

M/004.466

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique

Université de Guelma

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Département d'informatique



Mémoire de fin d'études Master

13/854

Filière : Informatique

Option : Ingénierie des Médias

Thème :

Systeme de reconnaissance de visage par l'hybridation de la DCT et les réseaux de neurones

Encadré par :

Mr Hallaci Samir

Présenté par :

Arjuoui Nahila

Benassou Nabila



Juin 2013



Remerciements



Voilà, nous sommes enfin de parcours. Et comme c'est la phase finale : la phase de maturation - que nous espérons de bonne fructification- il nous est très agréable de remercier toutes les personnes qui nous ont aidé de loin ou de près.

Au premier rang, nous tenons à remercier vivement notre encadreur Mr Hallaci Samir, à la fois pour avoir accepté de nous guider et surtout pour les précieux conseils qu'il nous a fournis.

Nos remerciements les plus sincères vont également à Mme Bencheriat Chams Ennahare qui n'a pas manqué de nous orienter chaque fois qu'un obstacle fut immergé.

Nous devons exprimer profondément notre gratitude et reconnaissance à messieurs : Bouallague Halim, Oujuni Brahim et Amara Madi Ahcene pour le grand volume de temps qu'ils nous ont consacré et trouverons ici nos vifs remerciements.





Dédicace



*Le présent travail est le fruit d'une longue
période soldant cinq années de déploiement des efforts continus
pour un avenir meilleur impatientement souhaité par mes parents.
C'est pourquoi, il m'est très agréable de partager la joie de ce
Fruit avec le papa, la maman et tous les membres de ma famille.*

Nahla





Dédicace



Je dédie ce modeste travail à ...



*A ma adorable mère à la source de tendresse et d'amour,
A mon exemplairement mon père,
Je vous dois ce que je suis aujourd'hui grâce à votre amour, à
votre patience et vos innombrables sacrifices.*

A mes très chères sœurs, frère (Hizia, Imene, Hatem)

*A ma petite sœur amel qu'il est départe au dieu le crée de
monde, tu as une place particulière spéciale dans mon cœur je
t'aime beaucoup rahamaki Allah.*

A tout ma grande famille et mes amies.



Nabila



Résumé

La reconnaissance de visage ou la reconnaissance faciale est l'un des problèmes qui présentent le plus grand défi dans la vision par ordinateur et la reconnaissance de formes. La difficulté provient essentiellement des variations de l'apparence du visage causées par des facteurs tels que l'expression, l'éclairage, l'occlusion partielle du visage, et l'écart de temps entre la formation et les tests de capture de données. Par ailleurs, les performances des algorithmes de reconnaissance de visage dépendent fortement de l'étape de localisation des caractéristiques du visage.

L'algorithme de reconnaissance de visage proposé est basé sur la représentation des régions locales du visage en utilisant la transformée en cosinus discrète (DCT). La représentation locale fournit une robustesse contre les variations de l'apparence dans les régions locales causées par l'occlusion partielle du visage ou l'expression du visage. Alors que l'utilisation de l'information de fréquence fournit une protection robuste contre les changements d'éclairage. En outre, l'algorithme évite l'étape de localisation des traits du visage et formule l'alignement du visage comme étant un problème d'optimisation dans l'étape de classification. Cette dernière est faite à l'aide d'un réseau de neurones multicouches.

Mots Clés : Biométrie, reconnaissance faciale, DCT, apprentissage, réseaux de neurones, PMC.

Sommaire

Introduction générale

Chapitre 1 :Généralité se la reconnaissance faciale

I. Introduction.....	1
II. La vision par ordinateur	1
1. Définition	1
2. Catégories de la vision par ordinateur	1
III. L'image numérique	1
1. Définition	2
2. Caractéristiques d'une image numérique	2
2.1. Pixel	2
2.2. Dimension	2
2.3. Résolution.....	2
2.4. Histogramme	2
2.5. Contours et textures	3
2.6. Luminance	3
2.7. Contraste	3
2.8. Image en couleur	3
2.9. Bruit	3
3. Les types des images	4
3.1. L'image en pixel (Bitmap)	4
3.2. L'image vectorielle	4
4. La compression des images	5
4.1. La compression sans perte (réversible).....	5
4.2. La compression avec pertes (irréversible).....	5
4.3. Rapport et taux de compression	6
IV. La reconnaissance de visage dans une image fixe.....	6
1. Les systèmes biométriques	6
1.1. Définition.....	6

1.2. Les technologies biométriques existantes	7
1.2.1. Techniques morphologiques.....	7
1.2.2. Techniques comportementales	10
2.3. Architecture d'un système biométrique.....	11
2.4. Mesure de performances d'un système biométrique.....	12
2.5. Vision futur de la biométrie	14
3. Systèmes biométriques basés sur la reconnaissance faciale	15
3.1. Descripteurs et caractéristiques du visage dans une image.....	15
3.2. Processus de reconnaissance de visage	16
2.2.1. Acquisition de l'image.....	16
2.2.2. prétraitement.....	17
2.2.3. Détection du visage.....	17
2.2.4. Extraction des paramètres et classification.....	17
2.2.5. Apprentissage	17
2.2.6. Test et décision.....	18
2.3. Difficultés de la reconnaissance de visage.....	18
2.3.1. Changement d'illumination.....	18
2.3.2. Variation de pose.....	18
2.3.3. Expressions faciales.....	19
2.3.4. Présence ou absence des composants structurels	19
2.3.5. Occultations partielles.....	20
V. Conclusion.....	20

Chapitre 2 :Techniques de reconnaissance faciale

I. Introduction.....	23
II. Les étapes de reconnaissance du visage	23
II.1. la détection de visage	23
II.1.1. Définition.....	23
II.1.1.1 Approches de détection du visage	23
II.1.1.1.1. Approche basée sur les connaissances acquises.....	24
II.1.1.1.2. Approches basées sur la mise en correspondance	24

II.1.1.1.3. Approches basées sur l'apparence	24
II.1.1.1.4. Approche basée sur des caractéristiques invariables	25
II.2. l'extraction de signatures du visage.....	25
II.2.1. Définition.....	25
II.2.2. Méthodes d'extraction des caractéristiques de visage	26
II.2.2.1. Méthode statistique	26
II.2.2.2. Méthode géométrique.....	26
II.3. La reconnaissance de visage.....	27
II.3.1. Définition.....	27
II.3.2. Les méthodes de reconnaissance de visage.....	27
II.3.2.1. Les méthodes globales	27
II.3.2.1.1 Définition	27
II.3.2.1.2. Les techniques linéaires.....	27
a) la transformée en cosinus discrète	28
b) Eigenface	28
c) FisherFaces.....	29
d) L'analyse en composantes indépendantes ACI.....	29
II.3.2.1.3. Les techniques non linéaires.....	29
a) Réseaux de neurones	29
II.3.2.1.4. Évaluation des méthodes globales	30
II.3.2.2. Les méthodes locales.....	30
II.3.2.2.1. Définition	30
a) les méthodes basées sur les caractéristiques locales.....	30
b) les méthodes basées sur les apparences locales.....	31
II.3.2.2. Évaluation des méthodes locales	31
II.3.2.3. Les méthodes hybrides	31
II.3.2.3.1. Définition	31
II.3.2.3.2. Évaluation des méthodes hybrides.....	31
III. La méthode adoptée.....	32
IV. Conclusion	33

Chapitre 3 : Conception

I. Introduction.....	34
II. Les méthodes développées	34
II.1. La transformée en cosinus discrète.....	34
II.1.1. Définition.....	34
II.1.2. Variantes de la DCT.....	34
II.1.3. Propriété de la DCT	35
II.1.4. Application de la DCT à la problématique de reconnaissance faciale.....	36
II.2. Les réseaux de neurones.....	36
II.2.1. Définition.....	36
II.2.2. Architecture d'un réseau de neurone	37
II.2.3. L'apprentissage dans un réseau de neurones	37
II.2.3.1. Définition.....	37
II.2.3.2. Types d'apprentissage	38
II.2.4. Le perceptron multicouche.....	38
II.2.5. Fonction d'activation	40
III. Conception du système.....	42
III.1. prétraitement.....	44
III.1.1. Normalisation	44
III.1.2. Egalisation de l'histogramme.....	44
III.1.3. Filtre médian.....	45
III.2. Extraction des caractéristiques du visage par la DCT.....	45
III.2.1. Découpage de l'image en blocs.....	45
III.2.2. Application de la DCT sur chaque bloc.....	46
III.2.3. Récupération des coefficients DCT (scan zigzag)	47
III.2.4. Sélection des premiers coefficients de la DCT et extraction de vecteur de chaque bloc	48
III.2.5. Concaténation des vecteurs des blocs et formation du vecteur final.....	48

III.3. Classification	48
III.3.1. Architecture du modèle PMC.....	48
III.3.2. Phase d'apprentissage	50
III.3.3. phase de reconnaissance	52
IV. Conclusion	53

Chapitre 4 : Implémentation

I. Introduction.....	54
II. Environnement de développement	54
II.1. Environnement matériel	54
II.2. Environnement logiciel	54
II.2.1. Java.....	54
II.2. Matlab.....	55
II.3. Bases de données utilisées.....	56
II.3.1. Base ORL	56
II.3.2. Base Yale faces	58
III. Présentation de l'interface graphique.....	59
IV. Tests et résultats	62
IV.1. Test réalisé sur les paramètres de la DCT :	62
IV.1.1. Taille de blocs	62
IV.1.2. Nombre de coefficients.....	62
IV.2. Test réalisées sur les paramètres du PMC.....	62
V. Conclusion	64
Conclusion générale.....	65

Liste de figure

Figure I.1 : Synthèse additive.....	03
Figure I.2 : Principe de la compression d'image.....	05
Figure I.3 : Principales techniques de reconnaissance d'individus.....	07
Figure I.4 : Photo d'iris.....	08
Figure I.5 : Spectre d'un signal voix.....	09
Figure I.6 : Empreinte digitale.....	09
Figure I.7 : Schéma de vérification du visage.....	12
Figure I.8 : Schéma d'identification du visage.....	12
Figure I.9 : Courbe du point d'équivalence des erreurs dans un système biométrique.....	13
Figure I.10 : Courbe ROC.....	14
Figure I.11 : Exemple d'extraction des contours du visage.....	15
Figure I.12 : Processus de reconnaissance du visage.....	16
Figure I.13 : Excmple de détection du visage.....	17
Figure I.14 : Exemples de variation d'illumination de la base YaleB.....	18
Figure I.15 : Exemples de variation de poses.....	19
Figure I.16 : Exemples de variation d'expressions.....	19
Figure I.17 : Exemple la présence des composants structurels.....	19
Figure 2.1 : Les étapes de la reconnaissance de visage.....	21
Figure 2.2 : Visage de la première personne de la base de données Yale Faces.....	24
Figure 2.3 : Visage accompagné de mesures géométriques.....	24
Figure 2.4 : Mesures géométriques réalisées sur le visage.....	25
Figure 2.5 : Les différents modules d'un système de reconnaissance de visage basé sur l'ACP.....	26
Figure 2.6 : Distance entre points caractéristiques d'un visage.....	29
Figure 3.1 Composition d'un réseau de neurones.....	35
Figure 3.2 : Perceptron multicouche.....	37
Figure 3.3 : Modèle d'un perceptron multicouche pour la reconnaissance de visage.....	37
Figure 3.4 : Fonction Heaviside.....	38
Figure 3.5 : Fonction signe.....	38
Figure 3.6 : Fonction linéaire.....	39
Figure 3.7 : Fonction linéaire à seuil.....	39
Figure 3.8 : Fonction sigmoïde.....	40

Figure 3.9 : Conception du système de reconnaissance de visage avec la méthode hybride DCT et réseau de neurones.....	41
Figure 3.10 : Normalisation d'une image.....	42
Figure 3.11 : Egalisation de l'histogramme	42
Figure 3.12 : Extraction des caractéristiques du visage par DCT.....	43
Figure 3.13 : découpage de l'image en blocs.....	44
Figure 3.14 : Image sans découpage en blocs.....	44
Figure 3.15 : Extraction des valeurs d'un bloc 8x8.....	44
Figure 3.16 : passage du domaine spatial au domaine fréquentiel par la DCT.....	45
Figure 3.17 : Parcours en ZIGZAG.....	45
Figure 3.18 : Organisme de l'apprentissage selon le modèle PMC proposé.....	47
Figure 3.19: Organisme de la reconnaissance de visage.....	51

Liste des tableaux

Tableau 2.1 : Comparaison des méthodes basées sur les caractéristiques locales et globales.....32

Tableau 4.1 : Caractéristiques du matériel utilisé..... 54

Tableau 4.2 : Tests réalisés sur les différentes tailles de blocs.....62

Tableau 4.3 : Tests réalisés sur le nombre de coefficients DCT.....62

Tableau 4.4:Taux de reconnaissance obtenus en fonction du nombre des itérations.....63

Tableau 4.5 : Taux de reconnaissance en fonction de nombre des neurones de la couche.....63

Liste des abréviations

ACI: L'analyse en composantes indépendantes.

ACP: l'analyse en composantes principales.

ADF: Analyse discriminante de Fischer.

CMC: Cumulative Match Characteristic.

DCT: transformée en cosinus discrète.

EER: Le taux d'égale erreur.

FAR: Le taux de fausse acceptation.

FRR: Le taux de faux rejet.

KLT: Karhmen-loeve Transform.

LDA: Analyse discriminante linéaire.

PMC: Perceptron Multi-Couche.

RNA: Réseau de neurones artificiels.

ROC: Receiver Operating Characteristic.

Introduction Générale

La vision joue un rôle primordial dans notre vie, en effet c'est la meilleure façon de comprendre le monde extérieur. L'homme est considéré comme un expert dans la vision, il peut facilement voir et reconnaître les objets et les personnes qu'il voit et avec une très grande précision. Cette capacité a permis à plusieurs chercheurs de s'en inspirer pour réaliser des systèmes et créer des machines capables de " voir ", d'où vient la naissance de la vision par ordinateur.

La vision par ordinateur est devenue depuis plusieurs années un domaine très actif, ces recherches se basent sur la simulation du processus de la vision humaine pour la création des systèmes de vision automatiques. Ainsi, des caméras sont utilisées pour observer des scènes qu'elles reproduisent sous la forme images ou vidéos. Ce domaine inclut également le traitement et l'analyse qui est effectué sur les données brutes.

La biométrie est une branche de la vision par ordinateur, qui vise à établir l'identité d'une personne en mesurant l'une de ses caractéristiques physiques. La reconnaissance de visages tient une place centrale en biométrie. Cet intérêt est motivé par plusieurs raisons. D'abord, le visage est universel.

La reconnaissance faciale est une tâche particulièrement difficile et généralement considérée comme un problème ouvert. Elle possède de nombreuses applications dans les domaines de sécurité informatique, de multimédia et aussi en sciences cognitives.

Plusieurs algorithmes de reconnaissance de visage ont été proposés ces dernières années. Chacun repose sur une représentation particulière du visage. Nous pouvons identifier trois types d'approches: l'approche globale dans laquelle l'image du visage est considérée comme un vecteur de caractéristique, l'approche locale qui consiste à appliquer des transformations à des endroits spécifiques de l'image, tel que les coins des yeux, de la bouche, le nez, et l'approche hybride qui combine les deux.

Malgré la multitude des approches et des méthodes qui ont été proposées pour résoudre le problème de vérification automatique de visage humains, il demeure un problème extrêmement difficile, ceci est dû au fait que le visage de personnes différentes ont globalement la même forme alors que les images d'un même visage peuvent fortement varier du fait des conditions d'éclairage, de la variation de posture, des expressions faciales.

L'objectif de ce travail est la mise au point d'un système de reconnaissance de visage basée sur l'hybridation de deux méthodes globale de reconnaissance de visage. La transformée en cosinus discrète et les réseaux de neurones. La première nous permet de la réduction des dimensions toute en gardant un maximum d'information, alors que la deuxième sera utilisée dont le but d'effectuer l'apprentissage et la décision du système.

Notre mémoire s'articule autour de quatre chapitres, organisés comme suit :

Chapitre 1

Il est consacré à la présentation générale de la biométrie, A travers ce chapitre, nous voulons positionner le problème de la reconnaissance faciale et présenter ses enjeux et intérêts par rapport aux autres techniques.

Chapitre 2

Présente une analyse détaillée des différentes techniques développées au cours de ces dernières années dans les domaines de la détection de visage et de la reconnaissance. Notre objectif est de classifier ces méthodes de manière efficace afin de mettre en évidence les particularités ainsi que les avantages et les inconvénients de chacune d'entre elles.

Chapitre 3

Dans ce chapitre, nous nous intéressons aux méthodes utilisées, ainsi les différentes étapes de conception de notre système.

Chapitre 4

Ici, nous abordons la réalisation du projet, son implémentation, les tests réalisés et leur interprétation.

Chapitre 01 :

Généralités sur la reconnaissance faciale

I. Introduction

La grande disponibilité des ordinateurs puissants et des systèmes informatiques ont suscité un énorme intérêt dans le traitement automatique des images et des vidéos numériques au sein de nombreuses applications, incluant l'identification biométrique, la surveillance, l'interaction homme-machine et la gestion des données multimédia.

