pour différentes dimensions d'image:

dimension	10×10	15×15	20×20	25×25	30×30
Taux de reconnaissance en %	100	100	100	100	100

Tableau 3.4: Variation du taux de reconnaissance en fonction de la dimension de l'image.

On obtient un taux de reconnaissance de 100% avec la dimension 10×10 , 15×15 et 20×20 , 25×25 et 30×30

❖ **Base d'image contient 5 personnes :**

Pour tester les deux méthodes précédentes: en utilise une base d'image. Cette base contient 5 personnes, où chaque personne est représentée par un ensemble de 10 images différentes de taille 48×48 .

- Un ensemble représentant la base d'apprentissage contient 25 images répartie sur 5 personnes (c à d 5 exemples de chaque personne).
- Un ensemble représentant la base de test contient 25 images (5 images de chaque personne) et sera utilisé pour tester et comparer les différentes classes.

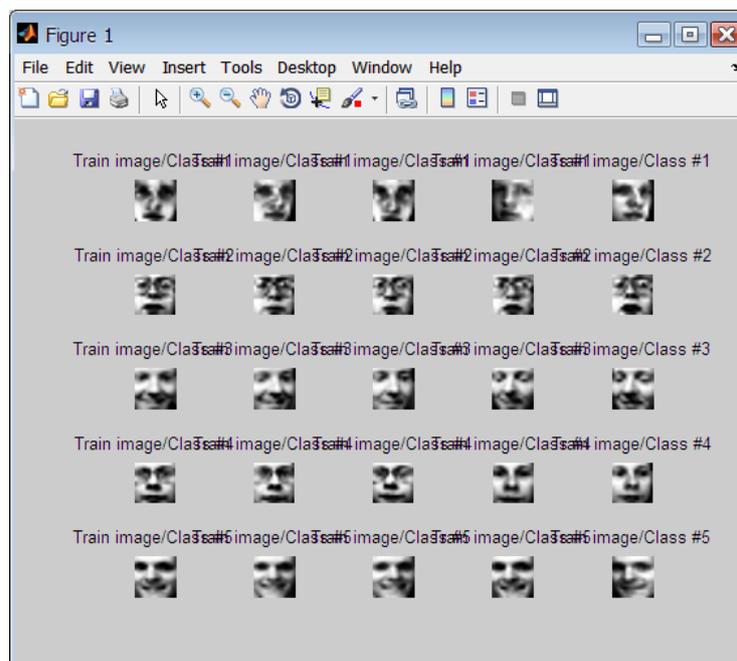


Figure 3.13: La base d'apprentissage contient 25 images

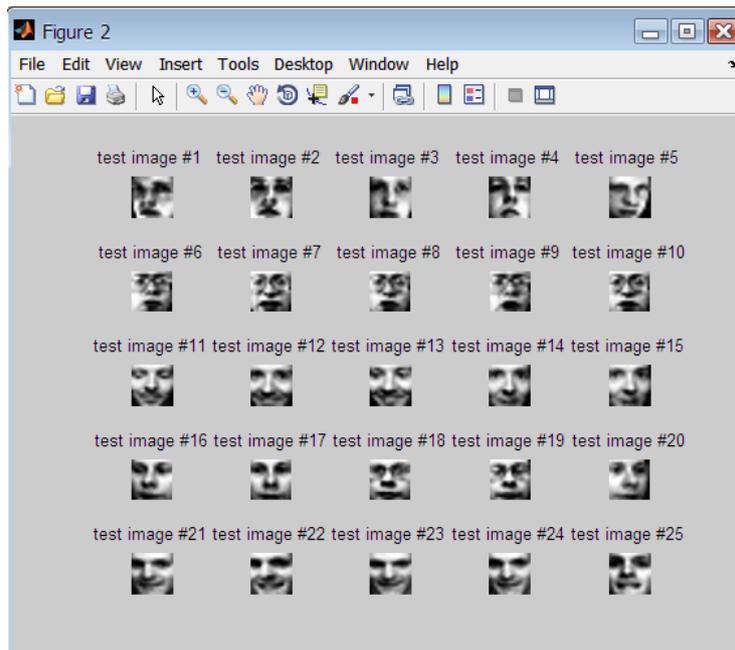


Figure 3.14: La base de test contient 25 images

Méthode de la distance Euclidienne et Méthode PCA:

On calcule le taux de reconnaissance pour différentes dimensions d'image:

dimension	10×10	15×15	20×20	25×25	30×30
Taux de reconnaissance en %	100	100	100	100	100

Tableau 3.5: Variation du taux de reconnaissance en fonction de la dimension de l'image.

On obtient un taux de reconnaissance de 100% avec la dimension 10×10,15×15 et 20×20,25 × 25 et 30× 30

3.8 CONCLUSION

Dans ce chapitre, nous avons présenté le processus générale de reconnaissance de visage, et les différentes méthodes de reconnaissance existants. Nous avons présenté une étude détaillée sur deux méthodes, la méthode de la distance euclidienne et la méthode ACP (Eigen faces). Les résultats que nous avons obtenus sont encourageants, dans le chapitre suivant nous essayons la reconnaissance de visages par réseau de neurones.

Reconnaissance de visage par Réseau de Neurones MLP

4.1 Introduction :

Depuis plusieurs années, les réseaux de neurones artificiels, et particulièrement les perceptrons multicouches, se sont montrés très efficaces dans le domaine de la reconnaissance statistique de forme. Nous allons maintenant voir la méthode de modélisation neuronale appliquée à l'identification de personnes, le but étant "simplement" de déterminer l'identité de la personne à partir d'une image de visage. Les réseaux de neurones artificiels sont capables d'apprendre l'association entrée/sortie, du fait de cette capacité, ils vont donc être utilisés pour résoudre ce type de problème. On utilise de type de réseaux : réseaux multicouches (MLP),.

4.2 Généralité sur les Réseaux de neurones Artificiels

4.2.1. Introduction et Historique

Les réseaux de neurones artificiels, nés il y a environ une cinquantaine d'années, sont toujours en cours de développement. Cet axe de recherche a tiré l'attention de beaucoup de chercheurs de différentes disciplines. Ces réseaux de neurones artificiels (RNA) sont issus de la combinaison entre la neurobiologie comme idée de base, la physique comme champs d'application et les mathématiques avec l'informatique en tant que moyens de réalisation. Nous présentons dans ce qui suit un bref historique des réseaux de neurone artificiels.

1943: J. McCulloch & W. Pitts [34]

- Proposent un modèle simple de neurone capable de reproduire la machine de Turing,

- Démontrent qu'un assemblage synchrone de tels neurones est une machine universelle de calcul (c'est-à-dire que toute fonction logique peut être représentée par des unités à seuil),

1948: D. Hebb

- Propose une règle d'apprentissage pour des réseaux de neurones. *The Organization of Behaviour*, Wiley, New York 1949.

1958: F. Rosenblatt

- Propose le modèle du perceptron et démontre son théorème de convergence.

1969: M. Minsky & S. Papert [35]

- Démontrent les limitations du modèle du perceptron. *Perceptrons*, the MIT Press, Cambridge 1969.

1972 :T. Kohonen

- Présente ses travaux sur les mémoires associatives.
- Propose des applications à la reconnaissance de formes.

1982: J.J. Hopfield [36]

- Présente une théorie de fonctionnement et des possibilités des réseaux de neurones. *Neural Networks and Physical Systems with Emergent Collective Computational Abilities*, *Proceedings of the National Academy of Sciences, USA*, 1982, pp. 2,554.

1986: D.E. Rumelhart et al [37]

- Apprentissage par rétro-propagation pour les réseaux multicouches.

Learning representations by back-propagating errors, *Nature*, vol. 323 (1986).

4.2. 2. Du neurone biologique au modèle mathématique

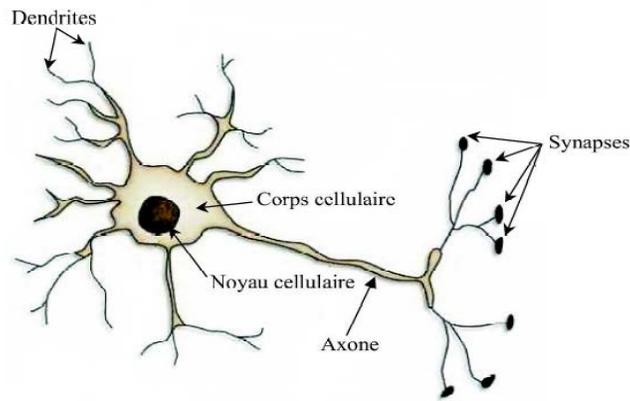


Figure .4. 1 : Représentation schématique d'un neurone biologique.

L'architecture biologique du neurone est présentée sur la **Figure 4.1**. Les dendrites sont considérées comme les portes d'entrées des neurones. Elles sont les réceptrices principales du neurone permettant de capter les signaux lui parvenant. Ensuite, un traitement de toutes ces informations a lieu au niveau du corps cellulaire. Les principales lignes de transmission du système nerveux sont les axones. Ils servent de moyen de transport pour les signaux émis par le neurone, vers d'autres neurones. L'information va transiter de neurone en neurone le long de l'axone jusqu'aux muscles via un potentiel d'action (un signal électrique). La synapse désigne les zones de contact entre neurones, zones spécialisées dans la transmission de l'information. Elles peuvent être inhibitrice ou excitatrice de signal pondérant l'information arrivant au corps cellulaire par dendrites.