La reconnaissance faciale, en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de la recherche, ceci étant dû aux avancées rapides dans les technologies existantes.

II. La vision par ordinateur

II.1. Définition

La vision par ordinateur appelée aussi vision artificielle, vision numérique ou récemment vision cognitive, c'est une branche de l'intelligence artificielle, son but est de permettre à une machine de comprendre ce qu'elle 'voit' lorsqu'elle est connectée à une ou à plusieurs cameras, en d'autres termes essayer de modéliser et reproduire la vision humaine[1].

II.2. Catégories de la vision par ordinateur

La vision artificielle se divise en deux catégories [2]:

- **Le traitement d'images bas niveau** : Elle est représentée par les applications suivantes : le suivi des formes, la détection des contours, la segmentation, la super résolution, la restauration, la transmission et la compression.
- **Le traitement d'image de haut niveau** : Elle est représentée par des techniques pour l'interprétation, la reconstruction et la compréhension d'une scène 3D à partir d'une ou plusieurs images 2D : compréhension de l'image, reconnaissance des formes et stéréovision.

III. L'image numérique

L'acquisition d'une image se fait par moyen d'un appareil photo ou une prise d'un film sur une caméra. La nature brute de l'image à la sortie des capteurs est un signal analogique qui doit être numérisé pour pouvoir l'exploiter par un processeur de données numériques. La source peut être aussi une image artificielle, dans ce cas, le signal de l'image est d'origine discrète.

III.1. Définition

Une image numérique est un signal bidimensionnel fini discret qui peut être assimilé tel vue sur un écran d'ordinateur à une matrice $I(M, N)$ finie où chaque élément $I(M_i, N_i)$ de cette matrice représente un point ou un élément de l'image appelé pixel, qui appartient à un espace de couleur fini [3].

III.2. Caractéristiques d'une image numérique

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants :

III.2.1. Pixel

Le pixel est le plus petit point de l'image, c'est une entité calculable qui peut recevoir une structure et une quantification. Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels d'affichage ou d'impression. La quantité d'information que véhicule chaque pixel donne des nuances entre images monochromes et images couleur. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image [4].

III.2.2. Dimension

C'est la taille de l'image, elle se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses. Le nombre de lignes de cette matrice multiplié par le nombre de colonnes donne le nombre total de pixels dans l'image [5].

III.2.3. 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 pixels par unité de mesure. Le terme résolution est utilisé aussi pour désigner le nombre total de pixels affichés horizontalement ou verticalement sur un moniteur [4].

III.2.4. 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 dans l'image. Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différents, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant [5] [6]. Il permet de donner un grand nombre d'informations sur la distribution des niveaux de gris et de voir entre quelles bornes est répartie la majorité des niveaux de gris dans les cas d'une image trop claire ou d'une image trop foncée.

III.2.5. Contours et textures

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative. Les textures décrivent la structure de ceux-ci. L'extraction de contours consiste à identifier dans l'image les points qui séparent deux textures différentes [4]

III.2.6. Luminance

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet [7].

III.2.7. Contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images [8].

III.2.8. Image en couleur

L'image en couleur est une image où chaque pixel est codé dans l'espace de couleur RGB (rouge, vert, bleu). 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 [3].

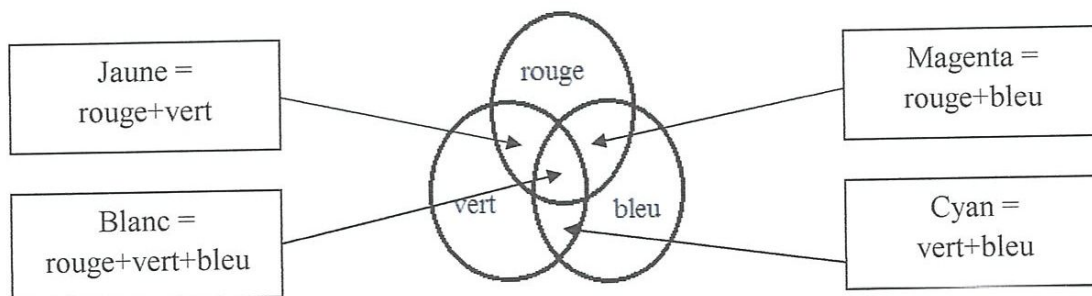


Figure 1.1 : Synthèse additive.

III.2.9. 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 [6].

III.3. Les types des images

La numérisation d'une image peut s'effectuer selon deux procédés différents de codage, ayant chacun leurs applications propres et produisant deux modes d'images : en pixel ou vectoriel.

III.3.1. L'image en pixel (Bitmap)

L'image bitmap est représentée par une trame de pixels. Ce ne sont pas des formules mathématiques qui définissent les formes, mais un ensemble de pixels qui agissent comme un tableau pointilliste [9].

Avantages

- Les images bitmaps peuvent facilement être créées et stockées dans un tableau de pixels représentant l'image.
- Lecture/écriture d'un pixel aisée de part la représentation de l'image comme une grille.
- Les images bitmaps peuvent facilement être affichées sur un écran ou être imprimées.

Inconvénients

- Fichiers volumineux.
- Ne conserve pas les proportions lors des modifications exemple : agrandissement ou réduction.
- Risque de disparition de pixels lors des modifications de taille.

III.3.2. L'image vectorielle

L'image vectorielle est une représentation conceptuelle de forme calculée par des formules mathématiques, par exemple, un cercle n'est pas déterminé par des pixels, mais par une formule mathématique qui détermine sa forme, sa taille et son emplacement [9].

Avantages

- Fichiers de volume peu important.
- Conservation des proportions lors des modifications de taille de l'image.
- Dessin aux contours nets.

Inconvénients

- Moindre possibilité de travail sur les couleurs.
- L'apparence d'une image vectorielle peut énormément varier en fonction du logiciel qui l'interprète ce qui dépend en fait des algorithmes utilisés pour l'affichage.
- La reconstitution d'une image à partir des vecteurs peut prendre beaucoup de temps qu'une image bitmap de complexité égale.

III.4. La compression des images

III.4.1. Définition

La compression des données ou le codage source, permet on appliquant des algorithmes spécifiques de décompression de réduire la taille d'une image sur une mémoire ou de manière équivalente de réduire son temps de transmission [10].

L'utilisation des données sous leurs formes numériques ne serait pas possible aujourd'hui sans la compression pour plusieurs raisons [11] :

- Les capacités de stockages des utilisateurs, même si elles ne cessent d'augmenter, ne sont pas infinies.
- La durée des transmissions de ces données numériques est conditionnée par le débit du réseau qui est utilisé et qui est parfois relativement faible.
- La compression des données numériques permet de diminuer la taille de stockages et de rendre possible leurs transports.

Image originale

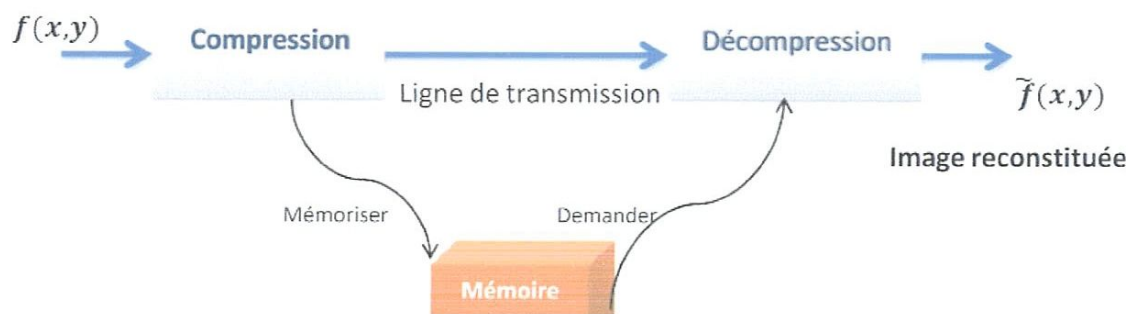


Figure 1.2 : Principe de la compression d'image.

III.4.2. Méthodes de compression

Les méthodes de compression d'images sont classées en deux catégories :

III.4.2.1. La compression sans perte

Restitue l'image originale après un cycle de compression / décompression. La compression sans pertes est très efficace avec les images 1 à 8 bits/pixel du type cartes, dessins et plans. Cette compression a une efficacité réduite pour les photographies 24 bits/pixel ou plus. Elle est utilisée pour les formats GIF, TIFF, PNG [11] [12].

III.4.2.2. La compression avec pertes

Élimine certaines informations qui ne sont pas indispensables. Les pertes dues à la compression se traduisent par des artefacts tels que du flou sur les transitions ou des distorsions géométriques ou des réductions des couleurs. La compression irréversible montre

toute son efficacité lorsqu'elle est appliquée aux images numérisées. Cette compression est utilisée pour les formats JPEG par exemple [12].

III.4.3. Rapport et taux de compression

Le rapport de compression est l'une des caractéristiques les plus importantes de toutes les méthodes de compression, il représente le rapport entre le nombre de bits avant le codage au nombre de bits après codage, il est défini par [13] [14]:

$$R = \frac{\text{nombre de bits avant compression}}{\text{nombre de bits après compression}} \dots\dots\dots (1.1)$$

Le taux de compression est un pourcentage de l'espace obtenu après la compression par rapport à l'espace total requis par les données avant la compression [15].

$$T = \left(1 - \frac{1}{\text{rapport de compression}}\right) * 100 \dots\dots\dots (1.2)$$

IV. La reconnaissance de visage dans une image fixe

L'être humain a des capacités naturelles pour reconnaître et différencier les visages, car notre cerveau est très évolué, peut mémoriser un nombre important d'individus avec les différentes caractéristiques. En revanche, pour une machine, la difficulté de reconnaissance de visage varie énormément suivant les différentes conditions d'acquisition de l'image. La reconnaissance de visage est classée parmi les principales techniques de la biométrie.

IV.1. Les systèmes biométriques

IV.1.1. Définition

La biométrie est toutes caractéristiques physiques ou traits personnels automatiquement mesurables, robustes et distinctives qui peuvent être utilisées pour identifier ou pour vérifier l'identité prétendue d'un individu [16].

Autre définition, la *biométrie* consiste à identifier une personne à partir d'une ou de plusieurs caractéristiques physiologiques ou comportementales.

Ces caractéristiques n'ont pas le même degré de fiabilité, mais ils doivent satisfaire les conditions suivantes [17]:

- **Universelles** : caractéristiques présentes chez tous les individus.
- **Unicité** : caractéristiques permettant de différencier un individu par rapport à tout autre.
- **Permanent** : présentes tout au long de la vie.
- **Collectables** : possibilité d'enregistrer les caractéristiques d'un individu avec son accord.

L'avantage principal de la biométrie est que chaque personne a ses propres caractéristiques qui ne peuvent pas être subtilisées ou détournées.

IV.1.2. Les technologies biométriques existantes

Il existe plusieurs techniques biométriques utilisées dans plusieurs applications et secteurs, et qui exploitent diverses informations, elles peuvent essentiellement se diviser en deux grandes catégories [18] :

- **Physiologiques** : empreintes digitales, visage, iris, contour de la main, etc.
- **Comportementales** : signature, dynamique de frappe au clavier.

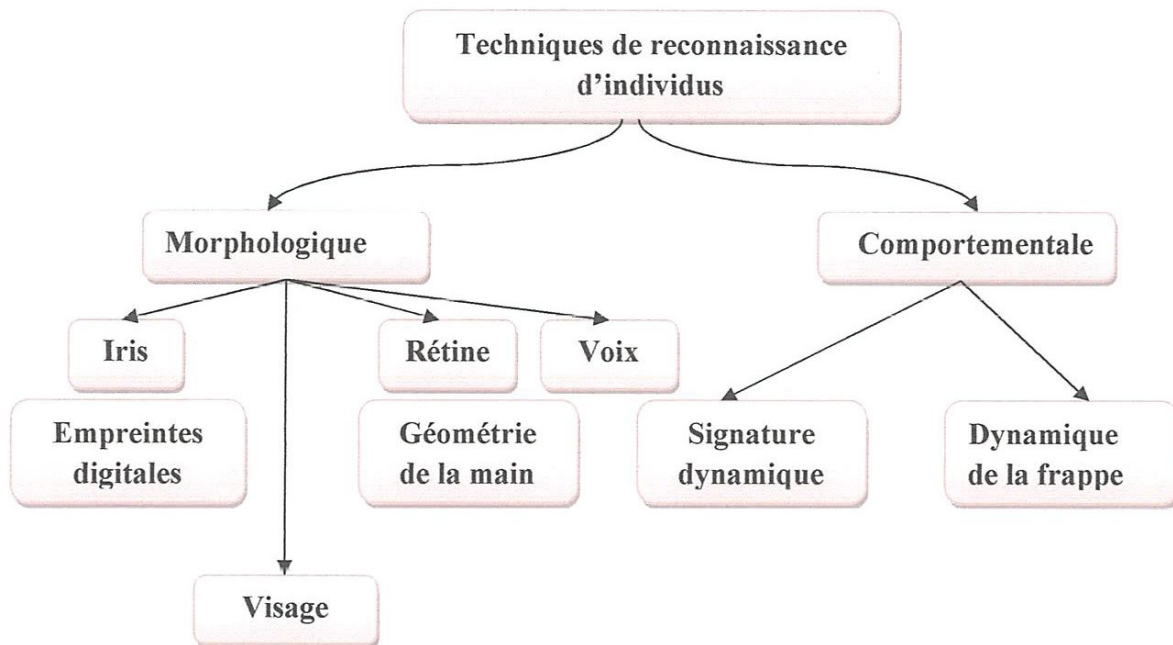


Figure 1.3 : Principales techniques de reconnaissance d'individus.

IV.1.2.1. Techniques morphologiques

- **Le visage** : La reconnaissance basée sur la vérification de l'identité, on compare les données caractéristiques extraites des visages des différentes personnes à partir de vidéo ou d'images fixes [19].

Avantages

- C'est une technique commune, populaire, simple.
- Ne nécessite pas des équipements chers.

Inconvénient

- Des éléments externes comme l'éclairage, le bruit, l'expression faciale et la posture dégradent le taux de reconnaissance.

- **L'iris** : La reconnaissance par iris est très utilisée dans les applications d'identification et de vérification. Le traitement s'effectue en trois phases [20] :
 - Recherche de la position de l'iris dans l'image de l'œil
 - Extraction des paramètres caractéristiques
 - Comparaison avec les éléments connus

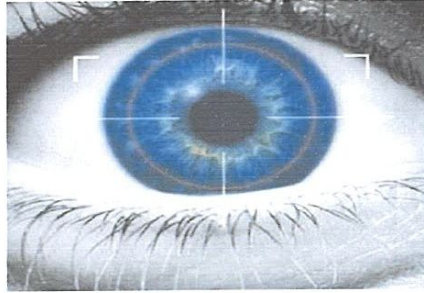


Figure 1.4 : Photo d'iris.

Avantages

- Fiable
- Pas de risque identifié sur la santé

Inconvénients

- Système intrusif mal accepté psychologiquement (hygiène, proximité de l'objectif)
- Contraintes d'éclairage
- **La rétine** : Cette technique mesure la forme de vaisseaux sanguins qui se trouvent dans l'arrière de l'œil. Le dispositif utilisé emploie une source lumineuse rayonnante en direction de l'œil de l'utilisateur, qui doit se tenir fixe devant le dispositif [21].

Avantages

- Résistant à la fraude, difficile et long à imiter
- Unicité même chez les vrais jumeaux
- Technique fiable
- La cartographie de la rétine est la même tout au long de la vie, en absence de maladie spécifique

Inconvénients

- Nécessité de placer les yeux à très faible distance du capteur, ce qui le rend un système intrusif mal accepté psychologiquement
- Couteuse
- Difficile à utiliser en cas de contrôle d'une population importante (temps important)

- **La voix :** La reconnaissance par voix utilise les caractéristiques vocales pour identifier les personnes en utilisant des phrases mot de passe « pass-phrase ». ces caractéristiques de la parole sont constituées par une combinaison des facteurs comportementaux (vitesse, rythme...) et physiologique (tonalité, âge, sexe, fréquence, accent, harmonique...) [22].

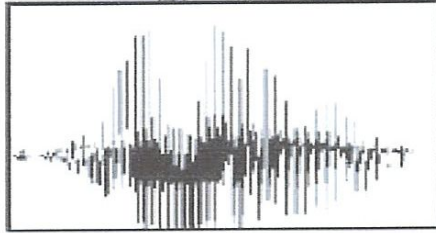


Figure 1.5 : Spectre d'un signal voix.

Avantages

- Technique très facile à mettre en œuvre et peu intrusive.
- Elle permet une reconnaissance à distance.

Inconvénient

- Technique très sujette à des variations volontaires (imitation) ou involontaires (maladie, émotion, vieillissement...) de la voix et très dépendante de la qualité de l'enregistrement.
- **L'empreinte digitale :** La reconnaissance par empreinte digitale se base sur le fait que chaque personne a ses empreintes uniques. après la capture de l'image de l'empreinte, on fait un rehaussement de l'image ensuite on identifie et on extrait les minuties, qui vont être comparées avec l'ensemble des minuties sauvegardées des autres utilisateurs [21] [22].



Figure 1.6 : Empreinte digitale.

Avantage

- Facilité de mise en place

Inconvénient

- Dépend d'une qualité optimale des appareils de mesure
- **La géométrie de la main :** La géométrie de la main ou du doigt est une mesure automatisée de plusieurs dimensions, notamment la largeur de la main et des doigts et la longueur des doigts [20].

Avantages

- Technologie rapide et bien développée et qui est facilement acceptée par les utilisateurs.
- Faible volume de stockage par fichier

Inconvénients

- Elle offre un taux d'erreur relativement haut et elle n'est pas utilisable avec des personnes jeunes ou âgées
- Cette technique n'a pas évolué depuis plusieurs années

IV.1.2.2. Techniques comportementales

• La signature

La signature est saisie avec un stylo optique sur une tablette graphique. Les paramètres qui caractérisent sa dynamique sont enregistrés : inclinaison du stylo, vitesse et accélération de l'écriture, pression exercée, nombre de fois que le stylo se soulève. Ces informations sont stockées sous forme de fichier pour être comparées lors de la reconnaissance [19] [20].

Avantage

- Technique repose sur un geste naturel facile à faire et à enregistrer.

Inconvénient

- La signature n'est pas parfaitement reproductible, elle peut varier suffisamment d'un jour à l'autre.

• Dynamique de la frappe au clavier

La dynamique de la frappe au clavier c'est une caractéristique de l'individu, c'est la transposition de la graphologie aux moyens électroniques.

Les paramètres suivants sont généralement pris en compte [22] :

- Vitesse de frappe
- Suite de lettres
- Mesure de temps de frappe
- Pause entre chaque mot

Avantage

- Moyen non intrusif qui exploite un geste naturel

Inconvénient

- Dépendance de l'état physique de la personne

IV.1.3. Architecture d'un système biométrique

Un système biométrique typique peut être représenté par les principaux modules suivants [18] [20] [22] :

- **Module de capture :** Responsable de l'acquisition des données biométriques d'un individu.
- **Module d'extraction de caractéristiques :** Prend en entrée les données biométriques acquises par le module de capture et extrait seulement l'information pertinente afin de former une nouvelle représentation des données.
- **Module de correspondance :** Compare l'ensemble des caractéristiques extraites avec le modèle enregistré dans la base de données du système et détermine le degré de similitude entre les deux.
- **Module de décision :** Vérifie l'identité affirmée par un utilisateur ou détermine l'identité d'une personne basée sur le degré de similitude entre les caractéristiques extraites et les modèles stockés.
- **Module base de données :** Stocke les modèles biométriques des utilisateurs enrôlés.

IV.1.4. Modes de fonctionnement d'un système biométrique

Les systèmes biométriques peuvent fournir trois modes de fonctionnement, à savoir : l'enrôlement, l'authentification ou vérification et l'identification.

➤ L'enrôlement

C'est la première phase de tout système biométrique, il s'agit de l'étape pendant laquelle un utilisateur est enregistré dans le système pour la première fois où une ou plusieurs modalités biométriques sont capturées et enregistrées dans une base de données. Cet enregistrement peut s'accompagner par l'ajout d'information biographique dans la base de données [18].

➤ Authentification (vérification)

Dans une application de vérification, l'utilisateur annonce son identité par l'intermédiaire d'un mot de passe, d'un numéro d'identification, d'un nom d'utilisateur, ou toute la combinaison des trois [20] [22].

Le système sollicite également une information biométrique provenant de l'utilisateur et compare les données caractéristiques obtenues à partir de l'information entrée, avec la donnée enregistrée correspondante à l'identité prétendue.

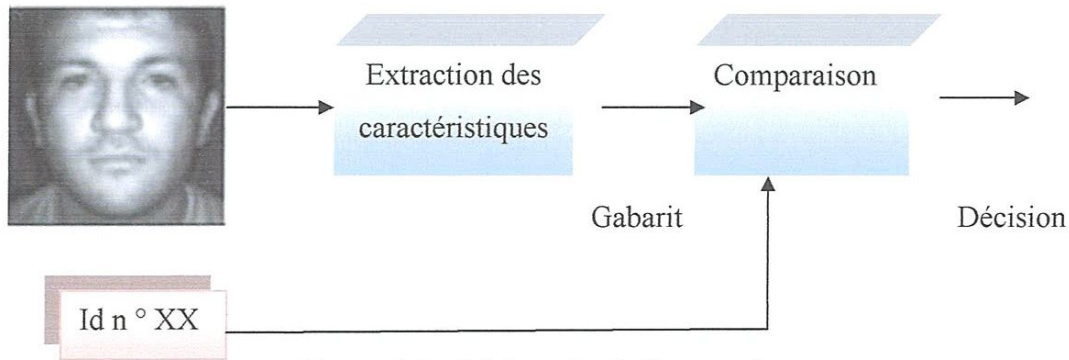


Figure 1.7 : Schéma de vérification du visage.

➤ Identification

C'est la reconnaissance à partir d'une base de données sans autre connaissance à priori de l'identité de la personne [18].

L'utilisateur ne dévoile pas explicitement son identité. Cependant, l'affirmation implicite faite par l'utilisateur est qu'elle est une des personnes déjà enrôlées par le système. Ainsi, l'échantillon biométrique de l'individu est comparé avec les modèles de toutes les personnes de la base de données [23].

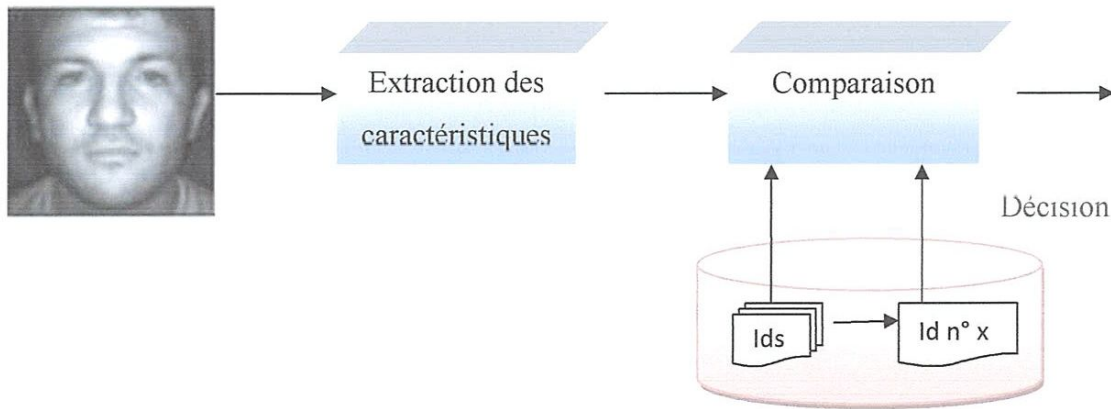


Figure 1.8 : Schéma d'identification du visage.

IV.1.5. Mesure de performances d'un système biométrique

Afin de mesurer la fiabilité d'un système biométrique en modes de vérification et/ou d'identification, trois critères principaux sont définis pour mesurer la fiabilité d'un système biométrique: le taux de faux de rejet, le taux de fausse acceptation, le taux d'égale erreur.

- **Le taux de faux rejet** ("False Reject Rate" ou **FRR**)

Ce taux représente le pourcentage de personnes censées être reconnues, mais qui sont rejetées par le système, il est défini par [22]:

$$FRR = \frac{\text{nombre de faux rejets}}{\text{nombre total d'accès client}} \dots\dots\dots (1.3)$$

- **Le taux de fausse acceptation** (« False Accept Rate » ou **FAR**).

Ce taux représente le pourcentage de personnes censées ne pas être reconnues, mais qui sont tout de même acceptées par le système, défini par [23] :

$$FAR = \frac{\text{nombre de fausses acceptations}}{\text{nombre total imposteurs}} \dots\dots\dots (1.4)$$

- **Le taux d'égale erreur** ("Equal Error Rate" ou **EER**)

Ce taux est calculé à partir des deux premiers critères et constitue un point de mesure de performance courant. Ce point correspond à l'endroit où $FRR = FAR$, c'est-à-dire le meilleur compromis entre les faux rejets et les fausses acceptations [22].

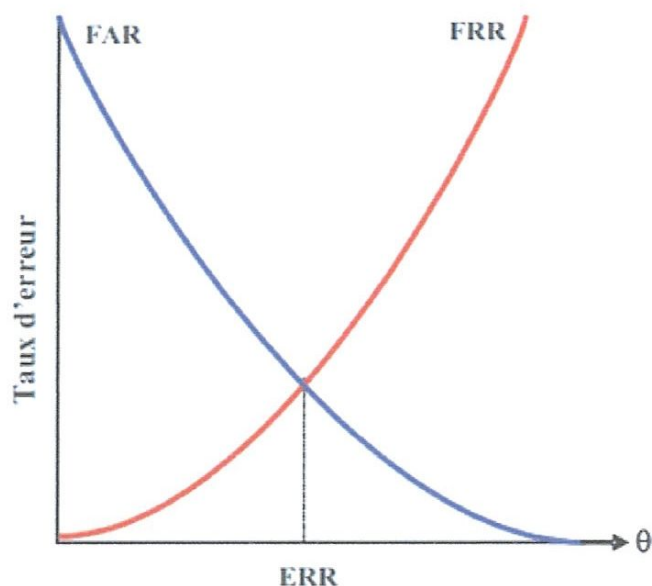


Figure 1.9 : Courbe du point d'équivalence des erreurs dans un système biométrique

Selon la nature d'authentification ou d'identification du système biométrique, il existe deux façons d'en mesurer la fiabilité :

- **En mode authentification :** On utilise une **courbe ROC** « Receiver Operating Characteristic ». La courbe ROC trace le taux de faux rejet en fonction du taux de fausse acceptation [24]. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé.

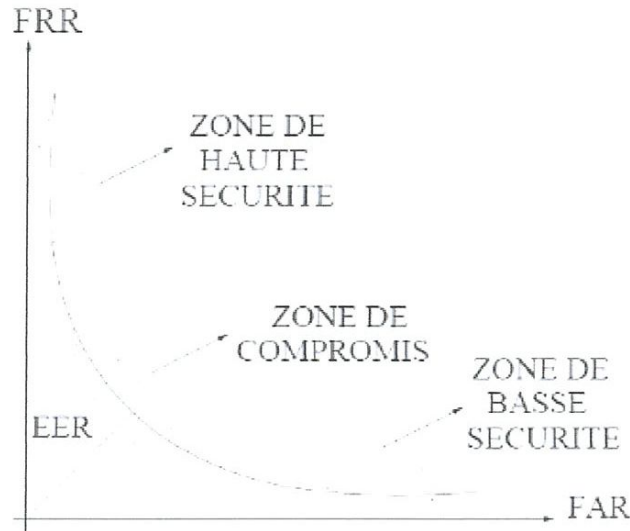


Figure 1.10 : Courbe ROC.

- **En mode identification :** On utilise une **courbe CMC** « Cumulative Match Characteristic ». La courbe CMC donne le pourcentage de personnes reconnues en fonction d'une variable appelée le rang [24] :
 - ✓ Un système est reconnaît au rang 1 lorsqu'il choisit la plus proche image comme résultat de la reconnaissance.
 - ✓ Système est reconnaît au rang 2, lorsqu'il choisit, parmi deux images, celle qui correspond le mieux à l'image d'entrée, etc.

Plus le rang augmente, plus le taux de reconnaissance correspondant est lié à un niveau de sécurité faible.

IV.1.6. Vision futur de la biométrie

Les systèmes biométriques utilisant une seule biométrie ont trois limitations principales :

- Une limitation en termes de performances en reconnaissance
- Une limitation en termes d'universalité d'utilisation
- Une limitation en termes de détection des fraudes.

Ces limitations peuvent être réduites, voire supprimées par l'utilisation conjointe de plusieurs systèmes biométriques formant alors un système biométrique multimodal. Les systèmes multimodaux permettent d'améliorer les performances de reconnaissance en combinant les décisions de plusieurs systèmes. Ils permettent également de régler le problème de non-universalité de certaines biométries en proposant une alternative aux personnes ne pouvant utiliser certaines biométries. Et enfin, ils permettent de limiter les possibilités de

fraude, car il est plus difficile d'obtenir et de reproduire plusieurs caractéristiques qu'une seule.

IV.2. Systèmes biométriques basés sur la reconnaissance faciale

Les différentes technologies de la reconnaissance des individus telles que : l'iris, la rétine...Etc. ont montré leurs limites.

Toutes fois, ces méthodes présentent l'inconvénient majeur d'être intrusive, ce qui limite énormément leurs domaines d'application.

Par contre, les technologies de reconnaissance basées sur le visage possèdent plusieurs avantages par rapport aux autres techniques, elle est :

- Naturelle
- Non intrusive
- Facile à utiliser

La reconnaissance faciale en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de recherche.

IV.2.1. Descripteurs et caractéristiques du visage dans une image

Pour que deux images des visages soient comparées, il faut tout d'abord trouver une caractéristique pour décrire chacune d'elle. Les caractéristiques les plus couramment utilisées sont classées en deux approches:

- **Approche basée sur les caractéristiques invariables** : L'objectif de ces méthodes est de trouver les caractéristiques du visage dans une image, tel que : les traits du visage, la couleur de la peau [25].



Figure 1.11: Exemple d'extraction des contours du visage.

- **Approche basée sur la mise en correspondance** : Cette approche consiste à construire des modèles de visage qui représentent des classes de visages [26]. Le problème de l'identification de visage est résolu par une comparaison entre l'ensemble complet des

sous-images de l'image traitée et les patterns de visage. Bien que les visages aient la même structure, ils peuvent être différents en taille, en termes d'orientation.

IV.2.2. Processus de reconnaissance de visage

Un schéma général utilisé en reconnaissance de visage est illustré en *Figure 1.12* [26] [27] [28] :

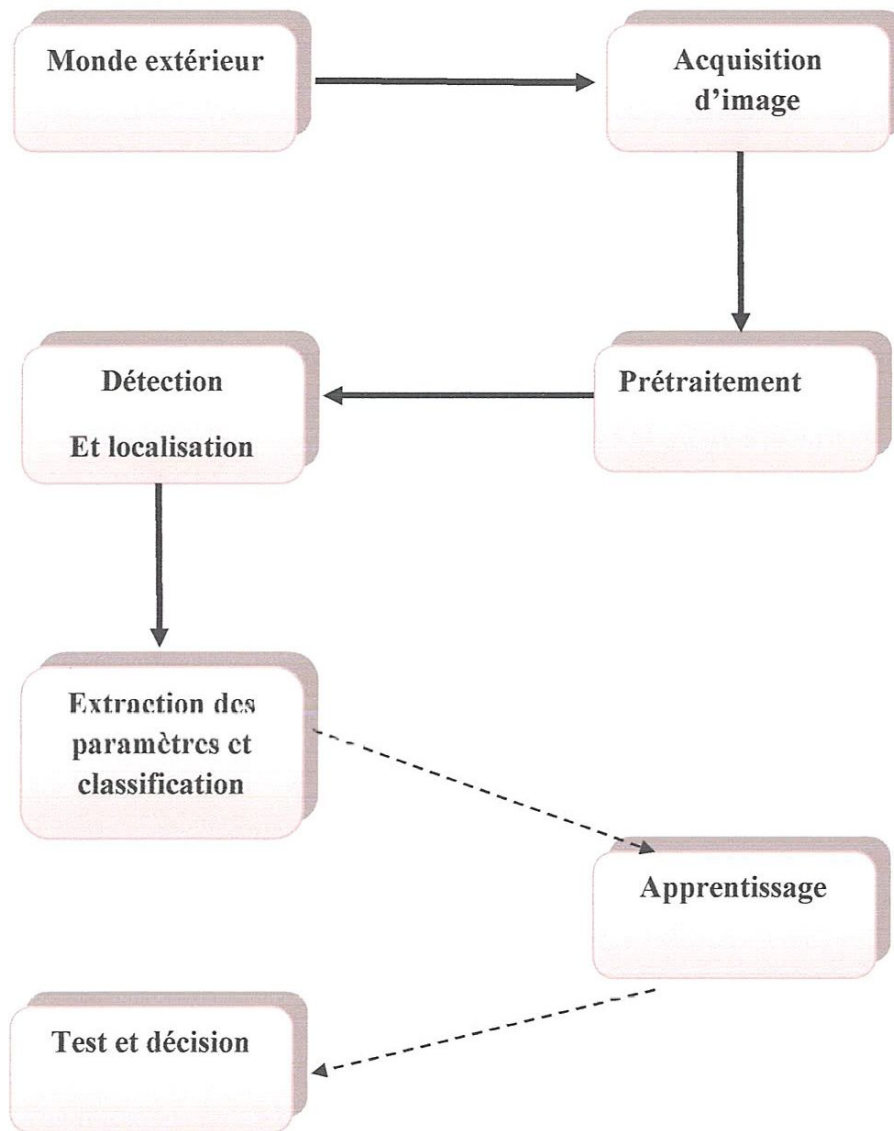


Figure 1.12 : Processus de reconnaissance du visage.