Par analogie avec la biologie, un neurone artificiel **Figure. 4.2**. Reçoit l'information provenant des entrées x_i ($i = 1, 2, \dots, n$) par l'intermédiaire des connexions dont on affecte à chacune d'elles un poids ω (en anglais weights) pondérant l'information reçue, appelée aussi poids synaptique. Il fonctionne en deux étapes :

- La première phase représente le prétraitement des données reçues en calculant le potentiel v_j des neurones j par la fonction suivante :

$$v_j = b_j + \sum_{i=1}^n \omega_{j,i} x_i \quad (4.1)$$

Où $\omega_{j,i}$ désigne le poids de la connexion liant le neurone j à l'entrée i , et b_j est un terme constant appelé *biais* qui peut être considéré comme le poids d'une entrée x_0 égale à 1. Ainsi, la relation 4.1 s'écrit plus simplement :

$$v_j = \sum_{i=0}^n \omega_{j,i} x_i \quad (4.2)$$

Dans la deuxième phase, une fonction de transfert g appelée également *fonction d'activation* calcule la valeur de l'état interne s_j du neurone j à partir de la valeur du potentiel v_j . Cette valeur désignera la sortie du neurone :

$$s_j = g(v_j) = g\left(\sum_{i=0}^n \omega_{j,i} x_i\right) \quad (4.3)$$

Ainsi, le neurone peut être défini mathématiquement comme étant une fonction algébrique, a priori non linéaire (suivant g) et bornée, des entrées x_i et paramétrée par les poids $\omega_{j,i}$ [38]-[39].

$$s_j = g(x_i, \omega_{j,i}) = g_{\omega_{j,i}}(x_i) \quad (i = 1, 2, \dots, n) \quad (4.4)$$

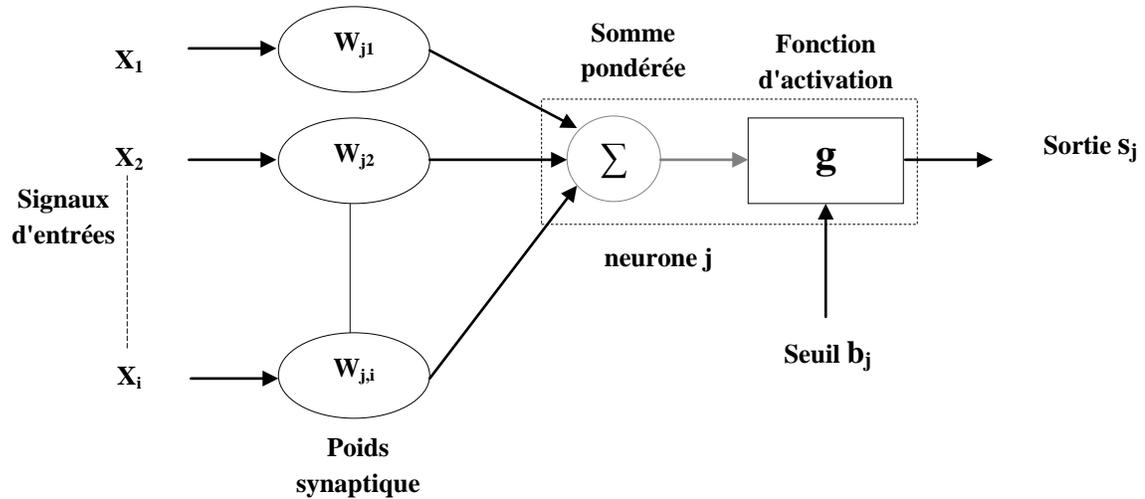


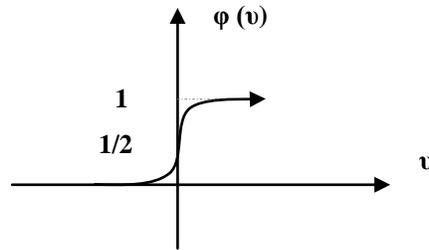
Figure .4.2 : Représentation d'un neurone mathématique. Les variables x_i désignent les entrées, la sortie est notée s_j , le biais b_j et les poids synaptiques sont désignés par $\omega_{j,i}$.

Plusieurs possibilités existent pour spécifier la nature de la fonction d'activation du modèle. Les fonctions mathématiques les plus souvent utilisées sont:

- la fonction sigmoïde ;
- fonction gaussienne;
- fonction à seuil (fonction échelon) ;
- fonctions linéaires par parties.

A. La fonction sigmoïde

La fonction sigmoïde, définie par la formule (4.5) et représentée par le graphe (Graphe : 4.3) Possède les propriétés importantes évoquées précédemment (elle n'est pas polynomiale et est indéfiniment continûment dérivable).

**Graphe. 4.3.**

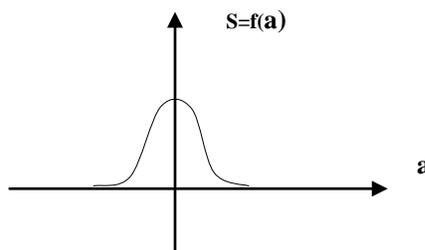
$$\varphi(v) = 1 / (1 + e^{-v}) \quad (4.5)$$

En outre, une propriété simple permet d'accélérer le calcul de sa dérivée, ce qui réduit le temps calcul nécessaire à l'apprentissage d'un réseau de neurones. On a en effet

$$\frac{d\varphi(v)}{dv} = \varphi(v) \cdot (1 - \varphi(v)) \quad (4.6)$$

On peut donc calculer la dérivée de cette fonction en un point de façon très efficace à partir de sa valeur en ce point. De plus, la fonction sigmoïde est à valeurs dans l'intervalle $[0,1]$, ce qui permet d'interpréter la sortie du neurone comme une probabilité. Elle est aussi liée au modèle de régression logistique et apparaît naturellement quand on considère le problème de la séparation optimale de deux classes de distributions gaussiennes avec la même matrice de covariance.

B. La fonction gaussienne

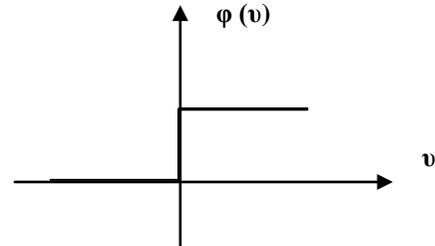
**Graphe. 4.4.**

$$s = \exp(-a^2 / \sigma) \quad (4.7)$$

C. La fonction à seuil

Certains auteurs utilisent les neurones à sortie binaire. La fonction d'activation est alors la fonction signe (-1 ou +1) ou la fonction échelon (0 ou +1) (Graphe: 4.5).

$$\varphi(v_k) = \begin{cases} 1 & \text{si } v_k \geq 0 \\ 0 & \text{si } v_k < 0 \end{cases} \quad (4.8)$$

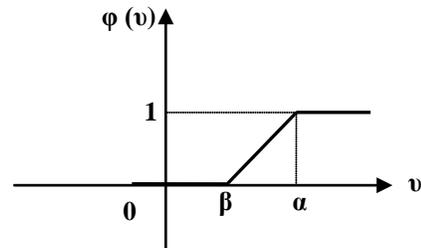


Graphe. 4.5.

D. La fonction linéaire par parties

Ses modèles peuvent être décrits par une partie linéaire à pente constante et une valeur de saturation (Graphe : 4.6)

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq \alpha \\ v & \text{si } \alpha > v > \beta \\ 0 & \text{si } v < \beta \end{cases} \quad (4.9)$$



Graphe. 4.6.

4.2. 3. Architecture des réseaux de neurones

Un neurone réalise une fonction non linéaire. L'intérêt des neurones réside dans les propriétés qui résultent de leurs associations en réseaux. C'est-à-dire de la composition des fonctions non linéaires réalisées par chacun des neurones constitutifs.

La plupart des modèles de réseaux de neurones artificiels (RNA) ne retirent du fonctionnement réel que les principes suivants :

- Ceux caractérisés par des interconnexions denses entre des unités de traitement simples agissant en parallèle.
- Ceux qu'à chaque connexion est associé un poids qui détermine l'influence réciproque des deux unités connectées.
- Ceux pour lesquels les poids de connexions sont modifiables et c'est cette plasticité qui donne lieu aux facultés d'adaptation et d'apprentissage.

Les neurones sont donc fortement connectés entre eux. Mais ces connexions ne sont pas aléatoires. Elles correspondent à des réseaux dont les architectures sont assez connues.

4.2. 3.1. Connectivité

La connectivité des réseaux, c'est à dire la manière dont ils sont reliés, peut être totale (tous les neurones connectés entre eux) ou par couche (les neurones d'une couche ne sont connectés qu'aux couches suivantes) ; il existe des réseaux monocouches ou multicouches (multi layer networks). Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle [39], [40].

Selon le type d'interconnexion on peut classer les réseaux comme réseaux à connexion locale (partielle) ou totale (complète).

- Réseaux à connexions locales (Figure : 4.7) : Chaque neurone entretient des relations avec un nombre réduit et localisé de neurones de la couche en aval.

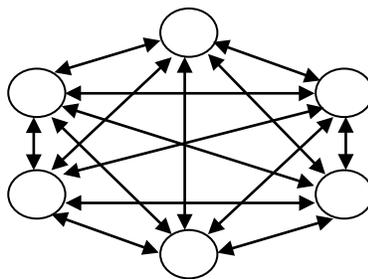


Figure. 4.7: Réseau à connexions locales.

b. Réseaux à connexions totales (figure : 4.8) : Chaque neurone est connecté à tous les neurones du réseau.

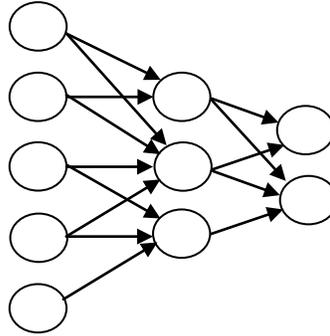


Figure. 4.8 : Réseau à connexions totales.