IV.2.2.1 Acquisition de l'image

Cette étape donne lieu à une représentation 2D (la matrice des niveaux de gris) pour un objet 3D (le visage), l'acquisition de l'image et sa digitalisation comporte un risque de bruit.

Cette opération peut être statique : Appareil Photo, Scanner... etc. ou dynamique : Caméra, Webcam, dans ce cas on aura une séquence vidéo. À ce niveau on aura une image brute.

IV.2.2.2. prétraitement

Permet d'éliminer le bruit, il consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image. Cette opération est très complexe, surtout dans le cas où l'image contient plusieurs visages ou le cas où l'arrière-plan n'est pas neutre. Pour pallier à ces problèmes, il existe plusieurs méthodes de traitement et d'amélioration des images, telles que : la normalisation, l'égalisation de l'histogramme, etc.

IV.2.2.3. Détection du visage

C'est un problème de classification où on assigne l'image à la classe visage ou à la classe non visage. Dans l'étape de détection, on identifie et on localise le visage dans l'image acquise au départ, indépendamment de la position, l'échelle, l'orientation et l'éclairage...

Son efficacité a une influence directe sur les performances du système de reconnaissance de visages.



Figure 1.13 : Exemple de détection du visage.

IV.2.2.4. 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, les individus ayant des similarités sont regroupés dans la même classe. Ces classes varient selon le type de décision.

IV.2.2.5. 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.

IV.2.2.6. Test et décision

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

IV.2.3. Difficultés de la reconnaissance de visage

Tout processus automatique de reconnaissance de 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. Voici les différentes difficultés de reconnaissance de visage :

IV.2.3.1. Changement d'illumination

L'apparence d'un visage dans une image varie énormément en fonction de l'illumination de la scène lors de la prise de vue. Les variations d'éclairage rendent la tâche de reconnaissance de visage très difficile [29][30]. En effet, le changement d'apparence d'un visage dû à l'illumination se révèle parfois plus critique que la différence physique entre les individus, et peut entraîner une mauvaise classification des images d'entrée.



Figure 1.14 : Exemples de variation d'illumination de la base YaleB.

IV.2.3.2. Variation de pose

La variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Quand le visage est de profil dans le plan image (orientation $< 30^\circ$), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux). Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible [31].

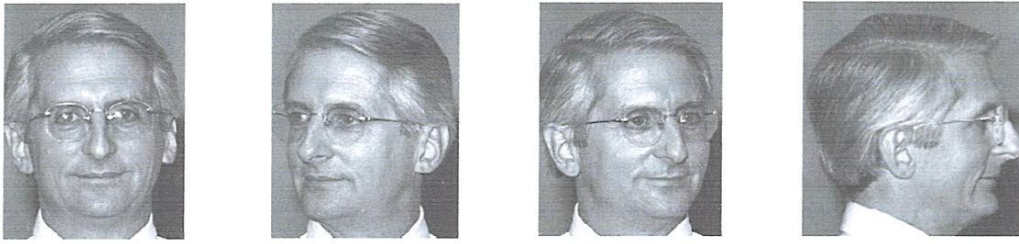


Figure 1.15 : Exemples de variation de poses.

IV.2.3.3. Expressions faciales

La déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. L'information faciale se situant dans la partie supérieure du visage reste quasi invariable, Elle est suffisante pour effectuer une identification. Toutefois, étant donné que l'expression faciale modifie l'aspect du visage, elle entraîne forcément une diminution du taux de reconnaissance [32]. L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui reste non résolu.



Figure 1.16 : Exemples de variation d'expressions.

IV.2.3.4. Présence ou absence des composants structurels

La présence des composants structurels telle que la barbe, la moustache, ou bien les lunettes peuvent modifier énormément les caractéristiques faciales telles que la forme, la couleur, ou la taille du visage. De plus, ces composants peuvent cacher les caractéristiques faciales de base causant ainsi une défaillance du système de reconnaissance [29].



Figure 1.17 : Exemple la présence des composants structurels

IV.2.3.5. Occultations partielles

Les visages peuvent être partiellement occultés par d'autres objets. En effet, dans une image contenant un groupe de personnes par exemple, des visages peuvent partiellement masquer d'autres visages [23] [29].

V. Conclusion

Dans ce chapitre, nous avons présenté les technologies utilisées dans les systèmes biométriques pour l'identification des personnes. Nous avons aussi donné un aperçu sur les techniques de mesure de leurs performances. Cette étude nous a permis de constater que la reconnaissance de visage suscite de plus en plus l'intérêt de la communauté scientifique, car elle présente plusieurs challenges et verrous technologiques.

Enfin, nous avons mis en évidence le processus de la reconnaissance automatique de visage, ainsi les différentes difficultés liées à cette tâche. Les techniques utilisées aux différentes étapes de la reconnaissance faciale seront détaillées dans le chapitre suivant.

Chapitre 02 :

Techniques de reconnaissance faciale

I. Introduction

Le visage constitue une catégorie de stimulus unique par la richesse des informations qu'ils véhiculent, ils sont à la fois des vecteurs visuels principaux de l'identité individuelle, et des vecteurs essentiels de communication, d'intentions et d'émotions entre les individus, via, en particulier, la direction du regard et les expressions faciales.

La reconnaissance de visage constitue l'une des capacités les plus performantes du système visuel humain. Le but d'un système de reconnaissance de visage est de stimuler le système de reconnaissance humain par la machine pour automatiser certaines applications telles que : la télésurveillance, le contrôle d'accès à des sites, l'accès à des bâtiments sécurisés, etc. Malgré la multitude des méthodes proposées et les progrès réalisés, le problème reste non résolu parfaitement.

II. Les étapes de reconnaissance du visage

La reconnaissance automatique de visage s'effectue en trois étapes principales :

- détection du visage.
- extraction et des caractéristiques du visage.
- reconnaissance du visage (identification et/ou vérification).

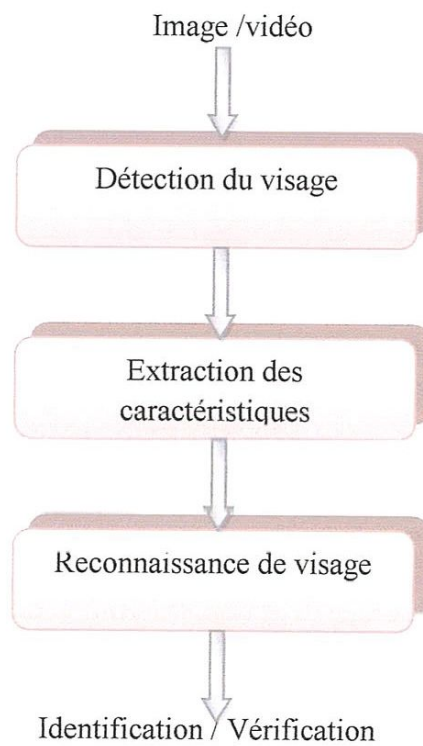


Figure 2.1 : Les étapes de la reconnaissance de visage.

II.1. la détection de visage

II.1.1. Définition

La détection de visage est une technologie informatique qui détermine les emplacements et les tailles des visages humains dans une image numérique, ou une vidéo. Le système détecte le soin du visage représenté et ignore autre chose, comme des constructions, des corps...Etc. [33].

II.1.1.1 Approches de détection du visage

Les approches de détection de visages peuvent être subdivisées en quatre catégories [2] [24] [34] :

- Approche basée sur les connaissances acquises
- Approches basées sur la mise en correspondance
- Approches basées sur l'apparence
- Approches basées sur des caractéristiques invariables

II.1.1.1.1. Approche basée sur les connaissances acquises

Cette approche s'intéresse aux parties caractéristiques du visage comme le nez, la bouche et les yeux. Elle est basée sur la définition des règles strictes à partir des rapports entre les caractéristiques faciales. Ces méthodes sont conçues principalement pour la localisation de visage.

II.1.1.1.2. Approches basées sur la mise en correspondance

Cette approche considère le problème de détection du visage comme un problème de mise en correspondance des formes : on vérifie à chaque endroit de l'image si une fenêtre est un candidat de visage, en comparant la différence entre celle-ci et les modèles de visage. Malgré le fait que les visages ont la même structure, ils peuvent être en différentes tailles et positions. La mise en correspondance est basée sur des fonctions de corrélations croisées de deux dimensions entre une fenêtre de l'image et le modèle. Des modèles de visage normalisés sont soit prédéfinis manuellement par un spécialiste, soit paramétrés automatiquement par des fonctions.

II.1.1.1.3. Approches basées sur l'apparence

L'idée principale de ces méthodes est d'utiliser une procédure de formation qui classe les modèles en deux classes :

- classe de visages
- classe de non visages

Ces approches appliquent généralement des techniques d'apprentissage automatique. Ainsi, les modèles sont appris à partir d'un ensemble d'images représentatives de la variabilité de l'aspect facial. Ces modèles sont employés pour la détection.

II.1.1.1.4. Approche basée sur des caractéristiques invariables

Cette approche a pour objectif de trouver les caractéristiques structurelles même si le visage est dans différentes positions, conditions lumineuses ou changement d'angle de vue.

Il existe principalement trois familles d'approches basées sur les caractéristiques invariables :

- La première utilise la propriété de la peau humaine pour capter des régions contenant des visages.
- La deuxième famille vise à détecter les composantes faciales.
- La troisième combine une multitude de caractéristiques de natures différentes.

II.2. l'extraction de signatures du visage

II.2.1. Définition

L'extraction de signatures est une phase d'importance circulaire dans un système de reconnaissance faciale. C'est un ensemble des caractéristiques de visage, ces dernières doivent être très différentes d'une personne à une autre et invariables pour les images d'une même personne [29]. La définition de la signature est un problème très complexe, du fait que les images d'une même personne peuvent être différentes en fonction des expressions faciales et/ou des différentes conditions de prise de vue.

Les facteurs modifiant l'apparence de l'image du visage d'une même personne sont qualifiés d'intrinsèques ou d'extrinsèques [35] :

- **Les facteurs intrinsèques** : Sont des facteurs indépendants du contexte dans lequel évolue la personne. Il s'agit essentiellement des changements de son expression faciale lors de la prise de vue, et ce pour différentes raisons : les émotions, le bavardage, l'ouverture de la bouche, etc.
- **Les facteurs extrinsèques** : Sont des facteurs ayant un effet sur l'apparence du visage et indépendants de l'état émotionnel de la personne. Ces facteurs peuvent correspondre à un masquage de traits importants du visage, ou à une modification des conditions de prise de vue, par exemple lorsque la source d'éclairage est modifiée.

La figure 2.2 présente trois différentes conditions de prise de vue qui fournit des images très différentes d'une même personne :



Figure 2.2 : Visage de la première personne de la base de données Yale Faces.

II.2.2. Méthodes d'extraction des caractéristiques de visage

On peut distinguer deux pratiques différentes [35] [36] :

II.2.2.1. Méthode statistique

Extrait des régions entières du visage, consiste à appliquer une transformation sur les données disponibles pour chaque personne. Cette transformation permettant de conserver certaines propriétés de la distribution initiale des valeurs des pixels. La transformation vise à réduire la dimension de l'espace contenant les vecteurs représentant une image de visage. Ainsi, la signature associée à un visage présente le nouveau vecteur résultant de la transformation appliquée à la donnée initiale.

II.2.2.2. Méthode géométrique

Extrait des points particuliers des différentes régions caractéristiques du visage : tels que les coins des yeux de la bouche et du nez. Elle est utilisée avec une méthode locale de reconnaissance faciale.

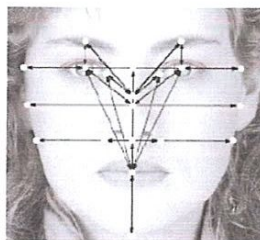


Figure 2.3 : Visage accompagné de mesures géométriques.

Il existe plusieurs techniques de prise de mesures, mais leur point commun est qu'elles se concentrent toutes sur la mesure des trois organes importants, à savoir : les yeux, le nez et la bouche.

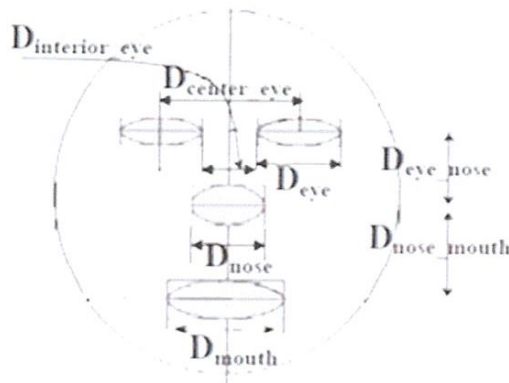


Figure 2.4 : Mesures géométriques réalisées sur le visage.

II.3. La reconnaissance de visage

II.3.1. Définition

L'étape de la reconnaissance faciale utilise les caractéristiques du visage extraites pour créer et stocker une signature dans une base de données. Ainsi, à chaque visage de la base est associée une signature unique qui caractérise la personne correspondante. La reconnaissance d'un visage requête est obtenue par l'extraction de la signature requête correspondante et sa mise en correspondance avec la signature la plus proche dans la base de données. La reconnaissance dépend du mode de comparaison utilisé : vérification ou identification [27].

II.3.2. Les méthodes de reconnaissance de visage

II.3.2.1. Les méthodes globales

II.3.2.1.1 Définition

Les méthodes globales sont basées sur des techniques d'analyse statistique bien connues. Il n'est pas nécessaire de repérer certains points caractéristiques du visage : les centres des yeux, les narines, le centre de la bouche, etc. à part pour la normalisation des images. Dans ces méthodes, les images de visage, qui peuvent être vues comme des matrices de valeurs de pixels, sont traitées de manière globale et sont transformées en vecteurs, plus faciles à manipuler [36].

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

II.3.2.1.2. Les techniques linéaires

Le traitement consiste essentiellement à projeter l'espace d'entrée dans un autre espace de plus faible dimension et à ne conserver que les données considérées comme significatives selon des critères de sélection définis a priori [18].

En revanche, la dimension de l'espace final de représentation est fixée a posteriori, après la sélection. Les méthodes les plus connues de cette famille sont [35] [36] :

- La transformée en cosinus discrète
- l'Eigenface
- FisherFaces
- L'analyse en composantes indépendantes

a) la transformée en cosinus discrète

La transformée en cosinus discrète est une transformation mathématique complexe qui a pour but de transformer le domaine de représentation des données, c'est-à-dire de passer d'un domaine spatial à un domaine fréquentiel [35]. Le but de la transformée en cosinus discrète est de supprimer les hautes fréquences de l'image tout en gardant les données majeures qui sont représentées par les basses fréquences.

b) Eigenface

Eigenface est l'une des premières techniques fonctionnelles de reconnaissance de visages. Elle repose sur l'analyse en composantes principales (ACP) d'un ensemble des images de visage stockées dans une base de données [37]. L'ACP consiste à rechercher une transformation linéaire de l'espace initial dans lequel sont représentées les données aboutissant à un nouvel espace, dans lequel l'information est concentrée sur un nombre limité de directions orthogonales.

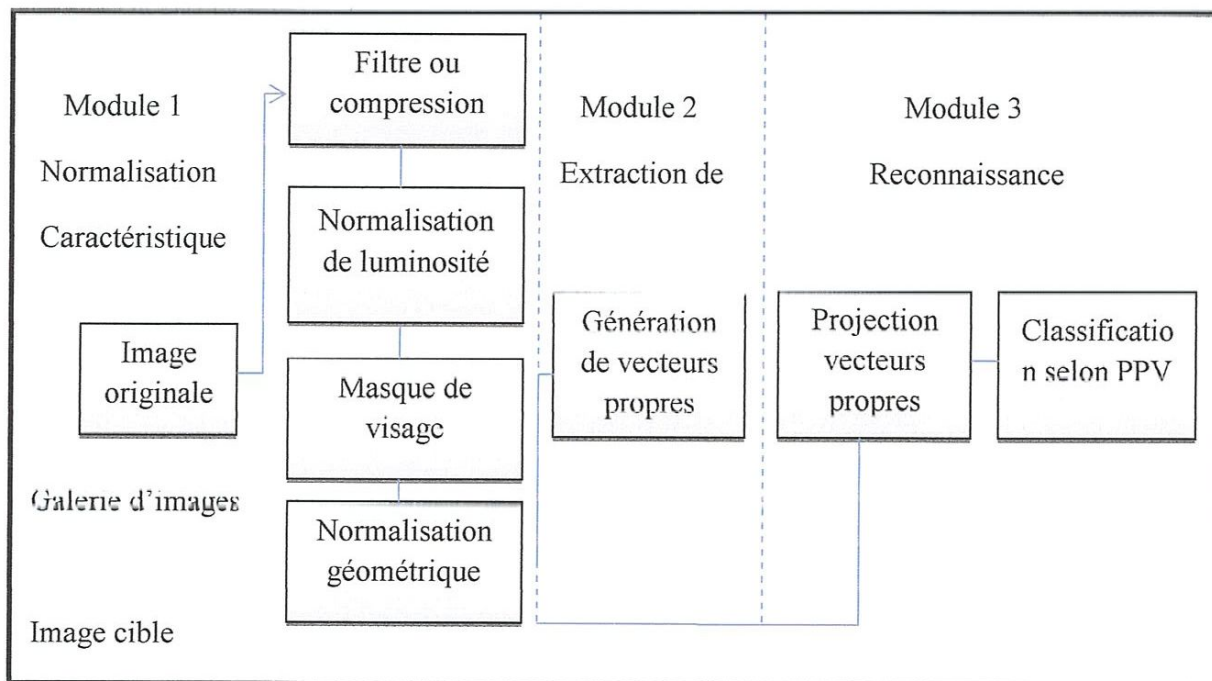


Figure 2.5 : Les différents modules d'un système de reconnaissance de visage basé sur l'ACP

c) FisherFaces

FisherFaces est une méthode similaire à l'Eigenface, dans le sens qu'elle vise à réduire la taille de l'espace contenant les images afin de constituer des signatures. Cependant, elle tend à construire un sous espace plus approprié pour différencier les différents individus connus dans la base initiale [38]. Les images de la base sont regroupées en classes, chacune regroupe toutes les images correspondant à un même individu. La méthode vise à déterminer une transformation linéaire pour chaque classe de la base. Dans l'espace résultant de cette transformation, les projections des visages de la classe seront clairement séparées par des projections des visages correspondantes à d'autres individus.