4.2 .4. Apprentissage

Une caractéristique des réseaux de neurones est leur capacité à apprendre (par exemple à reconnaître une lettre, un son...). Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple [40].

On appelle « phase d'apprentissage » des réseaux de neurones la procédure qui consiste à estimer les paramètres des neurones du réseau. Afin que celui-ci remplisse au mieux la tâche qui lui est affectée. La phase de l'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré.

4.2 .4.1. Types d'apprentissage

Selon le critère d'apprentissage les réseaux de neurones se subdivisent en deux classes, les réseaux à apprentissage supervisé (Supervised learning) et les réseaux à apprentissage non supervisé (Unsupervised Learning) [41], [37].

- Pour les réseaux à apprentissage supervisé, on présente au réseau des entrées et au même temps les sorties que l'on désirerait pour cette entrée. Le réseau doit alors se reconfigurer. C'est-à-dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie désirée.

- Pour les réseaux à apprentissage non supervisé, on présente une entrée au réseau et on le laisse évoluer librement jusqu'à ce qu'il se stabilise.

Il existe aussi des réseaux à apprentissage dit semi supervisé (renforcement Learning) qui ne tiennent compte que d'une évaluation partielle ou qualitative des sorties.

4.2 .4.2. Application des réseaux de neurones selon le mode d'apprentissage

Les réseaux à apprentissage supervisé servent à :

- La modélisation statique des processus non linéaires décrits par des équations algébriques.
- La modélisation dynamique des processus décrits par des équations récurrentes.
- La commande de processus. Ils servent aussi à imposer une dynamique de réponse à une sollicitation.

Les réseaux à apprentissage non supervisé permettent l'analyse, la classification et la visualisation des données.

4.2 .5. Les types de réseaux

Par définition, un réseau de neurones est constitué d'un assemblage d'éléments, d'unités ou de nœuds processeurs pour lequel un sous-groupe effectue un traitement indépendant et transmet le résultat à un deuxième sous-groupe et ainsi de suite (cas d'un réseau à couches multiples). Les capacités de traitement du réseau dépendent des poids w_i auxquels sont affectées des valeurs produisant un filtre affectant la capacité d'apprentissage du réseau. Dans un réseau de neurones, les neurones sont regroupés en couches. Habituellement, chaque neurone dans une couche est connecté à tous les neurones dans la couche précédente et la couche suivante (excepté dans la couche d'entrée et celle de sortie du réseau). L'information donnée à un réseau de neurones est

propagée couche par couche de la couche d'entrée à la couche de sortie en passant par une ou plusieurs couches intermédiaires (couches cachées).

4.2 .5.1. Le perceptron

Le perceptron est un réseau de neurones très simple. Il est linéaire et monocouche, acceptant uniquement des valeurs d'entrées et de sorties binaires (0 ou 1). Il est inventé par le psychologue *Frank Rosenblatt* et inspiré du système visuel. Le procédé d'apprentissage est supervisé et le réseau est capable de résoudre des opérations logiques simples comme les opérations "ET logique" ou "OU logique". Ces réseaux sont efficaces pourvu que les données traitées soient des vecteurs orthogonaux ou linéairement indépendants. Ce que l'on entend par méthode d'apprentissage supervisé est que les sorties du réseau sont comparées avec des sorties cibles. Ce type de réseau est aussi utilisé pour la classification. Les opérations logiques plus compliquées comme un problème "XOR" (OU Exclusif) ne peuvent pas être résolues par un perceptron. L'algorithme d'apprentissage est la règle d'apprentissage de *Hebb* qui affecte le changement des poids en multipliant l'entrée d'un neurone par sa sortie et le taux d'apprentissage du réseau .

Un réseau "feedforward" à une seule couche (perceptron) peut uniquement représenter des fonctions linéairement séparables. C'est-à-dire celles pour lesquelles la surface de décision séparant les cas positifs des cas négatifs est un (hyper-)plan.

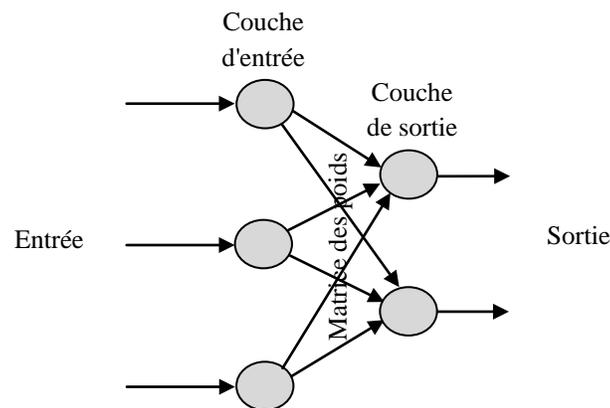


Figure. 4.9 : Architecture du perceptron.

La sortie du perceptron dépend de la somme des composantes x_i du vecteur d'entrée, pondérées par des poids réels w_i . Conformément aux notations utilisées dans ce travail, nous appellerons un potentiel, la somme pondérée notée par 'net', suivante.

$$net_k = \sum_{i=0}^n W_{ki} X_i \quad (4.10)$$

La sortie du réseau sera désignée par 'O'. Elle sera de la forme suivante :

$$O_k = \begin{cases} 1 & \text{si } net_k \geq \theta_k \\ 0 & \text{si } net_k < \theta_k \end{cases} \quad (4.11)$$

w_{ji} : Coefficients synaptiques de j vers i, x_i : sortie(ou état) de j, et θ_k : seuil

4.2 .5.2. Perceptron multicouche

Les réseaux multicouches sont actuellement les plus employés. Plusieurs couches de traitement leur permettent de réaliser des associations non linéaires entre l'entrée et la sortie. Le perceptron à niveaux multiples (perceptron multi couche'PMC' ou 'MLP' en anglais) est une extension du perceptron qui a une ou plusieurs sous-couches cachées entre ses couches d'entrée et de sortie.

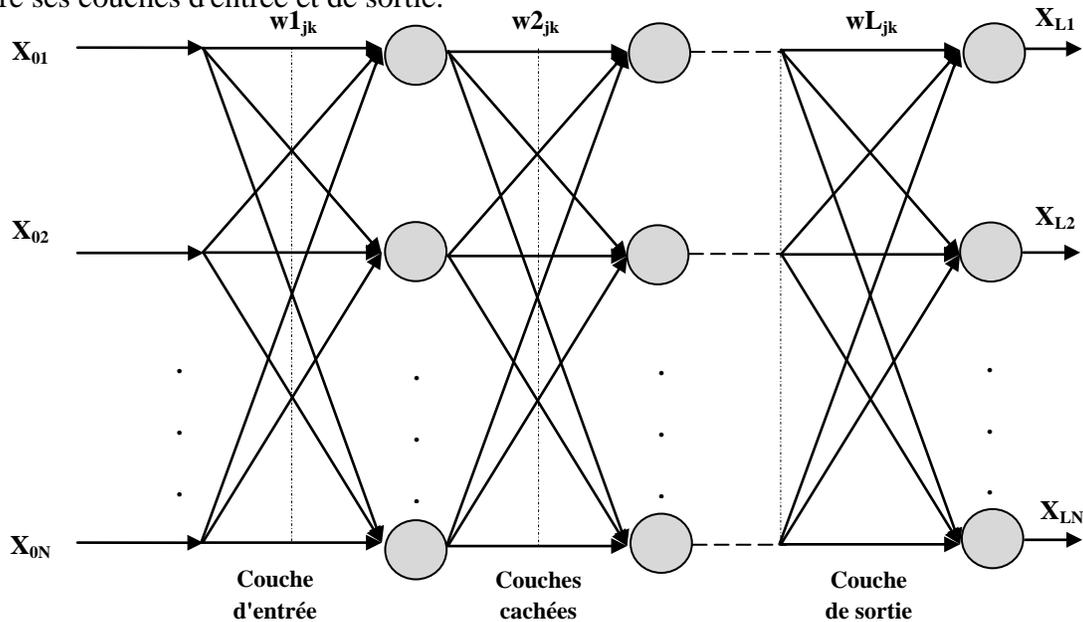


Figure. 4.10 : Architecture du perceptron multicouche.

Le réseau MLP contient N_0 entrées, $L-1$ couche cachées de N_i neurones chacune ($1 < i < L-1$) et une couche de sortie contenant N_L neurones. L'indice de la couche est i , X_{ik} est la sortie du neurone k de la couche i , w_{ijk} est le poids reliant la sortie x_{i-1j} au neurone k de la couche i , N_i est le nombre de neurones dans la couche i . Avec ces notions, la sortie X_{ik} du neurone (i, k) sera donnée par :

$$y_{ik} = \sum_{j=1}^{N(i-1)} w_{ijk} x_{i-1j} + b_{ik} \quad (4.12)$$

$$X_{ik} = f(y_{ik})$$

4.2 .5.2.1. Apprentissage du perceptron multicouche

Pour la couche de sortie, on peut appliquer l'apprentissage du perceptron, mais comment modifier les poids pour les connexions qui ne sont pas en relation avec un neurone de sortie? Le problème est l'obtention d'une estimation de la valeur désirée pour chaque neurone de la couche cachée. Ils utilisent pour modifier leurs poids, un algorithme de rétro-propagation du gradient (back-propagation). Il s'agit toujours de minimiser l'erreur quadratique.

4.2 .5.2.2. Réseau de rétro-propagation

Le réseau de rétro-propagation est un réseau à plusieurs couches constitué au minimum de trois couches (Figure 4-10), une couche d'entrée, une couche de sortie et une ou plusieurs couches cachées. Chaque couche contient un ou plusieurs neurones. Cela dépend du nombre de données qu'on désire faire apprendre au réseau, et du nombre de sortie qu'on désire avoir. Il n'existe pas une méthode bien précise pour déterminer le nombre de neurones dans la couche cachée, car il dépend de la complexité du problème.

4.2 .5.2.3. Algorithme de rétro propagation

C'est un algorithme supervisé dont l'objectif est d'ajuster les poids du réseau MLP de façon à minimiser une fonction de coût différentiable, telle que l'erreur quadratique entre la sortie du réseau et la sortie désirée :

$$E(n) = \|d(n) - Y_L(n)\|^2 \quad (4.13)$$

Où $Y_L(n)$ est la sortie du réseau au temps n et $d(n)$ la sortie désirée. L'algorithme de BP assure la descente du gradient sur le critère d'erreur pour atteindre le minimum. Le gradient de $E(n)$ est calculé pour tous les poids de la manière suivante :

$$\frac{\partial E(n)}{\partial w_{ijk}(n)} = \frac{\partial E(n)}{\partial y_{ik}} \cdot \frac{\partial y_{ik}}{\partial w_{ijk}} = \frac{\partial E(n)}{\partial y_{ik}} x_{i-1j} \quad (4.14)$$

Dans le cas de la couche de sortie ($i = L$) on évalue le terme d'erreur de sortie

δ_{Lk} comme suite :

$$\delta_{Lk} = \frac{\partial E(n)}{\partial y_{Lk}} = 2f'(y_{Lk})(d_k - x_{Lk}) \quad (4.15)$$

Où $f'(\cdot)$ est la dérivée de la fonction d'activation :

$$f'(x) = \frac{df(x)}{dx}.$$

Pour les couches cachées, le terme d'erreur δ_{ik} du neurone (i, k) est donnée par :

$$\delta_{ik} = f'(y_{ik}) \sum_{j=1}^{N(j+1)} \delta_{i+1j} w_{i+1kj} \quad (4.16)$$

La modification des poids et biais est obtenue selon les équations suivantes :

$$\begin{aligned} w_{ijk}(n+1) &= w_{ijk}(n) + \eta \delta_{ik} x_{i-1j} + \Omega (w_{ijk}(n) - w_{ijk}(n-1)) \\ b_{ik}(n+1) &= b_{ik}(n) + \eta \delta_{ik} \end{aligned} \quad (4.17)$$

η est le pas d'apprentissage et Ω le terme d'inertie (momentum).

4.2 .5.2.4. Les paramètres d'apprentissage

Tous les modèles de réseaux de neurones requièrent un apprentissage. Plusieurs types d'apprentissages peuvent être adaptés à un même type de réseau de neurones. Les critères de choix sont souvent la rapidité de convergence ou les performances de généralisation. L'efficacité d'apprentissage dépend de plusieurs paramètres:

- **Taux d'apprentissage η**

Ce paramètre détermine la vitesse de convergence. Si la valeur de démarrage de η est grande, alors on aura un apprentissage très rapide mais au prix de la création d'oscillations dans l'erreur totale moyenne qui empêcheront l'algorithme de converger vers le minimum désiré. Le réseau devient instable. Dans la plupart des cas, si la fonction d'erreur possède plusieurs minimums locaux, le réseau subira un blocage dans l'un d'eux. Toutes ces conditions nous obligent à commencer l'apprentissage avec une petite valeur de η , si on veut atteindre un minimum global, même si l'apprentissage est long.

- **Momentum Ω :**

Ce paramètre permet l'introduction de l'ancien poids dans le calcul de sa nouvelle valeur. Le momentum empêche l'algorithme de rester bloqué dans un minimum local et il augmente encore la vitesse de convergence. La valeur de Ω est toujours inférieure à 1 (valeur typique $\Omega = 0.9$) [42].

- **Seuil de tolérance:**

Ce paramètre critique détermine la précision dans la réponse du réseau de neurones.

La phase d'apprentissage est souvent arrêtée lorsque l'erreur calculée sur l'ensemble de la base d'apprentissage est inférieure à un seuil déterminé par l'utilisateur. Il est possible d'arrêter l'apprentissage en fixant une limite au nombre d'itérations.

4.2 .5.3. Le réseau à fonction de base radiale

Le réseau à fonction de base radiale RBF (Radial Basis Function) est basé sur une architecture qui s'organise en deux couches seulement ; une couche cachée et une couche de sortie comme le montre la Figure 4.11. [43]

La couche cachée, constituée des noyaux (ou neurones) RBF effectue une transformation non linéaire de l'espace d'entrée. La couche de sortie calcule une combinaison linéaire des sorties de la couche cachée. Chaque noyau élémentaire calcule la distance entre l'entrée et son centre qu'il passe ensuite dans une non linéarité concrétisée par une fonction d'activation $\varphi (\cdot)$ qui est généralement de type gaussienne [44].

La valeur que prend la sortie du noyau gaussien est d'autant plus importante que l'entrée est plus proche de son centre et tend vers zéro, lorsque la distance entrée centre devient importante. La sortie du réseau RBF est donnée par :

$$y_i = \sum_{k=1}^{N_1} W_{kj} \varphi_k (\|x - c_k\|) \quad \text{et} \quad \varphi(\xi) = \exp\left(-\frac{\xi^2}{2\eta^2}\right) \quad (4.18)$$

$\|\cdot\|$ dénote la norme euclidienne, x le vecteur d'entrée, c_k est le centre associé au noyau k .

N_1 le nombre de noyaux de la couche cachée et W_{kj} les poids associés à la couche de sortie. Le paramètre ' η ' permet de contrôler la vitesse de décroissance de la fonction φ .

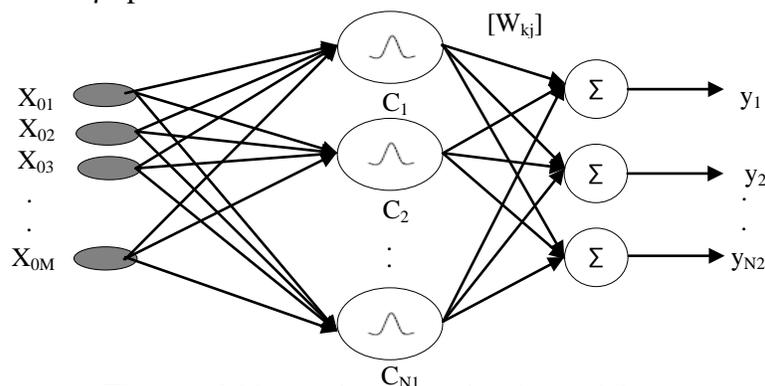


Figure. 4.11 : Architecture du réseau RBF.

L'utilisation habituelle des RBF conserve une fonction d'activation linéaire en sortie mais l'utilisation d'une fonction non linéaire reste possible [44], [45]. Le réseau RBF possède aussi la propriété d'approximation c'est à dire que le réseau est capable d'approximer n'importe quelle fonction continue avec une précision donnée pourvu qu'on fournisse un nombre de neurones suffisant [43], [46]. L'apprentissage des réseaux RBF est composé d'une règle non supervisée pour l'adaptation des centres et une règle d'apprentissage supervisé pour l'adaptation des poids. Le choix de la position des centres et le nombre de neurones reste généralement arbitraire.

4.3 Reconnaissance de visages : approche Mono- Réseau MLP

4.3.1 Base utilisée

Cette méthode repose sur l'apprentissage d'un réseau de neurones discriminant. On utilise un seul réseau à qui on donne en entrée tous les pixels de l'image test et on aura en sortie la personne correspondante. La base utilisée est ORL comprend 40 individus [47], et avec pour chaque personne Il y a 10 images différentes. On choisit, pour chaque personne, une base d'apprentissage de 6 images et une base de test de 4 images. Ces images seront prises aléatoirement parmi toutes les images disponibles. Il faut normaliser les valeurs des pixels avant de regrouper par base dans des fichiers différents grâce au programme de construction.

4.3.2 Structure du réseau neuronal:

Le réseau de neurones utilisé est un réseau MLP, à une seule couche cachée. L'entrée est composée de tous les pixels de l'image. En sortie du réseau, on a autant de neurones que de personnes à différencier, soit 40 sorties. Chaque neurone utilise comme fonction de transfert une sigmoïde. Ainsi, lorsqu'une personne est reconnue, le neurone correspondant sur la dernière couche aura une valeur en sortie proche de 1, et les autres neurones auront une valeur proche de 0. L'apprentissage des poids des neurones du réseau se fera avec l'algorithme de rétro propagation du gradient.

Les paramètres qui pourront être modifiés pour un meilleur apprentissage sont le nombre de neurones de la couche cachée, et le nombre d'itération d'apprentissage maximum.

Les valeurs des entrées sont simplement l'intensité du pixel en niveaux de gris (de 0 à 255), normalisés à des valeurs réels entre 0 et 1.

4.3.3 Reconnaissance pour un ensemble autorisé:

L'évaluation de tout système de reconnaissance revient à déterminer, par un test le taux de reconnaissance, ce taux de reconnaissance est la probabilité avec laquelle une personne est identifiée, si elle existe bien sur dans la base des visages. Une fois le réseau calculé, il faut procéder à des tests pour vérifier que le réseau réagit bien comme on le souhaite

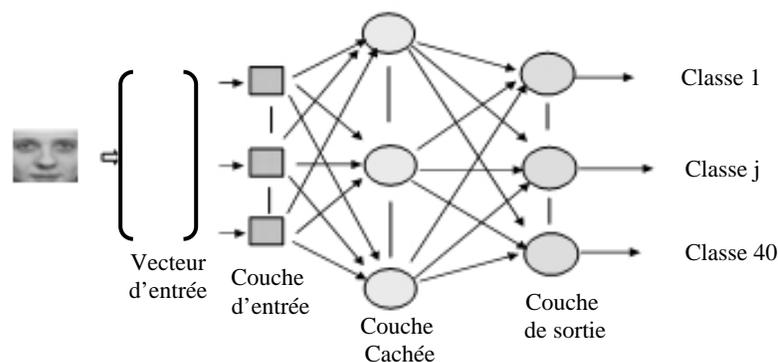


Figure 4.12: Modèle de reconnaissance mono réseau

Dans cette expérience, on va explorer plusieurs aspects du système de reconnaissance, et on va tester l'influence des différents paramètres qui entrent en jeu sur le résultat final.