L'ADF est une méthode utilisée pour déterminer les transformations associées à toutes les classes de la base. Son principe consiste à projeter des vecteurs représentant les images sur un axe particulier de l'espace initial, après la projection, chaque vecteur se transforme en une valeur unique [39].

d) L'analyse en composantes indépendantes ACI

L'ACI a été introduite afin de trouver une solution au problème de séparation des sources lorsque la fonction de mélange F est inconnue. Le traitement consiste à extraire les composantes linéaires d'une observation multi variée afin qu'elles soient aussi indépendantes que possible [40] [41].

II.3.2.1.3. Les techniques non linéaires

Le traitement consiste à trouver des espaces de séparation capables de représenter les données sous la forme de classes indépendantes, chaque classe correspondant à un individu. Ces méthodes sont généralement itératives avec une dimension de l'espace de représentation fixée a priori [18]. L'exemple le plus courant c'est les réseaux de neurones. Cette technique est également utilisée dans le module de classification et de décision.

a) Réseaux de neurones

Un réseau de neurones artificiels RNA est l'une des techniques non linéaires d'extraction de caractéristiques qui ont été largement utilisées pour la reconnaissance de visages. Ces derniers ont été initialement inspirés de la physiologie du système nerveux si parfaitement créée et conçue [42].

Un neurone formel est une fonction algébrique non linéaire paramétrée et à valeurs bornées de variables réelles appelées entrées, qui tente de reproduire ce mode de fonctionnement [43].

La méthode d'extraction par un RNA présente des similarités avec celle mise en œuvre par les techniques d'ACI qu'il vise à déterminer une fonction de séparation. L'image constitue

l'observation, c'est-à-dire la seule information permettant de déterminer la fonction de séparation.

Dans le cas du RNA, la fonction de séparation est en fait codée dans les paramètres des neurones de la couche cachée. Les sorties de ces mêmes neurones constituent la signature de l'image placée en entrée. La couche de sortie permet de reconstruire une version approchée de l'image aussi proche que possible de celle placée en entrée. Cette couche remplit un rôle similaire à la fonction F de mélange des sources. Les neurones de la couche cachée opèrent ainsi une réduction de dimension et ceux de la couche de sortie une reconstruction de l'image à partir de la signature [44].

II.3.2.1.4. Évaluation des méthodes globales

L'avantage principal des méthodes globales est qu'elles sont relativement rapides à mettre en œuvre et que les calculs de base sont d'une complexité moyenne. En effet, elles préservent implicitement les informations de texture et de forme nécessaire pour la reconnaissance de visages. En revanche, elles sont très sensibles aux variations d'éclairément, de pose et d'expression faciale.

II.3.2.2. Les méthodes locales

II.3.2.2.1. Définition

Les méthodes locales, basées sur des modèles, utilisent des connaissances a priori que l'on possède sur la morphologie du visage et s'appuient en général sur des points caractéristiques de celui-ci [36]. Ces méthodes constituent une autre approche pour prendre en compte la non linéarité en construisant un espace de caractéristiques local et en utilisant des filtres d'images appropriés, de manière à ce que les distributions des visages soient moins affectées par divers changements. Les méthodes locales peuvent être classifiées en deux catégories :

a) les méthodes basées sur les caractéristiques locales

Ces méthodes utilisent la description locale calculée au voisinage de points d'intérêt. La détection de points d'intérêt est une première étape dans le processus de la mise en correspondance et de la reconnaissance [36].

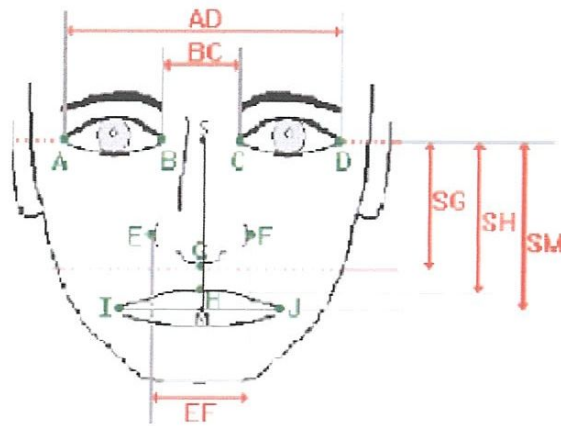


Figure 2.6 : Distance entre points caractéristiques d'un visage

b) les méthodes basées sur les apparences locales

Ces techniques sont utilisées de manière modulaire pour les différentes régions faciales. Un modèle global est défini à partir de la combinaison des différents modèles locaux. Ainsi, les différentes régions faciales ne seront plus affectées de la même manière par les différentes sources de variabilité [45] [46].

II.3.2.2. Évaluation des méthodes locales

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

II.3.2.3. Les méthodes hybrides

II.3.2.3.1. Définition

Les méthodes hybrides permettent d'associer les avantages des méthodes globales et locales en combinant la détection de caractéristiques géométriques avec l'extraction de caractéristiques d'apparence locales. Elles permettent d'augmenter la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales [47]. En effet, les caractéristiques locales et les caractéristiques globales ont des propriétés tout à fait différentes.

Le tableau 2.1 récapitule la différence entre les deux types de caractéristiques [48] :

Facteurs de variation	Caractéristiques locales	Caractéristiques globales
Illumination	Très sensible	Sensible
Expression	Pas sensible	Sensible
Pose	Sensible	Très sensible
Bruit	Très sensible	Sensible
Occlusion	Pas sensible	Très sensible

Tableau 2.1 : Comparaison des méthodes basées sur les caractéristiques locales et globales.

Les caractéristiques locales et globales réagissent différemment aux facteurs de variation. Par exemple, les changements d'illumination peuvent avoir plus d'influence sur les caractéristiques locales, tandis que les changements d'expression ont plus d'impact sur les caractéristiques globales [49] [50].

II.3.2.3.2. Évaluation des méthodes hybrides

Les méthodes hybrides peuvent constituer une approche efficace pour réduire la complexité des classifieurs et améliorer leur capacité de généralisation. Malgré ces avantages, les travaux de recherche dans ce domaine restent relativement peu nombreux, probablement en raison des difficultés mentionnées ci-dessus.

III. La méthode adoptée

Au vu de l'état de l'art et différentes approches de reconnaissance de visage, nous avons choisi une méthode hybride, combinant la transformée en cosinus discrète et les réseaux de neurones. L'avantage majeur de la DCT est qu'elle permet de réduire l'information redondante et peut être utilisée comme une étape d'extraction de caractéristiques.

De même, les réseaux de neurones sont des outils très robustes, ils ont été abondamment utilisés dans la reconnaissance des formes, reconnaissance de visage et la représentation des connaissances. L'emploi des réseaux de neurones peut se justifier par les arguments suivants :

- La tolérance par rapport aux fautes et robustesses par rapport à des défaillances internes du réseau.
- La capacité de classification et de généralisation.
- La capacité d'approximation universelle prouvée.

IV. Conclusion

Plusieurs algorithmes de reconnaissance de visage ont été proposés ces dernières années, chacun repose sur une représentation particulière du visage. Comme nous l'avons vu dans ce chapitre, on peut identifier trois types d'approches: l'approche globale dans laquelle l'image du visage est considérée comme un vecteur de caractéristique, l'approche locale qui consiste à appliquer des transformations à des endroits spécifiques de l'image, et l'approche hybride qui combine entre les deux.

Cette étude nous a permis de mettre en avant les critères de performances qui vont guider le choix la solution retenue dans le cadre de notre problématique, cette solution consiste à l'hybridation des deux méthodes globales à savoir : la transformée en cosinus discrète et les réseaux de neurones, qui font l'objet du chapitre suivant.

Chapitre 03 :

Conception

I. Introduction

L'objectif de ce travail est le développement d'une méthode hybride combinant la transformée en cosinus discrète et les réseaux de neurones. La DCT est exploitée afin d'extraire les caractéristiques décroissantes et non redondantes du visage, qui seront utilisées pour la classification. Pour réaliser cette tâche, nous avons choisi les réseaux de neurones multicouches comme classifieur.

II. Les méthodes développées

II.1. La transformée en cosinus discrète

II.1.1. Définition

La transformée en cosinus discrète est une transformation numérique appliquée à chaque bloc et pour chaque couleur. Cette transformée est une variante de la transformée de Fourier. Elle permet de décrire chaque bloc en une carte de fréquences et en amplitudes plutôt qu'en pixels et couleurs [17]. La valeur d'une fréquence reflète l'importance et la rapidité d'un changement, tandis que la valeur d'une amplitude correspond à l'écart associé à chaque changement de couleur.

La découverte de la DCT date de 1974 [49]. cela a constitué une avancée notable dans le domaine des traitements du signal, notamment après que furent publiés les travaux d'implémentation entrepris par Chen Smith et Fralick [50]. cette transformation est à la base de nombreuses techniques et technologies, notamment pour la compression avec perte des signaux temporels, des images fixes et vidéos.

II.1.2. Variantes de la DCT

Il existe 8 variantes de la DCT, ceux les plus connus sont DCTI, la DCT 2D ou DCT II et sa transformé inverse IDCT ou DCT III, qui sont la base de la compression JPEG [29]:

- **DCT I**

Pour un ensemble d'éléments discrets N , la DCT s'exprime par :

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \dots\dots\dots (3.1)$$

$$u = 0, \dots, N - 1$$

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{pour } u = 0 \\ \sqrt{2/N} & \text{pour } u \neq 0 \end{cases} \dots\dots\dots (3.2)$$

• **DCT II**

C'est une extension directe de la DCT I dans deux dimension $N \times N$, sa forme mathématique est donnée par :

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \dots\dots\dots (3.3)$$

Avec :

x et y : coordonnées de la matrice image

u et v : coordonnées de la matrice DCT

N : taille du bloc

• **IDCT**

IDCT ou DCT III c'est l'inverse de la DCT II, sa formule est donnée par:

$$f(x, y) = \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \dots\dots\dots (3.4)$$

II.1.3. Propriété de la DCT

• **Décorrélation**

La DCT permet de réduire la redondance inter pixel, les pixels sont décorrélés afin de réduire la variance des pixels voisins dans l'image ce qui permet d'une compression efficace en codant chaque pixel indépendamment.

• **Concentration des coefficients**

La DCT est très efficace pour des images fortement corrélées du fait qu'elle permet de compacter les coefficients qui représentent les basses fréquences dans une seule partition de la matrice image. Cela permet la séparation des fréquences basses et des fréquences hautes, et si une image est faiblement corrélée, les coefficients sont concentrés dans plusieurs différentes partitions de la matrice image.

• **Séparabilité**

La formule de la 2-D DCT peut être écrite de la façon suivante:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \cos\left(\frac{\pi(2x+1)u}{2N}\right) \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{\pi(2y+1)v}{2N}\right) \dots\dots\dots (3.5)$$

Cette propriété de séparabilité, a l'avantage de pouvoir calculer $C(u, v)$ en deux étapes par des exécutions successives de 1-D DCT sur les lignes puis les colonnes d'une image. Les arguments présentés peuvent être identiquement appliqués pour le calcul inverse de la DCT.

II.1.4. Application de la DCT à la problématique de reconnaissance faciale

En dépit des efforts employés pour développer des algorithmes rapides de calcul de la DCT, il a fallu attendre jusqu'à 2001 pour l'avoir appliquer à la problématique de reconnaissance de visage. Voici quelques utilisations de la DCT dans ce domaine :

La première utilisation de la DCT a été relatée pour comparer ses performances à celles obtenues avec la transformation KLT (Karhmenloeve Transform)[51].

Karhunen Loeve Transform(KLT) est une transformation linéaire qui diagonalise la matrice de covariance ou de corrélation d'une séquence de variables aléatoires discrètes.

D'autres études et comparaisons ont suivi pour analyser plus précisément ses performances vis-à-vis des variations géométriques et/ou d'illumination des visages durant l'acquisition des images [52].

Dans le même contexte, Yin et Coll [53] ont choisi de réduire les dimensions de l'image par une DCT suivie d'une compression des paramètres de la modélisation. Pour ce faire, seules les basses fréquences spatiales sont sélectionnées et analysées par ADF (Analyse discriminante de Fischer).

Une autre utilisation a été proposée par Abouzar et Coll [54], laquelle consiste à utiliser deux méthodes de transformation comme prétraitement avant la phase d'extraction de signature. Il s'agit de la transformée en ondelettes et de la DCT. La première sert à la réduction de dimensions et la deuxième à supprimer les fréquences élevées dans l'image de taille réduite.

Dans [55], il a été prouvé expérimentalement que la PCA et la LDA peuvent être appliquées sur un nombre réduits de coefficients DCT pour réaliser une meilleure reconnaissance avec un gain en temps de calcul et en espace mémoire.

II.2. Les réseaux de neurones

II.2.1. Définition

Un réseau de neurones est un processus distribué de manière massivement parallèle, qui a une propension naturelle à mémoriser des connaissances de façon expérimentale et de les rendre disponibles pour utilisation. Il ressemble au cerveau en deux points:

- la connaissance est acquise au travers d'un processus d'apprentissage.
- les poids des connections entre les neurones sont utilisés pour mémoriser la connaissance [56].

II.2.2. Architecture d'un réseau de neurone

Les réseaux de neurones est une technique inspirée des réseaux de neurones biologiques pour exécuter des tâches calculatoires. Elle a la particularité de s'adapter, d'apprendre, de généraliser pour classer les données en entrée. Un réseau de neurones se compose essentiellement de [57] :

- Un ensemble de cellules de calcul, appelées neurones. Chaque neurone possède un état d'activation mesuré à sa sortie Y_k .
- Un ensemble de connexions entre les neurones. Chaque liaison est caractérisée par un poids $W_{j,k}$ reliant l'unité j avec l'unité k .
- La loi de propagation du réseau : l'effet de l'entrée active S_k d'un neurone sur les différents neurones.
- La méthode de collecte de l'information : Règles d'apprentissage.
- La fonction F_k réalisée par la cellule.
- Une entrée externe qui peut être considérée comme un décalage θ_k

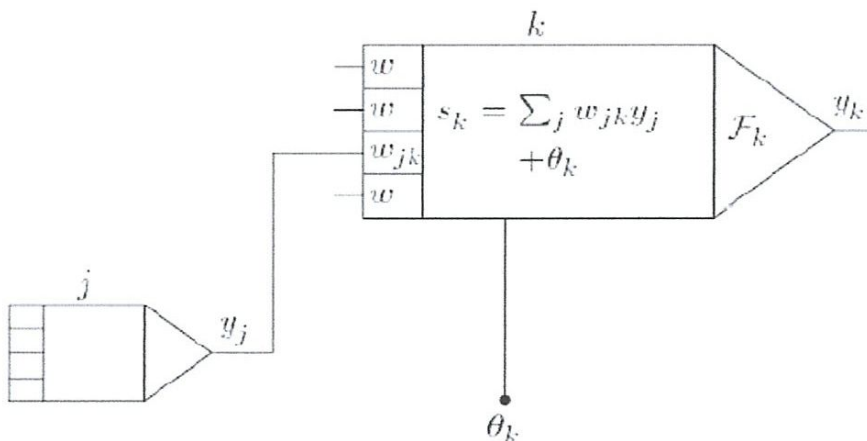


Figure 3.1 Composition d'un réseau de neurones

II.2.3. L'apprentissage dans un réseau de neurones

II.2.3.1. Définition

Dans un réseau de neurones, l'information est codée par les poids liés aux connexions. L'apprentissage est réalisé par des algorithmes de calcul dont le but est d'adapter ces poids en fonction des données d'apprentissage présentés à l'entrée du réseau [56].