Les paramètres investis sont notamment :

- L'influence du nombre de neurones dans la couche cachée,
- L'influence de la dimension de l'image.
- L'influence du seuil de classification.

Les paramètres d'apprentissages utilisés dans cette expérience sont :

Pas d'apprentissage : 0.2

L'erreur : epsilon=0.01

Nombre d'itération maximale : plusieurs itérations (très long)

Le temps d'apprentissage est très long (calcul en dizaines d'heures)

Le seuil de reconnaissance : 0.5 (associe la valeur 1 pour les réponses de neurone qui dépasse ce seuil et 0 sinon).

On répète la procédure d'apprentissage avec différents nombres de neurones dans la couche cachée, et différentes dimensions d'images, les taux de reconnaissance obtenus sont regroupés dans le tableau 4.1.

Neurones cachées	Taux de reconnaissance en (%)				
	10×10	15×15	20×20	25×25	30×30
30	82.5000	85	84.3750	82	84.8750
40	85.6250	86.2500	88.7500	84.3750	86.8750
50	88.7500	86.2500	88.7500	86.2500	86.8750
60	87.500	88.1250	86.8750	85	87.5000
70	86.25	87.5000	86.8750	83.7500	86.8750
80	90	91.6250	89.3750	85	90
90	88.1250	88.1250	85	89.3750	88.1250

Tableau 4.1 : Taux de reconnaissance en fonction de dimension de l'image, et nombre de neurones cachés.

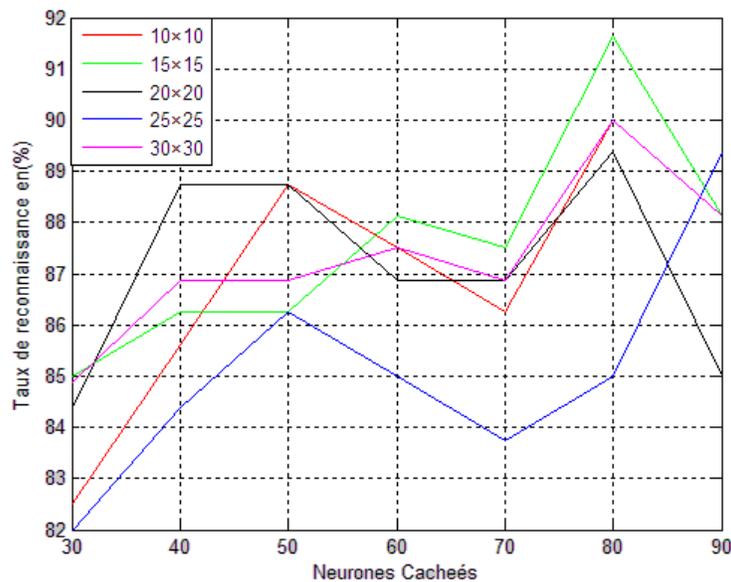


Figure 4.13: variation de taux de reconnaissance en fonction de nombre de neurones dans la couche cachée.

La figure 4.13 illustre la variation du taux de reconnaissance en fonction de la dimension de l'image pour différents nombres de neurones dans la couche cachée. On remarque que l'augmentation du nombre de neurones dans la couche cachée ne donne pas forcément un meilleur taux de reconnaissance. Le meilleur résultat est obtenu avec un nombre de neurone égal à 80 et une taille d'image égale à 15x15.

➤ Influence du seuil de classification:

Dans cette deuxième phase, on cherche à évaluer l'influence du seuil sur la qualité de classification et en déterminer une valeur optimale. Pour cette expérience, la structure du réseau et la taille des images sont fixées aux valeurs optimales trouvées dans phase précédente, à savoir :

Nombre de neurones cachées : 80

La dimension d'image : 15x15

Les taux de reconnaissance pour différents seuils sont donnés par le tableau 4.2.

seuil	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
T.de Rec(%)	84.3750	88.7500	91.6250	88.1250	83.7500	82.5000	78.1250	70	61.2500

Tableau 4.2: Taux de reconnaissance en fonction du seuil.

La variation du taux de reconnaissance avec le seuil est donnée dans la figure 4.14.

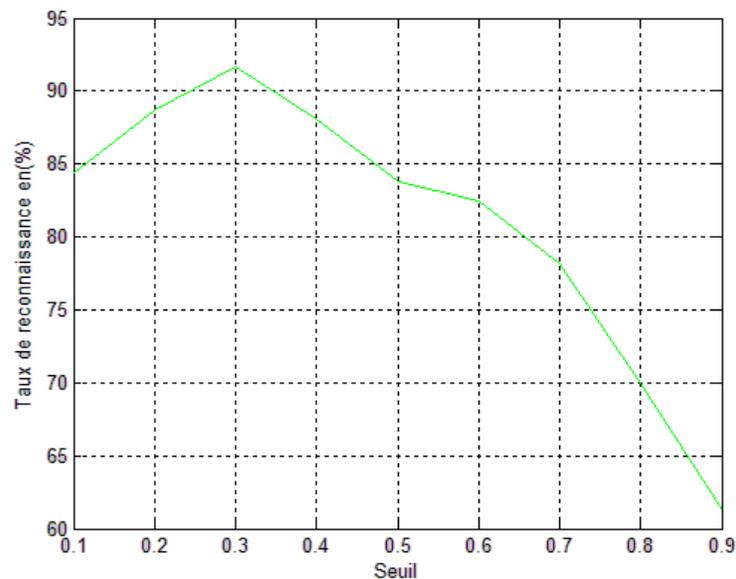


Figure 4. 14: Variation de taux de reconnaissance avec le seuil

On remarque, sur cette figure, qu'il y a une plage de taux de reconnaissance optimale correspondante à un seuil situé entre 0.2 et 0.3 ; ensuite l'augmentation du seuil entraîne une diminution très rapide du taux de reconnaissance.

4.4 Reconnaissance de visages par PCA + Réseau MLP

Malgré que l'utilisation des pixels des images brutes, comme entrées au réseaux de neurones, donne une très bonne performance, elle engendre aussi l'utilisation d'un réseau de grande taille (256 entrées pour une image 15*15) avec un grand nombre de paramètres à ajuster. Pour remédier à ce problème, on se propose dans cette section d'utiliser l'analyse en composantes principales (PCA ou ACP) pour réduire l'information présentée au

réseau de neurone, et n'en garder que la portion utile représentée par un certain nombre de composantes principales sélectionnées.

Après avoir calculé l'ensemble des vecteurs caractéristiques (Matrice de visages réduite), on utilise un réseau de neurones multicouches pour la phase de classification. Après le passage des vecteurs caractéristiques plusieurs fois dans le réseau (processus d'apprentissage), les paramètres (les poids,...) ainsi obtenus sont sauvegardés. L'algorithme d'apprentissage utilisé est l'algorithme de rétro propagation du gradient.

Le principe général de l'apprentissage est donné par l'organigramme de la figure 4.15

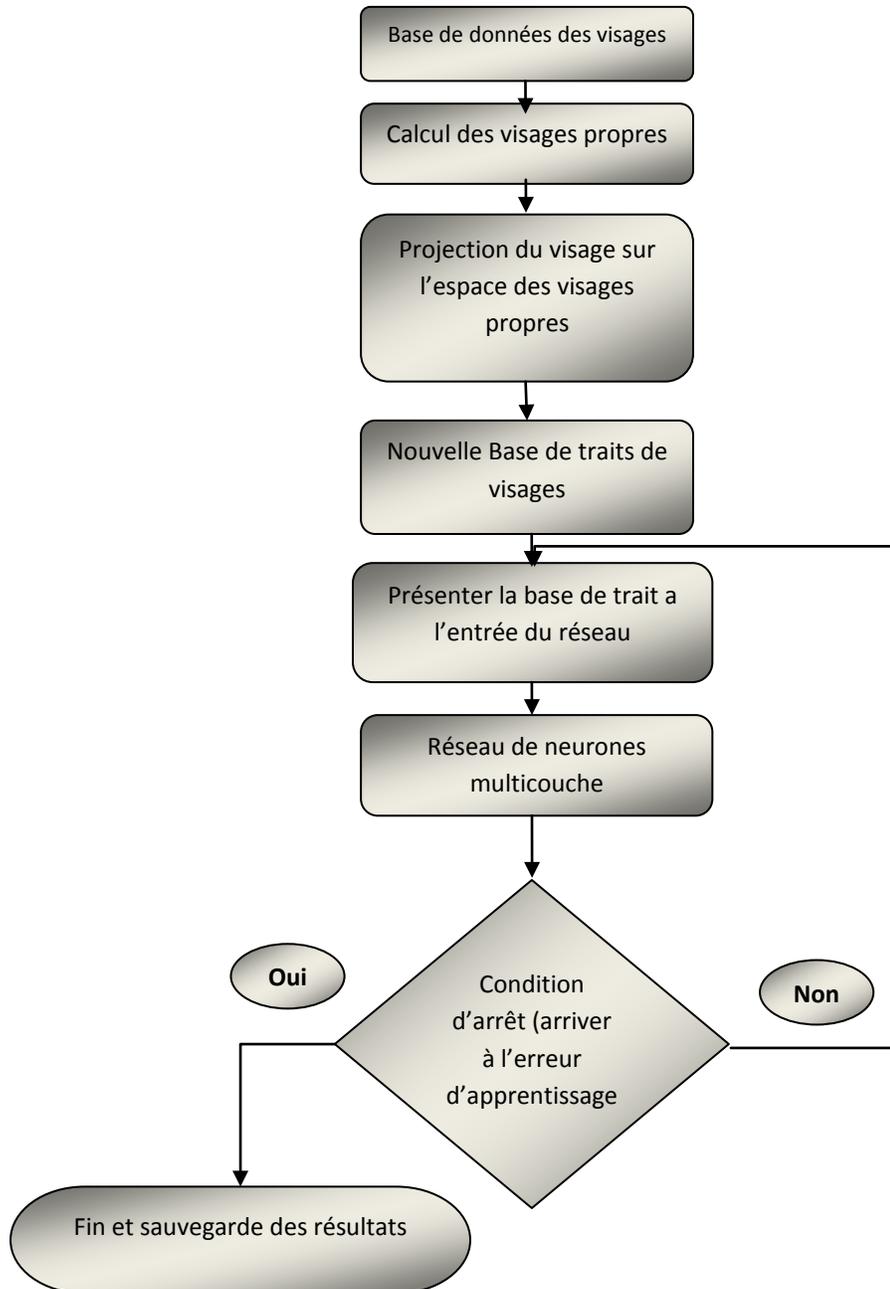


Figure 4.15: Organigramme générale d'apprentissage.

4.4.1 Processus d'identification:

Après avoir initialisé le système, le processus d'identification peut être lancé (figure 4.15) Le processus d'identification comporte les étapes suivantes:

- Restaurer les paramètres du système (les poids du réseau, les visages propres)
- Introduire un nouveau visage.

- Projeter ce visage sur l'espace des visages propres pour obtenir son vecteur trait.
- Présenter ce vecteur à l'entrée du réseau de neurones et calculer les sorties des neurones de la couche de sortie.
- La méthode d'estimation des poids est: rétro propagation de gradient.

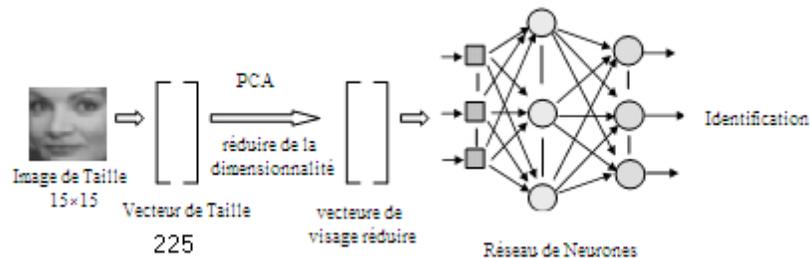


Figure 4.16: Processus d'identification

Les paramètres utilisés sont:

1. Type de Réseau de neurone: MLP mon réseau
2. Nombre de vecteurs propres utilisés: 80
3. Nombre de neurones dans la couche d'entrée: 80
4. Nombre de neurones dans la couche cachée: 80
5. Nombre de neurones dans la couche de sortie: 40
6. Pas d'apprentissage:0.2
7. Nombre d'itération : très grand !!

Les différents résultats obtenus donner dans le tableau suivant

Seuil	0.4	0.5	0.6	0.7	0.8	0.9
T.de Rec (%)	87.5000	86.8750	85	83.1250	80	68.7500

Tableau 4.7 : Taux de reconnaissance en fonction du seuil.

Le meilleur taux de reconnaissance lorsque en utilise un seuil est le : **87.5000 %**.

Le temps d'apprentissage est resté toujours long mais il est court para port la méthode de réseau de neurones seule.

4.5 CONCLUSION :

Le meilleur taux de reconnaissance est de **91.6250%** avec **80** neurones sur la couche cachée en dimension **15×15**,. Pour une classification avec un seuil, le résultat tombe à **87.5000%** avec **80** neurones sur la couche cachée en dimension **15×15**.

Alors que la méthode de la distance euclidienne avec la même taille d'image donne 92.50% comme taux de reconnaissance et 92.50 % pour la PCA et ceci pour 90 face propres.

Conclusion Générale

Dans notre travail nous avons étudié l'identification (reconnaissance) de l'identité employant des images faciales. Plusieurs techniques de reconnaissance de visage ont été testées sur une base de données standard de visage ORL.

La première méthode étudiée repose sur une comparaison simple des distances Euclidiennes entre les vecteurs représentant les images pour déterminer la personne la plus favorable (probables). Dans la deuxième méthode, nous avons utilisé une approche statistique, la PCA pour réduire la dimensionnalité des images. Durant les tests, nous avons constaté que ces méthodes produisent des résultats acceptables, mais n'atteignent pas une précision suffisante. Pour améliorer les performances du système d'identifications, nous avons étudié l'utilisation des réseaux de neurones de type MLP. Les essais ont montré aussi que l'algorithme de rétro propagation nécessite une très longue période d'entraînement. La discrimination représente l'étape finale de la reconnaissance. Le seuil présente l'avantage de pouvoir rejeter toute une base d'imposteurs (c.-à-d ne pas classer du tout). Cependant, le seuil présente l'inconvénient de ne pas être le même pour des bases autorisées et des bases non autorisées (comme vu dans les tests, le seuil doit être bas pour améliorer l'acceptance et élevé pour améliorer la rejection). plus élevés, présente la fâcheuse habitude de toujours classer. Aussi,

Les résultats que nous avons obtenus par l'utilisation de l'algorithme de rétro propagation sont très encourageants, comparés avec les autres méthodes d'apprentissage. Le tableau suivant illustre les différents résultats obtenus durant ce travail.

Méthode de Reconnaissance	Dimension d'images	Taux de Reconnaissance(%)	Temps d'apprentissage	Temps de test
Distance Euclidienne	15×15	92.5000	Court (quelques secondes)	Très court (quelques milli secondes)
PCA	15×15	92.5000	Court (quelques secondes)	Très court (quelques milli secondes)
MLP Mono Réseau	15×15	91.6250	Très long !!(15-20 heures environ) avec algorithme d'apprentissage BP	Court (quelques secondes)
PCA+ MLP Mono Réseau	15×15	87.5000	Long (plusieurs heures) mais court par rapport à Mono Réseau seul	Court (quelques secondes)

Tableau 5.1 : les meilleurs résultats de reconnaissance obtenus

Le Tableau 5.1 reprend les meilleurs résultats de reconnaissance fournis par chaque technique de classification.

Les performances atteintes par les réseaux de neurones montrent leur efficacité dans le domaine de reconnaissances, comparés aux techniques classiques tel que la distances Euclidienne, et PCA. L'inconvénient principal des réseaux de neurones demeure que le temps nécessaire à leur entraînement (jusqu'au 20 heures!!!).

Les résultats que nous avons obtenus par l'utilisation des réseaux de neurones sont très encourageants, ce qui encourage sans doute, les promotions à venir à essayer d'améliorer les méthodes d'identification de visage .Nous avons proposé à venir:

- Pour tester l'efficacité réelle de ce système, il faudra étendre la base de visages à un plus grand nombre de classes en choisissant une base reconnue sur le plan international comme la base Ferret.
- Pour amélioration la méthode de la rétro propagation, plusieurs approches ont été proposées: la modification du pas d'apprentissage, l'utilisation des algorithmes du type quasi newton.
- En utilisent seulement quelque composantes principales de visages telles que: des sourcils, des yeux, le nez, et la bouche.
- En remplace l'algorithme de rétro propagation a cause de leur inconvénients (la lenteur de la convergence, la sensibilité aux minima locaux et la difficulté à régler les

paramètres du réseau) par les algorithmes génétiques. Les algorithmes génétiques ont été utilisés de plusieurs façons pour optimiser les réseaux de neurones. Le principal avantage est que la recherche s'effectue en parallèle sur plusieurs points de l'espace de solution, ainsi la probabilité de tomber sur un optimum local est moindre que dans le rétro propagation.

Références Bibliographiques

- [1] A. Lalita "La biométrie et son usage par l'état" Division des sciences et de la technologie le 11 septembre 2006.
- [2] Commission Technique de Sécurité Physique."Technique de contrôle d'accès par biométrie".
- [3] C. Christian "les méthodes scientifiques d'identification des personnes à partir de données biométriques et les techniques de mise en œuvre" rapport, office parlementaire d'évaluation des choix scientifiques et technologique, 2003
- [4] M. Tayeb Laskri, D. Chefrour: « Who_is Système d'identification des visages Humain », Département d'Informatique - arima 2002 Université Badji Mokhtar Annaba2002.
- [5] Gonzalez, Rafael C « Digital image processing » volume2 1987.
- [6] S. Mori, H. Nishida and H. Yamada, "Optical Character Recognition", Copyright © 1999 by John Wiley and Sons.
- [7] P. Q. Dung" Réseaux de Neurones pour la reconnaissance des Formes" Travail d'intérêt Personnel Encadré par Hanoi, Juillet 2005.
- [8] H. Bredin "Vérification biométrique d'identité basée sur les visages parlantes. Apport de la mesure de synchronie audiovisuelle face aux tentatives d'imposture élaborées." Juin 2007
- [9] B. Miller "Vital signs of identity". IEEE Spectrum, Vol.31, no.2, pp.22, 30 February 1994.
- [10] <http://biometrie.online.fr/>