II.3.3.2. Types d'apprentissage

▪ Apprentissage supervisé

On enseigne au réseau ce qu'il doit présenter comme information à la sortie selon les données présentées à l'entrée. Cela force les poids à prendre des valeurs distinctes selon les formes présentées au réseau de neurones [42].

Avantage

Le réseau atteint une précision très élevée au niveau de la reconnaissance et la base d'exemple n'a pas besoin d'être très grande.

Inconvénient

Les formes sont reconnues d'une manière trop exacte, Si une forme proche est présentée au réseau, celle-ci ne sera peut-être pas reconnue. On peut ajouter un bruit blanc aux données pour permettre une meilleure généralisation.

▪ Apprentissage non supervisé

Le réseau regroupe les données selon des algorithmes de calcul de distance par la méthode des moindres carrés. À chaque groupe correspond à une classe. Après l'identification de toutes les classes, une valeur de sortie, dictée par l'utilisateur, est associée à chacune des classes. Et peut-être associé à une inférence bayésienne pour produire des probabilités conditionnelles pour chaque variable aléatoire étant donné les autres [42].

Avantage

Le réseau atteint une précision très élevée au niveau de la reconnaissance et la généralisation est très bonne.

Inconvénient

La base d'exemples se doit d'être très volumineuse et le temps d'apprentissage est très long si un nouvel exemple est présenté au réseau pour un nouvel apprentissage.

II.3.4. Le perceptron multicouche

Le perceptron multicouche est un réseau orienté de neurones artificiels organisé en couches où l'information voyage dans un seul sens, de la couche d'entrée vers la couche de sortie.

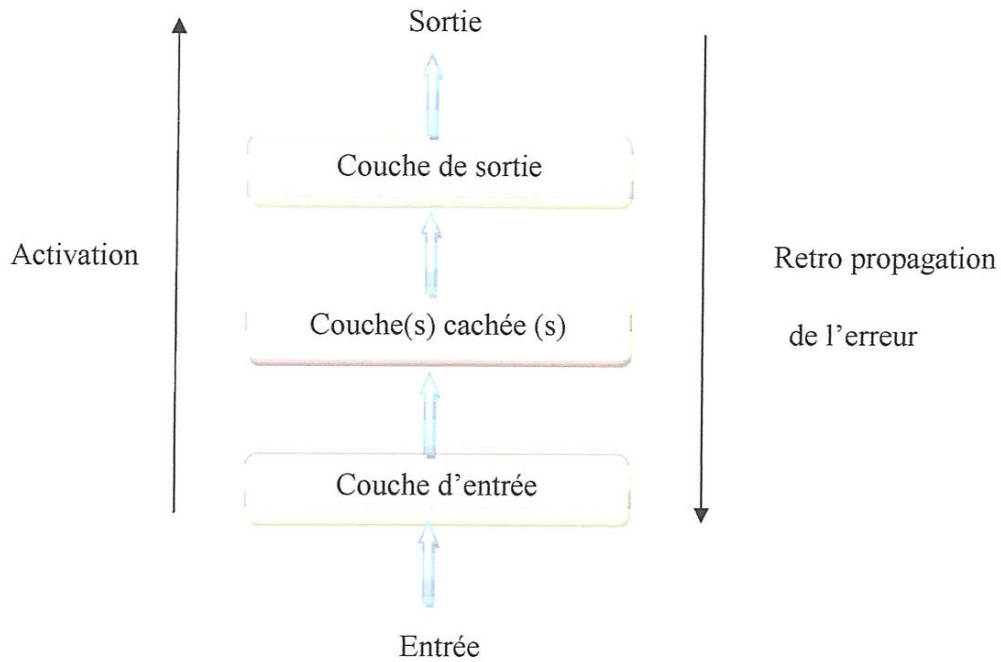


Figure 3.2: *Perceptron multicouche*

La couche d'entrée représente une couche virtuelle associée aux entrées du système, les neurones de cette couche sont reliés au monde extérieur et reçoivent tous le même vecteur d'entrée. Les neurones de la couche cachée n'ont aucun lien avec l'extérieur. Enfin, les sorties des neurones de la dernière couche correspondent aux sorties du système.

Les neurones sont reliés entre eux par des connexions pondérées appelées poids de ces connexions qui gouvernent le fonctionnement du réseau et "programment" une application de l'espace des entrées vers l'espace des sorties à l'aide d'une transformation non linéaire.

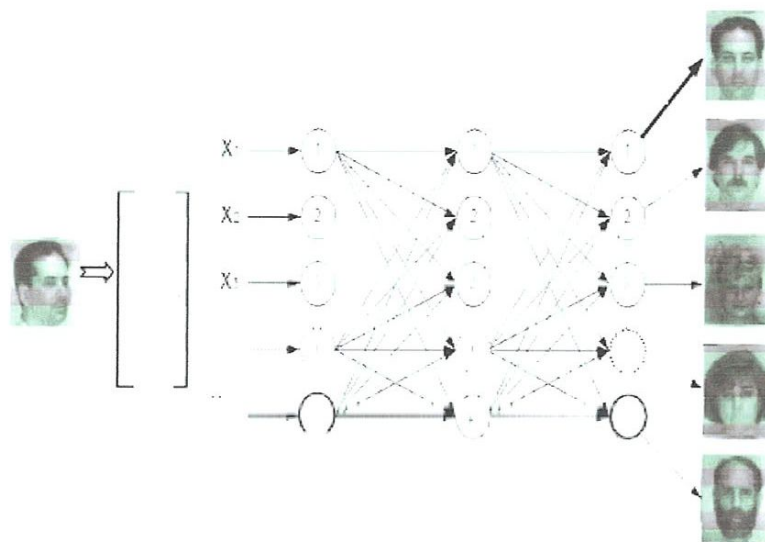


Figure 3.3 : *Modèle d'un perceptron multicouche pour la reconnaissance de visage*

L'activité dans le réseau se fait en propageant l'activation de la couche d'entrée jusqu'à la couche de sortie tandis que l'erreur est calculée dans le sens inverse.

II.2.5. Fonction d'activation

La fonction d'activation permet de définir l'état interne du neurone en fonction de son entrée totale, voici quelques fonctions souvent utilisées [43] :

- **Fonction binaire à seuil**

Le seuil introduit une nonlinéarité dans le comportement du neurone, cependant il limite la gamme des réponses possibles à deux valeurs.

- **Fonction Heaviside** définie par :

$$h(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

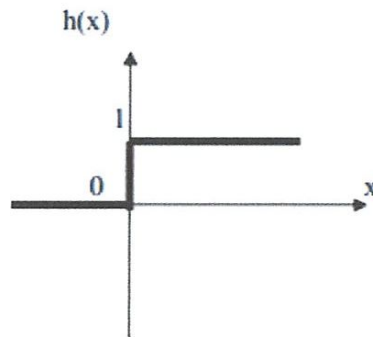


Figure 3.4: Fonction Heaviside

- **Fonction signe** définie par :

$$h(x) = \begin{cases} +1 & \text{si } x \geq 0 \\ -1 & \text{sinon} \end{cases}$$

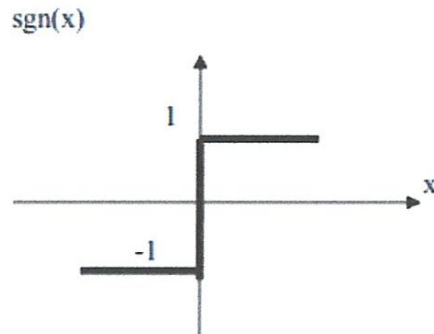


Figure 3.5 : Fonction signe

- **Fonction linéaire**

C'est l'une des fonctions d'activations les plus simples, sa fonction est définie par : $c(x) = x$

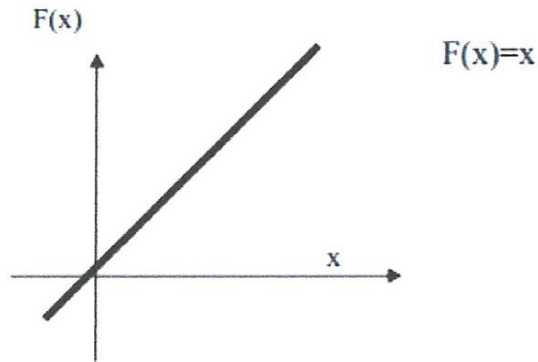


Figure 3.6 : Fonction linéaire

- **Fonction linéaire à seuil ou à multi seuil**

Défini comme suit :

$$f(x) = \begin{cases} x & \text{si } x \in [u, v] \\ u & \text{si } x \geq v \\ v & \text{si } x \leq u \end{cases}$$

Cette fonction représente un compromis entre la fonction linéaire et la fonction seuil : entre deux barres de saturation, elle confère au neurone une gamme de réponses possibles. En modulant la pente de la linéarité, on affecte la plage de réponse du neurone.

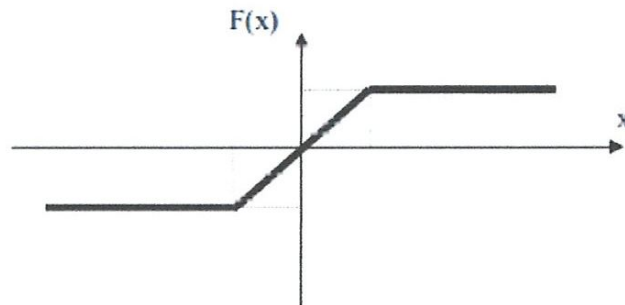


Figure 3.7 : Fonction linéaire à seuil

- **Fonction sigmoïde**

Elle est l'équivalent continu de la fonction linéaire. Étant continue, elle est dérivable, d'autant plus que sa dérivée est simple à calculer, défini par :

$$f(x) = \frac{1}{1 + e^{-x}}$$

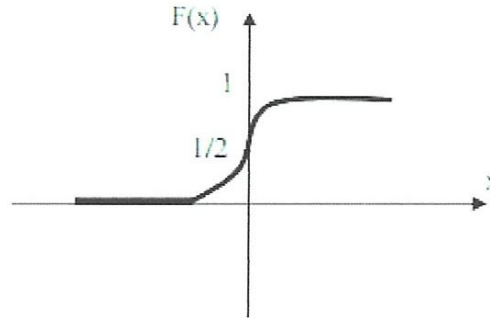


Figure 3.8 : Fonction sigmoïde

III. Conception du système

Les différentes étapes de conception de notre système sont présentées par 3 modules :
Module de prétraitement, module d'extraction des caractéristiques du visage et module de reconnaissance du visage.

1-Module de prétraitement

- Normalisation
- Egalisation de l'histogramme
- Filtre médian

2-Module d'extraction des caractéristiques du visage par la DCT

- Découpage de l'image en blocs
- Application de DCT sur chaque bloc
- Récupération des coefficients DCT (scan zigzag)
- Extraction et concaténation des premiers coefficients de la DCT pour chaque bloc afin de former un vecteur unifié

3-Module de reconnaissance de visage par le réseau de neurones

3.1-Apprentissage

3.2-Test et décision

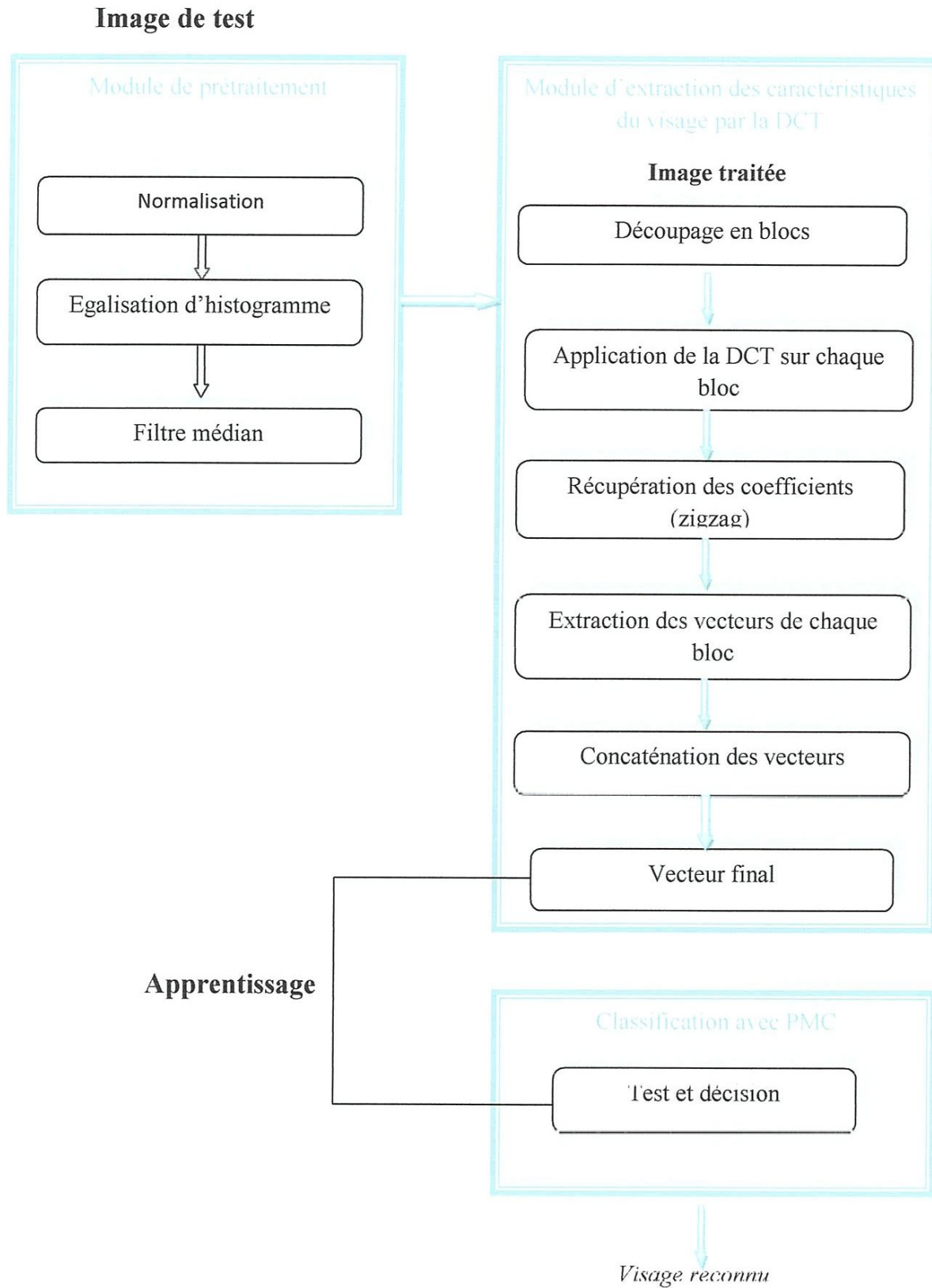


Figure 3.9 : Conception du système de reconnaissance de visage avec la méthode hybride DCT et réseau de neurones

III.1. prétraitement

Le but de prétraitement est d'améliorer la qualité de l'image en diminuant le bruit. Il permet de préparer l'image qui va être soumise à des traitements futurs.

III.1.1. Normalisation

La dynamique des niveaux de gris dans une image est l'intervalle $[a, b]$, correspondant aux niveaux de gris les plus petits et les plus grands. La normalisation consiste à obtenir la dynamique maximale $[0, 255]$. Pour cela on utilise la transformation suivante :

$$T : [a, b] \rightarrow [0, 255]$$

$$x \rightarrow \frac{(b-a)}{x-a} * 255$$



Figure 3.10 : Normalisation d'une image

III.1.2. Egalisation de l'histogramme

Un histogramme représente les niveaux de gris et fournit une description globale de l'image. L'objectif de l'égaliseur étant de rendre plus visible les détails peu perceptibles dans une zone contrastée de l'image ce qui revient à donner la même chance d'apparition aux niveaux de gris.

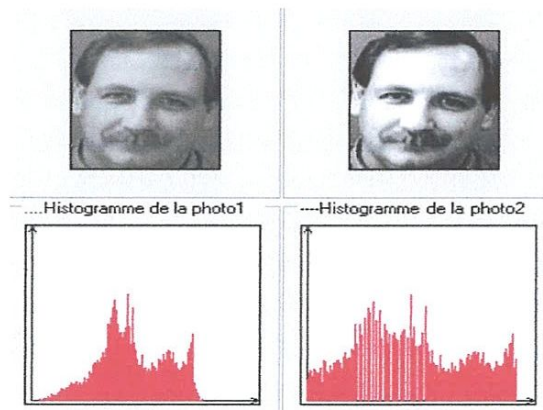


Figure 3.11 : Egalisation de l'histogramme

III.1.3. Filtre médian

Le filtre médian nous permet d'éliminer certains types de bruit (tache isolée), son principe est de remplacer le niveau de gris de chaque pixel par la médiane des niveaux de gris du voisinage du pixel en question, cette méthode est particulièrement pertinente quand la configuration du bruit est forte.