-
- [11] <http://www-asim.lip6.fr/~marzouki/perso/publi/cnil-biometrie01.html>
- [12] http://www.cite-sciences.fr/francais/ala_cite/expositions/biometrie/index2.php
- [13] <http://ophtasurf.free.fr/vue/biometrie.htm>
- [14] http://solutions.journaldunet.com/0410/041005_biometriques.shtml
- [15] <http://www.securiteinfo.com/conseils/biometrie.shtml>
- [16] http://www.souriez.info/article.php3?id_article=71
- [17] <http://big.chez.com/gipp/oraux/aal/index.html>
- [18] http://solutions.journaldunet.com/0411/041102_biometrie_get.shtml
- [19] <http://www.lexpress.fr/info/sciences/dossier/biometrie/dossier.asp>
- [20] <http://www.cl.cam.ac.uk/users/jgd1000/>
- [21] <http://perso.wanadoo.fr/fingerchip/index.htm>
- [22] R. Boumaza, "Cours de Traitement d'images", EMP, 1999.
- [23] R. Cyril, "prévision de température par réseaux neuronaux". dossier 2003-204 Lycée Victor Hugo Besançon.
- [24] C. Touzet " les réseaux de neurones artificiels, introduction au connexionnisme", cours, exercices et travaux pratiques, Juillet 1992
- [25] H. Abdi "les réseaux de neurones", livre, collection sciences et technologies de la connaissance 269
- [26] C. L. Giles "Méthodes d'authentification vocale d'utilisateurs dans les systèmes informatique" thème d'ingénieur C.N.A.M. en informatique, centre régional associe de Strasbourg, 2000.
- [27] K.Beghdad bey "Techniques de classification d'Empreintes Digitales" mémoire présenté pour l'obtention du diplôme de magister, Filière: Robotique, Automatique et Informatique Industrielle, option Informatique Industrielle. Ecole Militaire Polytechnique, EMP, 2003
- [28] A.K. Jain, L. Hong, S. Pankanti, and R. Bolle "An Identity-Authentication system Using Fingerprints ".Proceedings of the IEEE, Vol.85, no. 9, pp.1365-1388, September 1997.

-
- [29] Cours "Introduction aux réseaux de neurones" Master 2005-2006
- [30] Commission Technique de Sécurité Physique. "Technique de contrôle d'accès par biométrie".
- [31] A.Boutarfa" Reconnaissance de Formes 3D par approche Neuronale Associant la transforme de Hough en Robotique Mobile Application à la productique", Thèse Doctorat Es Sciences en Electronique industrielle, option: contrôle, 2006
- [32] Site Internet: <http://www.grappa.univ-lile3.fr/polys/apprentissage/sortie006.htm>
- [33] L. O'gorman. Biometrics Personal Identification in Networked society, chapter Fingerprint Verification, pages 43-64. Kluwer Academic Publishers. 1999
- [34] W.S. McCulloch and W. Pitts, "A logical calculus of ideas immanent in nervous activity", Bull. Mathematical Bio-physics, Vol. 5, 1943, pp. 115-133.
- [35] M. Minsky and S. Papert, "Perceptrons: An introduction to computational geometry", MIT Press, Cambridge. Mass, 1969.
- [36] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", in Roc. Nat, Academy of Sciences, USA 79, 1982, pp. 2,554 2,558.
- [37] D.E. Rumelhart and J.L. McClelland, "Parallel distributed processing: Exploration in the microstructure of cognition", MIT Press, Cambridge. Mass, 1986.
- [38] P.D.Power, "Non linear Multilayer Perceptron channel equalization", PHD Dissertation, University of Belfast, 2001.
- [39] G. Dreyfus, T. Cour, G. Giraud, A. Kodsi, T. Luong, R. Lauranson, C. Marcovici, K. Sadeghi, "Reconnaissance de formes par réseaux de neurones", Ecole polytechniques, (2002), France.
- [40] Claude Touzet, "les réseaux de neurones artificiels: introduction au connexionnisme", juillet 1992.

- [41] S. Siu, G.J. Gibson, and C.F.N. Cowan, “ Decision feedback equalisation using neural network structures and performance comparison with standard architecture”, IEEE Proceedings, Vol.137, Pt. I, No. 4, pp.221-225, August 1990
- [42] Jiant-Kang “neural networks and simulation methods”, CRC Press (Déc 1993) ISBN 0-8247-9181-9(acid-free paper)
- [43] M.Ibnnkahla, “Application of neural networks to digital communication”, Signal processing 80, pp 1185-1215, 2000.
- [44] B.Mulgrew, “Applying radial basis function networks”, IEEE Signal Processing Magazine, pp. 50–65, Mar. 1996.
- [45] S. Haykin, “Neural Networks: A Comprehensive Foundation”, 2nd Ed, New York: Macmillan College Publishing, 1998.
- [46] K. Hornik, “Multilayer feed forward networks are universal approximators”, Neural Networks, vol.2, pp .359-366. 1989.
- [47] Site Internet <http://www.cam-orl.co.uk/facedatabase.html>.

Annexe

A) Algorithme du rétro propagation de gradient:

On considère un réseau de neurones statique à trois couches: une couche d'entrée, une couche de sortie et une seule couche cachée. Supposons qu'on a un ensemble d'exemples constitué d'une suite de L vecteurs d'entrées, $X_p = [x_{p1} x_{p2} \dots x_{pn}]^T$, $p = 1 \dots L$ et n le nombre de neurones de la couche d'entrée. L'ensemble des réponses désirées constitué aussi de L vecteurs $T^{(p)} = [t_1^{(p)} t_2^{(p)} \dots t_m^{(p)}]^T$. Le critère des moindres carrés de l'erreur peut être utilisé pour définir la fonction de coût à minimiser, qui s'exprime alors:

$$E = \frac{1}{2} \sum_{p=1}^L \sum_{i=1}^m (t_i^{(p)} - y_i^{(p)})^2 \quad (\text{A.1})$$

- L : est le nombre d'exemples d'apprentissage.
- m : Le nombre de neurones que contient la couche de sortie.
- $y_i^{(p)}$: désigne la sortie du neurone i de la couche de sortie lorsque le vecteur X_p est présenté à l'entrée du réseau.
- $t_i^{(p)}$: représente la valeur de sortie désirée pour le neurone i de la couche de sortie lorsque le vecteur X_p est présenté à l'entrée du réseau.

La minimisation de cette fonction de coût non linéaire se fait de manière itérative, en utilisant une méthode de gradient

A chaque itération, un vecteur d'entrée $X_p = [x_{p1} x_{p2} \dots x_{pn}]^T$ ainsi que le vecteur associé de sorties désirées $T^{(p)} = [t_1^{(p)} t_2^{(p)} \dots t_m^{(p)}]^T$ sont présentés au système, les neurones d'entrées distribuent ces valeurs aux neurones de la couche cachée. La sortie du $j^{\text{ième}}$ neurone de la couche cachée est:

$$S_{pj} = f_j^c(z_{pj}^c) \quad (\text{A.2})$$

$$\text{Avec } z_{pj}^c = \sum_{i=1}^n w_{ji}^c x_{pi} \quad (\text{A.3})$$

Où w_{ji}^c est le poids de connexion entre le neurone i de la couche d'entrée et le neurone j de la couche cachée, c spécifie la couche cachée.

La sortie du neurone k de la couche de sortie est :

$$y_k^{(p)} = f_k^s(z_{pk}^s) \quad (\text{A.4})$$

$$\text{Avec } z_{pk}^s = \sum_{j=1}^{nc} w_{kj}^s S_{pj} \quad (\text{A.5})$$

Où s : signifie la couche de sortie

nc : Le nombre de neurones cachés

On définit l'erreur à un neurone de sortie par :

$$\delta_k^{(p)} = t_k^{(p)} - y_k^{(p)} \quad (\text{A.6})$$

Où p signifie le $p^{\text{ème}}$ exemple d'entraînement et k signifie le $k^{\text{ème}}$ neurone de sortie.

L'algorithme consiste à ajuster ou changer les poids de connexion en minimisant la fonction coût suivant :

$$E^{(p)} = \frac{1}{2} \sum_{i=1}^m (\delta_k^{(p)})^2 \quad (\text{A.7})$$

✓ Ajustement des poids de la couche de sortie

La loi d'adaptation des poids est donnée par :

$$w_{kj}^s(t+1) = w_{kj}^s(t) - \eta \frac{\partial E^{(p)}}{\partial w_{kj}^s(t)} \quad (\text{A.8})$$

Où: η constante positive appelé taux d'apprentissage $\frac{\partial E^{(p)}}{\partial w_{kj}^s(t)}$ est appelé *gradient*

instantané de l'erreur.

$$\frac{\partial E^{(p)}}{\partial w_{kj}^s(t)} = \frac{1}{2} \frac{\partial}{\partial w_{kj}^s} \left[\sum_{k=1}^m (t_k^{(p)} - y_k^{(p)})^2 \right] \quad (\text{A.9})$$

$$= \frac{1}{2} \frac{\partial}{\partial y_k^{(p)}} \left[\sum_{k=1}^m (t_k^{(p)} - y_k^{(p)})^2 \right] \frac{\partial y_k^{(p)}}{\partial w_{kj}^s} \quad (\text{A.10})$$

$$= \frac{1}{2} \frac{\partial}{\partial y_k^{(p)}} \left[\sum_{k=1}^m (t_k^{(p)} - y_k^{(p)})^2 \right] \frac{\partial y_k^{(p)}}{\partial (z_{pk}^s)} \frac{\partial (z_{pk}^s)}{\partial w_{kj}^s} \quad (\text{A.11})$$

$$\frac{\partial}{\partial y_k^{(p)}} \left[\sum_{k=1}^m (t_k^{(p)} - y_k^{(p)})^2 \right] = -2(t_k^{(p)} - y_k^{(p)}) \quad (\text{A.12})$$

$$\frac{\partial(z_{pk}^s)}{\partial w_{kj}^s} = \frac{\partial}{\partial w_{kj}^s} \left(\sum_{j=1}^L w_{kj}^s S_{pj} \right) = S_{pj} \quad (\text{A.13})$$