III.2. Extraction des caractéristiques du visage par la DCT

L'avantage majeur de la DCT est qu'elle permet de réduire l'information redondante et peut être utilisé comme une étape d'extraction de caractéristiques.

Les différentes étapes d'extraction des caractéristiques de visage sont illustrées par la figure suivante :

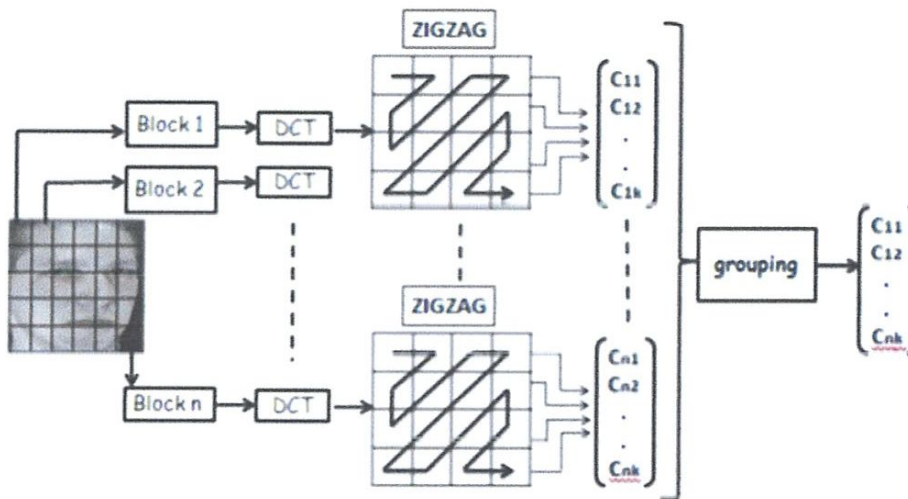


Figure 3.12 : Extraction des caractéristiques du visage par DCT

III.2.1. Découpage de l'image en blocs

Plusieurs facteurs sont à prendre en ligne de compte pour tester la DCT. Comme l'application de la transformée en cosinus discrète ce fait par bloc, on fera donc varié la taille des blocs DCT comme suit :

Application de la DCT Localement :

- ❖ Bloc de taille 8 x 8
- ❖ Bloc de taille 16 x 16
- ❖ Bloc de taille 32 x 32

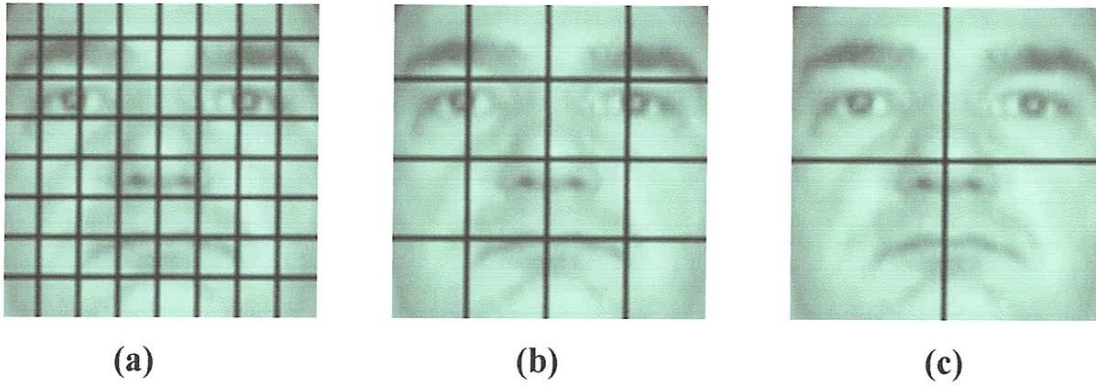


Figure 3.13 : Découpage de l'image en blocs

Application de la DCT globalement :

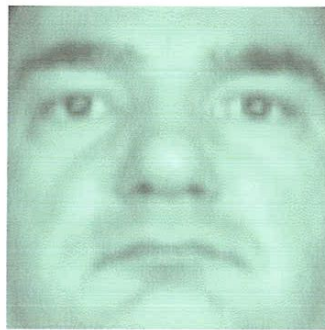


Figure 3.14 : Image sans découpage en blocs

III.2.2. Application de la DCT sur chaque bloc

C'est le fait d'appliquer une transformation en cosinus discrète aux blocs de pixels, pour retirer la redondance des données de l'image. Cette étape a pour but de regrouper l'énergie en passant de la notion de pixel à la notion de fréquences et amplitudes.

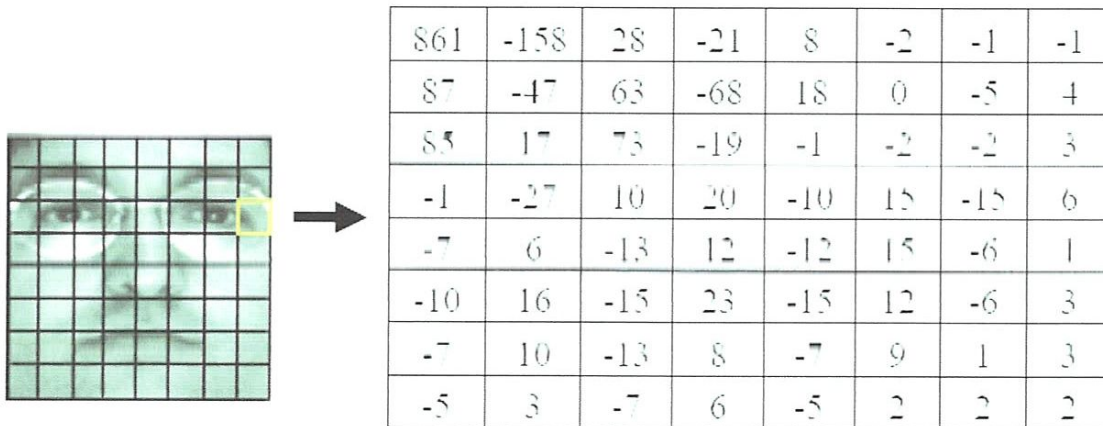


Figure 3.15 : Extraction des valeurs d'un bloc 8x8

182	182	182	181	181	179	180	178
178	178	179	181	181	181	182	179
182	181	181	181	181	182	182	179
182	182	182	181	181	179	180	179
178	178	178	182	181	182	182	179
182	181	181	181	180	182	182	178
182	181	181	181	181	182	182	178
182	182	182	181	181	179	180	179

↓

1445	2	2	2	1	3	1	0
1	1	0	0	0	0	0	0
0	2	0	0	0	0	0	1
0	0	0	0	0	0	0	0
0	2	0	2	1	1	0	0
2	4	1	0	0	0	0	1
1	3	1	0	0	0	0	0
0	2	0	1	0	0	0	0

Figure 3.16 : passage du domaine spatial au domaine fréquentiel par la DCT

III.2.3. Récupération des coefficients DCT (scan zigzag)

Une fois la DCT calculée sur un bloc, nous obtenons une matrice carrée des valeurs pour chacune des fréquences. Le codage de la matrice DCT va se faire en parcourant les éléments dans l'ordre imposé par une séquence particulière appelée séquence zigzag. Cette séquence a la propriété de parcourir les éléments en commençant par les basses fréquences et de traiter les fréquences de plus en plus hautes. Puisque la matrice DCT quantifiée contient beaucoup de composantes de hautes fréquences nulles, l'ordre de la séquence zigzag va engendrer de longues suites de 0 consécutifs.

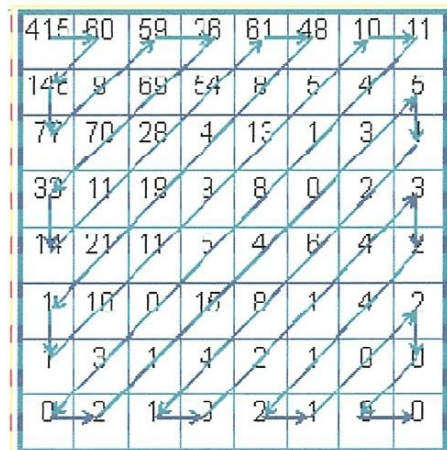


Figure 3.17 : Parcours en ZIGZAG

III.2.4. Sélection des premiers coefficients de la DCT et extraction de vecteur de chaque bloc

Afin de tester le nombre de coefficient DCT à prendre, On se fixe une taille de bloc de 8x8 et on fait varier le nombre de coefficient DCT. Voici les valeurs testées : 5, 8, 16, 20.

III.2.5. Concaténation des vecteurs des blocs et formation du vecteur final

Chaque image est représentée par un vecteur composé des premiers coefficients de la transformée. A chaque introduction d'une image test au module, sa transformée est calculé et un certain nombre de coefficients est retenu pour la comparaison avec ceux de la base de donnée.

III.3. Classification

III.3.1. Architecture du modèle PMC

Notre modèle utilise un apprentissage supervisé qui se base sur le principe de réglage des poids synaptiques pour l'entrée correspondante à la sortie désirée, il consiste à modifier les paramètres du réseau de façon que les réponses de ce dernier soient attendues.

Un algorithme de retro propagation de gradient a été utilisé afin de minimiser l'erreur quadratique moyenne entre la valeur de la sortie obtenue et celle de la sortie désirée.

Nous effectuons l'apprentissage d'un réseau possédant les caractéristiques suivantes :

- **Couche d'entrée :** un vecteur de 1024 valeurs, où les valeurs représentent les caractéristiques du visage.
- **Couche cachée :** le choix de ce paramètre est effectué en utilisant une heuristique, car il n'est n'existe pas une formule ou méthode directe permettant de déterminer le nombre de neurones dans cette couche.
- **Couche de sortie :** contient 40 neurones pour la base ORL, est 15 neurones associés à 15 classes pour la base YALE.
- **Fonction d'activation choisie :** fonction sigmoïde unipolaire.
- **Nombre des itérations :** 2000

Le schéma (3.18) présente l'organisme de l'apprentissage selon le modèle PMC proposé :

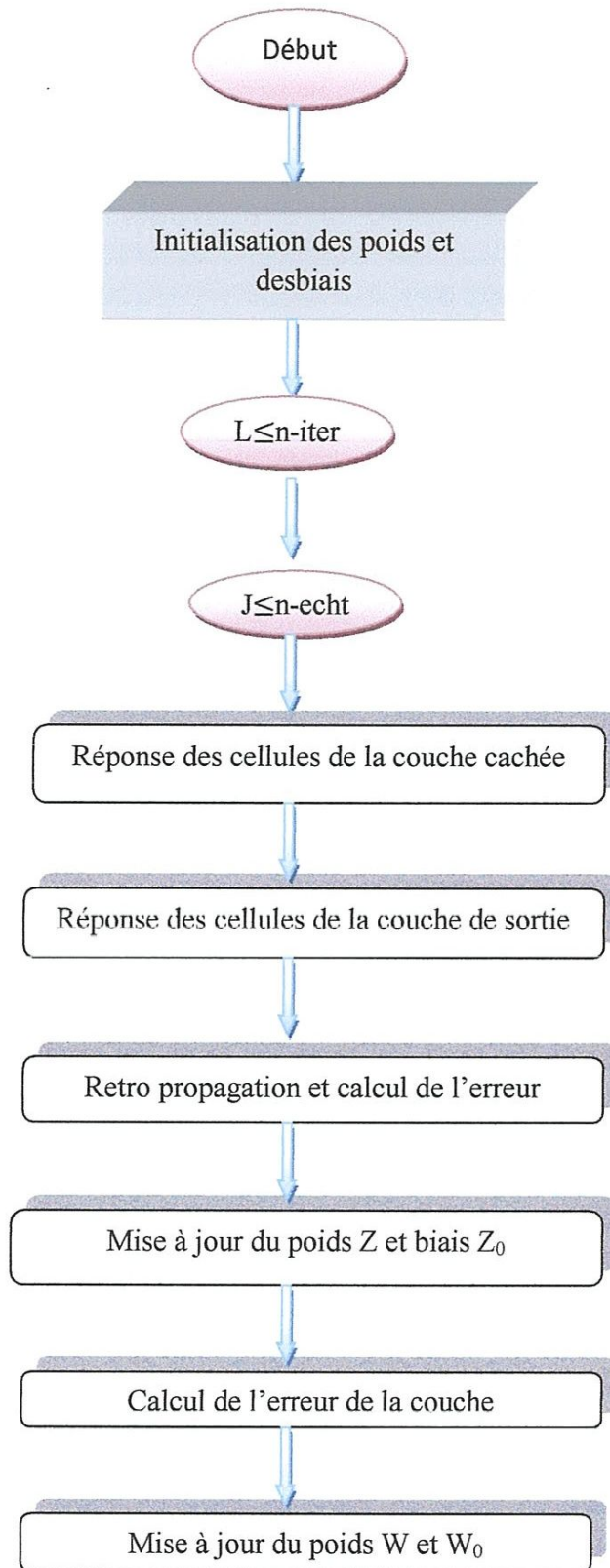


Figure 3.18 : Organisme de l'apprentissage selon le modèle PMC proposé

III.3.2. Phase d'apprentissage

l'apprentissage du perceptron multicouches est supervisé et consiste à adapter les poids des neurones de manière à ce que le réseau soit capable de réaliser une transformation donnée, présenté par un ensemble d'exemples constitué d'une suite de N vecteurs d'entrées $x_k = [x_{k1} \ x_{k2} \ \dots \ x_{kd}]^T$ associés à une autre suite de vecteurs de sorties désirées

$T^k = [x_1 \ x_2 \ \dots \ x_{hl}]^T$. voici la formalisation de cet algorithme [42][43][44] :

x : Matrice de stimulus de dimension $(1, N)$ ou N représente le nombre de stimulus à présenter au réseau.

x_k : Stimulus numéro k (colonne k de la matrice x)

T : Matrice cible ou réponse théorique de dimensions $(1, W)$

t_k : Cible n° k de la matrice T

h_k : Sortie de la couche cachée de dimension $(L, 1)$

o_k : Réponse du réseau, vecteur colonne à j composante

w : Matrice des poids des connexions liant la couche d'entrée à la couche cachée, de dimension (L, i)

w_0 : Vecteur de biais des cellules de la couche cachée de dimension $(L, 1)$

z : Matrice de poids de connexions liant la couche cachée à la couche de sortie de dimension (j, l)

z_0 : Vecteur de biais des cellules de la couche de sortie de dimension $(j, 1)$

A chaque présentation d'un stimulus n° k on réalise les étapes successives suivantes :

- **passage du stimulus de l'entrée vers la sortie du réseau**

$$O_k = g[z(f(wx_k + w_0))] + z_0 \dots \dots \dots (3.6)$$

Tel que :

- f et g sont les fonctions de transfert (activation), respectivement des couches cachée et sortie.

Le vecteur de sortie est comparé à celui que l'on désire obtenir :

$$e_k = t_k - o_k \dots \dots \dots (3.7)$$

• **la rétro propagation de l'erreur dans le réseau**

L'apprentissage du réseau consiste à modifier , à chaque pas d'apprentissage ,les poids et les biais afin de minimiser la somme des carrées des erreurs en sortie.la méthodes de rétro propagation est basée sur la technique du gradient :

$$E = e_k^t \quad e_k = \frac{1}{2} (t_k^t + o_k^t o_k - 2o_k^t t_k) \dots \dots \dots (3.8)$$

$$\nabla E_k / Z = \frac{1}{2} [\nabla o_k^t o_k - 2o_k^t t_k] / Z \dots \dots \dots (3.9)$$

Si l'on considère les fonctions de transferts identiques pour les deux couches ($f \equiv g$), la sortie du réseau est donnée par :

$$o_k = f[Z h_k + Z_0] \dots \dots \dots (3.10)$$

Le gradient de E_k par rapport à la matrice Z peut se calculer comme suit :

$$\nabla E_k / Z = \frac{\partial E_k}{Z} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial o_k} \frac{\partial o_k}{\partial (Z h_k + Z_0)} \frac{\partial (Z h_k + Z_0)}{\partial Z} \dots \dots \dots (3.11)$$

D'après l'expression de E_k la première dérivée partielle est :

$$\frac{\partial E_k}{\partial o_k} = o_k - t_k \dots \dots \dots (3.12)$$

La deuxième dérivée partielle dépend du type de la fonction de transfert utilisée .Dans le cas de sigmoïde unipolaire, la deuxième dérivée partielle a pour expression :

$$\frac{\partial o_k}{\partial (Z h_k + Z_0)} = f(Z h_k + Z_0) \cdot [1 - f(Z h_k + Z_0)] = o_k \cdot (1 - o_k) \dots \dots (3.13)$$

Où 1 représente un vecteur unitaire de même taille que le vecteur de sortie o_k et l'opérateur « .* » définit le produit terme ou produit de hadamard.