On remplace l'équation (4.13) et (4.14) dans (4.12), on obtient :

$$\frac{\partial E_p}{\partial w_{kj}^s} = -(t_k^{(p)} - y_k^{(p)}) f_k'^s(z_{pk}^s) S_{pj} \quad (\text{A.14})$$

On remplace l'équation (4.15) dans (4.9), L'équation d'ajustement des poids écrite comme :

$$w_{kj}^s(t+1) = w_{kj}^s(t) + \eta(t_k^{(p)} - y_k^{(p)}) f_k'^s(z_{pk}^s) S_{pj} \quad (\text{A.15})$$

✓ Ajustement des poids de la couche cachée

La loi d'adaptation des poids est donnée par

$$w_{ji}^c(t+1) = w_{ji}^c(t) - \eta \frac{\partial E^{(p)}}{\partial w_{ji}^c(t)} \quad (\text{A.16})$$

En utilisant l'équation (A.7) :

$$\begin{aligned} E^{(p)} &= \frac{1}{2} \sum_{i=1}^m (\delta_k^{(p)})^2 \\ &= \frac{1}{2} \sum_k (t_k^{(p)} - f_k^s(z_{pk}^s))^2 \\ &= \frac{1}{2} \sum_k (t_k^{(p)} - f_k^s \left(\sum_j w_{kj}^s S_{pj} \right))^2 \end{aligned} \quad (\text{A.17})$$

On sait que S_{pj} dépend des poids de la couche cachée à partir des équations (A.2) et (A.3), On peut calculer le gradient par rapport aux poids de la couche cachée comme

suit :

$$\begin{aligned} \frac{\partial E^{(p)}}{\partial w_{ji}^c(t)} &= \frac{1}{2} \sum_k \frac{\partial}{\partial w_{ji}^c} (t_k^{(p)} - y_k^{(p)})^2 \\ &= - \sum_k (t_k^{(p)} - y_k^{(p)}) \frac{\partial y_k^{(p)}}{\partial (z_{pk}^s)} \frac{\partial (z_{pk}^s)}{\partial S_{pj}} \frac{\partial S_{pj}}{\partial (z_{pj}^c)} \frac{\partial (z_{pj}^c)}{\partial w_{ji}^c} \\ \frac{\partial E_p}{\partial w_{ji}^c} &= - \sum_k (y_{pk}^d - y_{pk}) f_k'^s(z_{pk}^s) w_{kj}^s f_j'^c(z_{pj}^c) x_{pi} \end{aligned} \quad (\text{A.18})$$

L'équation d'ajustement des poids écrite comme :

$$w_{ji}^c(t+1) = w_{ji}^c(t) + \eta \sum_k (y_{pk}^d - y_{pk}) f_k'^s(z_{pk}^s) w_{kj}^s f_j'^c(z_{pj}^c) x_{pi} \quad (\text{A.19})$$

B) quelques faces de la base ORL (<http://www.cam-orl.co.uk/facedatabase.html>.)



Personne 1



Personne 2



Personne 3



Personne 4



Personne 5



Personne 6



Personne 7



Personne 8



Personne 9



Personne 10



Personne 11



Personne 12



Personne 13



Personne 14



Personne 15



Personne 16



Personne 17



Personne 18



Personne 19



Personne 20



Personne 21



Personne 22



Personne 23



Personne 24



Personne 25



Personne 26



Personne 27



Personne 28



Personne 29



Personne 30



Personne 31



Personne 32



Personne 33



Personne 34



Personne 35



Personne 36



Personne 37



Personne 38



Personne 39



Personne 40

- **Voici un exemple où l'acquisition se fait sous différentes orientations du visage :**



Exemple de changements d'orientations du visage

- **Voici un exemple où l'acquisition se fait sous différents éclairages :**



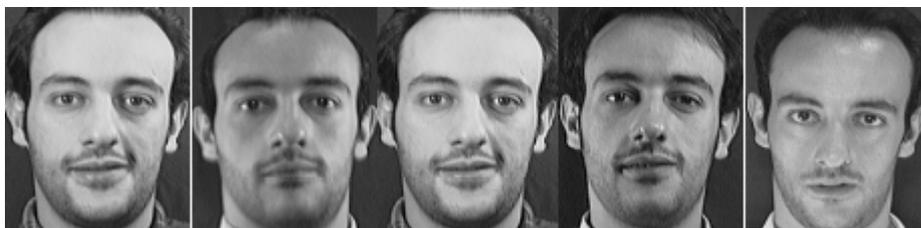
Exemple de changements d'éclairage

- **Voici un exemple où l'acquisition se fait sous différentes orientations du visage :**



Exemple de changements d'orientations du visage

- **Voici un exemple où l'acquisition se fait sous différents éclairages :**



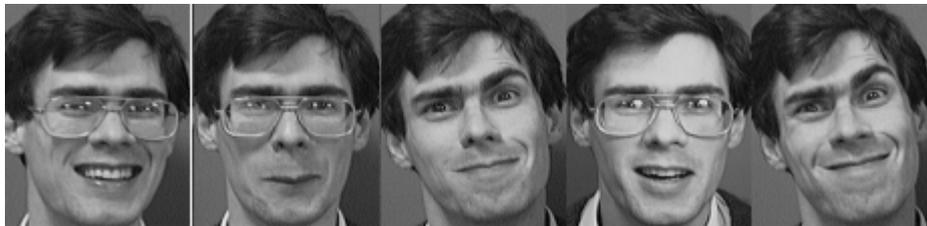
Exemple de changements d'éclairage

- Cet exemple montre les changements d'échelle dus à la distance entre le dispositif d'acquisition et l'individu :



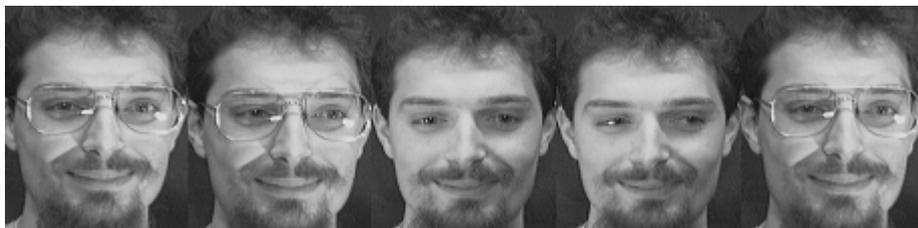
Exemple de changements d'échelle

- La base de donnée ORL prend aussi en considération les expressions faciales, telles que les grimaces. En voici un exemple :



Exemple de changements des expressions faciales

- La Base ORL prend en compte le fait qu'un individu peut porter ou ne pas porter des lunettes. Cet exemple en est l'illustration :



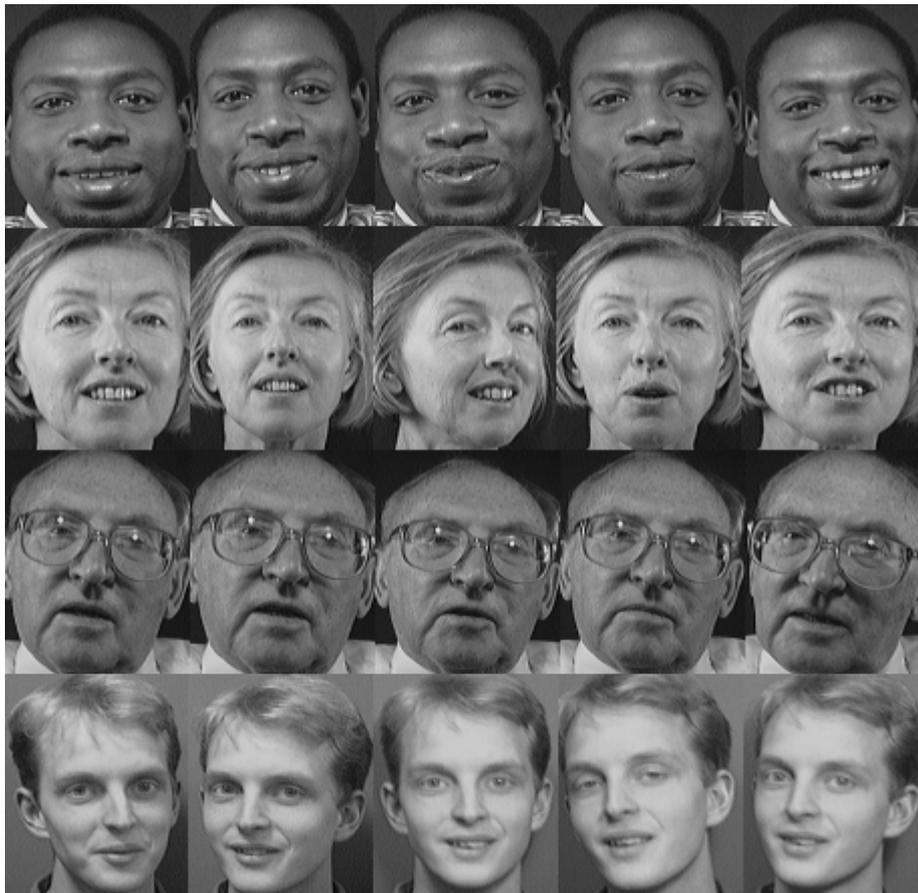
Exemple de port de lunettes

- Un individu peut aussi porter une barbe ou changer de coiffure, la base ORL prend en considérations ces particularités :



Exemple de changements de coiffure et de port de barbe

- La base ORL comprend aussi des individus de différents âges, sexe et couleurs de peaux :



Exemple d'individus de différents âges, races et sexes