La troisième dérivée partielle est simple à calculer :

$$\frac{\partial (Z h_k + Z_0)}{\partial Z} = h_k^T \dots \dots (3.14)$$

La mise à jour se faisant dans le sens inverse du gradient, la matrice de poids Z et l'étape future ($k + 1$) est :

$$Z(k + 1) = Z(k) - \eta \nabla E_k / Z = Z(k) - \eta [(t_k - o_k) \cdot o_k \cdot (1 - o_k) h_k^T] = Z(k) + \eta \partial_s h_k^T$$

Avec :

$$\partial_s = (t_k - o_k) \cdot o_k \cdot (1 - o_k)$$

η :Le gain d'apprentissage.

La même méthode sera utilisée pour la mise à jour du vecteur des biais Z_0 . En considérant que le vecteur d'erreur passant par les biais est unitaire, on obtient facilement :

$Z_0(k + 1) = Z_0(k) - \eta \nabla E_k / Z_0 = Z_0(k) + \eta (t_k - o_k) \cdot o_k \cdot (1 - o_k) = Z_0(k) + \eta \partial_s$
 L'erreur de sortie ∂_s est retro propagée à la sortie de la couche cachée à travers la transposée de la matrice Z .

Par analogie avec l'expression précédente la mise à jour de la matrice, où h_k est le vecteur de sortie de la couche cachée, on obtient :

$$w(k + 1) = w(k) - \frac{\eta \nabla E_k}{w} = w(k) + \eta Z(Z)^T \partial_s \cdot h_k \cdot (1 - h_k) x_k^t = w(k) + \eta \partial_h x^T$$

$$w_0(k + 1) = w_0(k) - \frac{\eta \nabla E_k}{w_0} = w_0(k) + \eta Z(k)^T \partial_s \cdot h_k \cdot (1 - h_k) = w_0(k) + \partial_h$$

Tel que:

$$\partial_h = \eta Z(k)^T \partial_s \cdot h_k \cdot (1 - h_k) \dots \dots \dots (3.15)$$

III.3.3. phase de reconnaissance

Après la phase d'apprentissage, on passe à la phase de reconnaissance.

On introduit un vecteur de test inconnu et cet algorithme doit être en mesure de prendre une décision d'affectation. Ceci serait visible au niveau de sorties. L'algorithme de cette phase est comme suit :

- Lire le vecteur candidat $x(x_1, x_2, \dots, x_n)$
- Lire les poids et les biais correspondants qui sont calculés lors de la phase d'apprentissage
- Calculer les sorties du réseau, celles de la couche cachée et la couche sortie de la même manière que lors de la phase d'apprentissage.
- Faire la décision à partir des valeurs obtenues alors l'affectation à la classe C_k sera fait si la sortie est proche de 1.

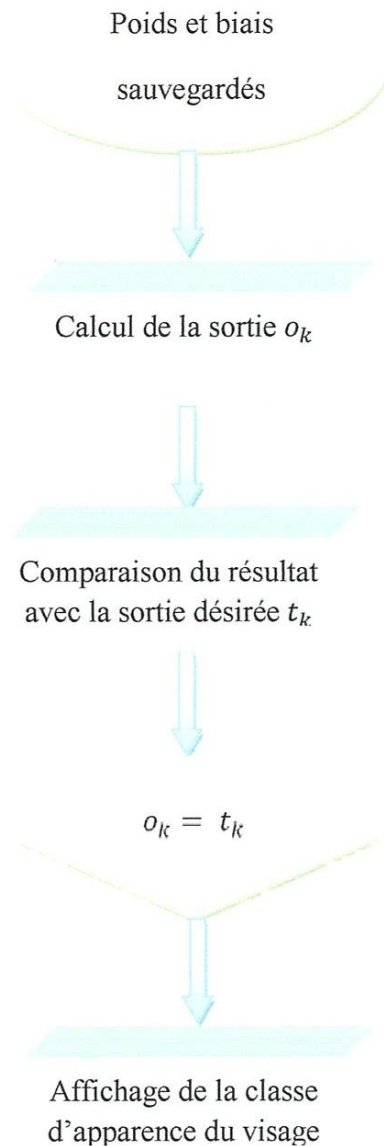


Figure 3.19: Organisme de la reconnaissance de visage

IV. Conclusion

Dans ce chapitre nous avons présenté d'une façon détaillée les trois méthodes utilisées dans notre travail ainsi que la partie conception de notre système allant de la phase de prétraitement, l'extraction des caractéristiques de visage utilisant la transformée de cosinus discrète, l'apprentissage et la reconnaissance de visage à l'aide d'un réseau de neurones multicouches.

Chapitre 04 :

Implémentation

I. Introduction

Pour valider notre travail et le situer par rapport à ce qui à été fait au par avant, des testes doivent être effectués. En effet, chaque méthode doit subir à une multitude de tests pour connaître ses avantages et ses faiblesses, le but est de trouver un compromis entre les différents paramètres qui nous permet de maximiser le taux de reconnaissance et de minimiser le temps d'exécution du système.

II. Environnement de développement

II.1. Environnement matériel

Le matériel utilisé est constitué d'un PC dont les caractéristiques sont présentées dans le tableau suivant:

Modèle	Pc portable
Processeur	Intel(R) core 2 duo (2 .10 GHZ)
Ram	2.96 Go
Disque dur	73.2Go

Tableau 4.1 : Caractéristiques du matériel utilisé

II.2. Environnement logiciel

Nous avons utilisé deux langages de programmation pour réalisé notre projet, à savoir : Java et Matlab.

II.2.1. Java

Java est un langage de programmation orienté objet et un environnement d'exécution, développé par Sun Microsystems. Il fut présenté officiellement en 1995. Java surgit sur la grande scène d'Internet et obtient immédiatement un énorme succès. La prétention de Java est de constituer la colle universelle capable de connecter les utilisateurs aux informations, que celles-ci proviennent de serveurs Web de bases de données, de fournisseurs d'informations ou de toute autre source imaginable, voici quelques caractéristiques de java :

- ❖ Java est devenu le meilleur langage de programmation, orienté objet, portable et gratuit.
- ❖ Java utilisé pour : jeux en ligne, logiciel de traitement d'image, encodage vidéo.
- ❖ Java est organisée, il contient des classes bien conçu et bien reparties.
- ❖ Java a été conçu pour être facile à utiliser et il est donc facile à écrire, compiler, déboguer par rapport aux autres langages de programmation.
- ❖ facile et flexible pour la programmation du traitement image : système de reconnaissance de forme, focale, visage.

II.2.2. Matlab

Notre logiciel a été développé avec le langage de programmation MATLAB version 7.9.

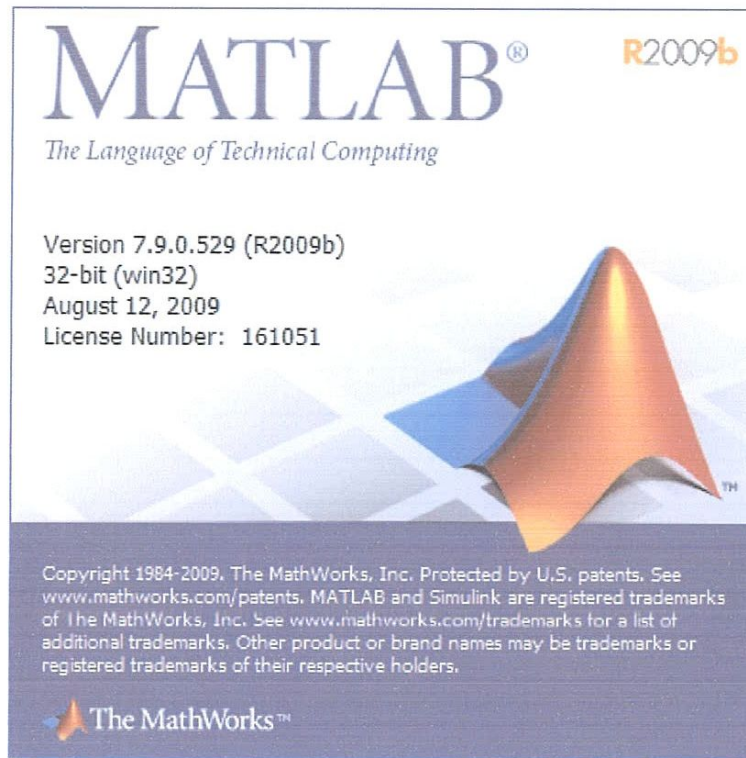


Figure 4.1 : Fenêtre du Matlab 7.9

Matlab est un logiciel de calcul numérique .Il est destiné à traiter des applications à partir des outils de l'analyse numérique matricielle. Matlab possède aussi tout un ensemble de fonctionnalités graphiques permettant de visualiser les résultats numériques .Il possède des boites à outils, c'est-à-dire des fonctionnalités supplémentaires, dédiées à des domaines particuliers du calcul scientifique, comme la résolution d'équations aux dérivées partielles, l'optimisation, l'analyse de données, etc.

En Matlab, les calculs sont effectués avec une arithmétique à précision finie .Ceci le différencie des logiciels de calcul symbolique tel que Maple, mais la comparaison n'a pas lieu d'être. Calcul numérique et calcul symbolique sont des outils complémentaires du calcul scientifique. Il existe deux modes de fonctionnement :

- **Mode interactif** : MATLAB exécute les instructions au fur et à mesure qu'elles sont données par l'utilisateur.
- **Mode exécutif** : MATLAB exécute ligne par ligne un "fichier .M" (programme en langage MATLAB).

II.3. Bases de données utilisées

II.3.1. Base ORL

La base ORL est conçue par AT&T laboratoires de l'université de Cambridge en Angleterre, la base de données ORL (Olivetti Research Laboratory) est une base de donnée de référence pour les systèmes de reconnaissances automatique des visages. En effet tous les systèmes de reconnaissances de visages trouvés dans la littérature ont été testés par rapport à l'ORL, cette popularité est due au nombre de contraintes imposées par cette base car la plus part des changements possibles et prévisibles du visage ont été pris en compte, comme par exemple : le changement de coiffure, la barbe, les lunettes, les changements dans les expressions faciales, etc. Ainsi que les conditions d'acquisition telles que : le changement d'illumination et le changement d'échelle du à la distance entre le dispositif d'acquisition et l'individu.

La base de données ORL est constituée de 40 individus, chaque individu possède 10 poses, donc la base contient 400 images. Les poses ont été prises sur des intervalles de temps différents pouvant aller jusqu'à trois mois. L'extraction des visages à partir des images a été faite manuellement. Nous présenterons dans ce qui suit les figures montrant les spécificités de la base de données de référence ORL.

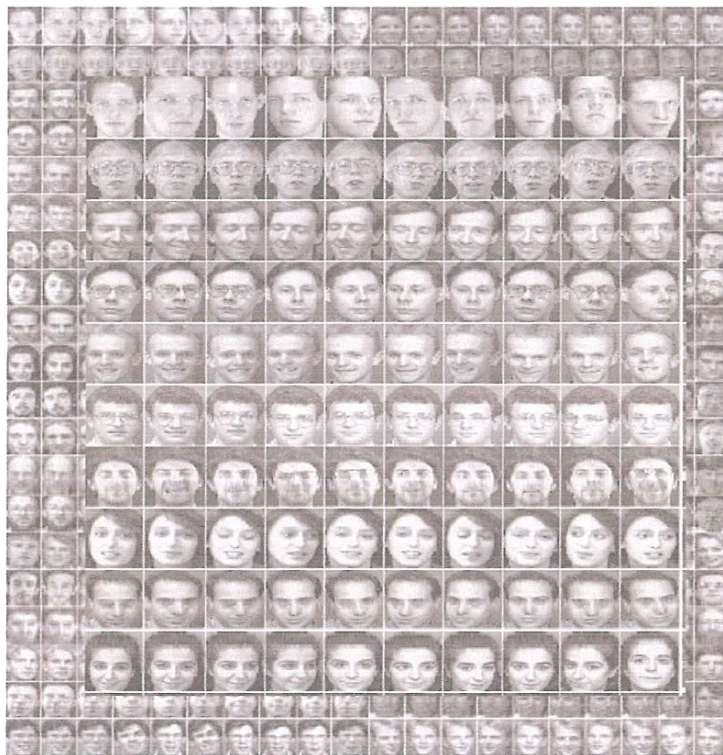


Figure 4.2 : Base de données ORL

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

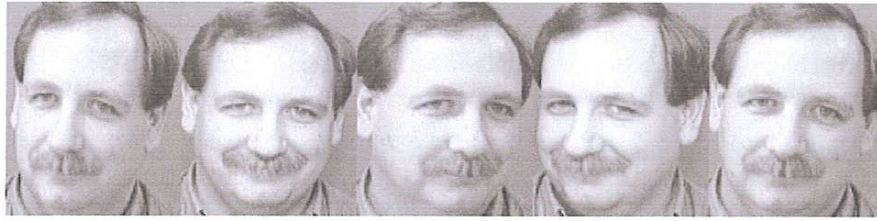


Figure 4.3 : Exemple de changements d'orientations du visage

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



Figure 4.4 : Exemple de changements d'éclairage

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



Figure 4.5 : Exemple de changements d'échelle

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

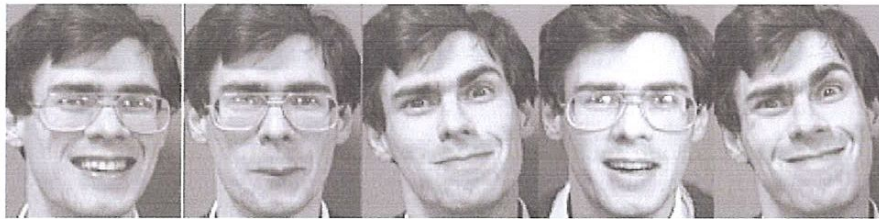


Figure 4.6 : Exemple de changements des expressions faciales

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



Figure 4.7 : Exemple d'individus de différents âges, races et sexes

Dans notre cas, nous avons divisée cette base en deux parties :

- **Base d'apprentissage** : composée de 200 images de 40 personnes, 5 images pour chaque individu.
- **Base de test** : la même division que la base d'apprentissage.

II.3.2. Base Yale faces

Cette base contient 165 images en niveaux de gris, représentant les visages de 15 individus. Il y a 11 images par individu, chacune correspondant à un type de prise de vue particulier. L'acquisition des images a été opérée au centre CVC (Center for Computational Vision and Control) de l'université Yale dans un environnement entièrement contrôlé. Les images contiennent des variations de luminosité et d'expressions faciales, elles ont été acquises toutes sur un fond blanc et avec une résolution de 320×243 .

Les 11 catégories d'images varient selon les critères suivants:

- **variation de l'expression du visage**: normale, triste, somnolent, surpris, clin d'œil, heureux.
- **variation des accessoires**: avec lunettes, sans lunettes.
- **variation de l'éclairage**: on a trois angles d'éclairage différent : centre, gauche, droite.

Nous avons divisée cette base en deux parties :

- **Base d'apprentissage** : composée de 90 images de 15 individus, 6 images pour chaque individu.
- **Base de test** : composée de 75 images de 15 individus, 5 images pour chaque individu.



Figure 4.8 : Images extraites de la base YALE FACES

III. Présentation de l'interface graphique



Figure 4.9: Interface principale du système réalisé

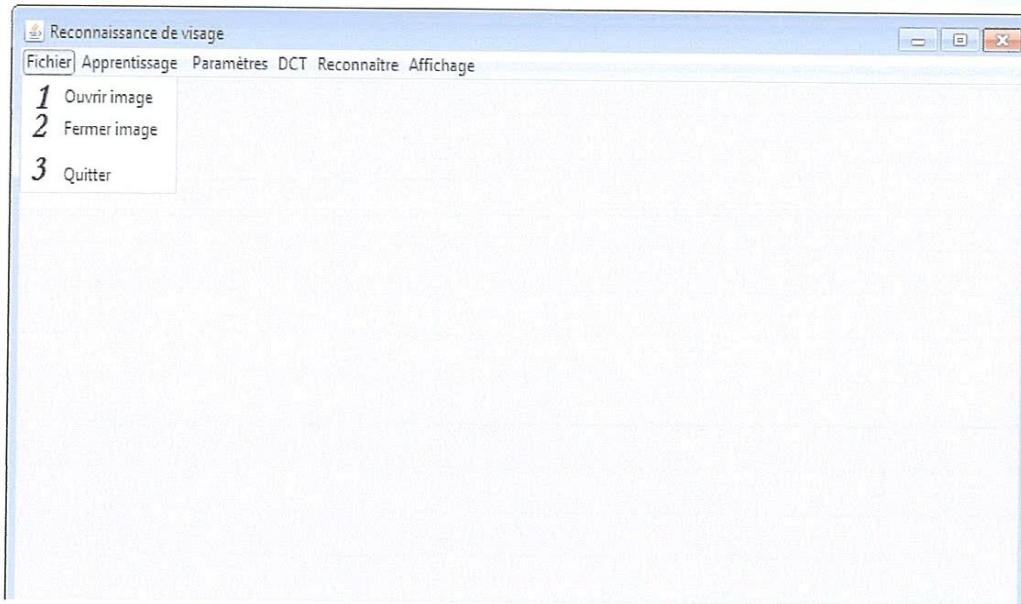


Figure 4.10 : Menu de chargement des images à partir de la base de données

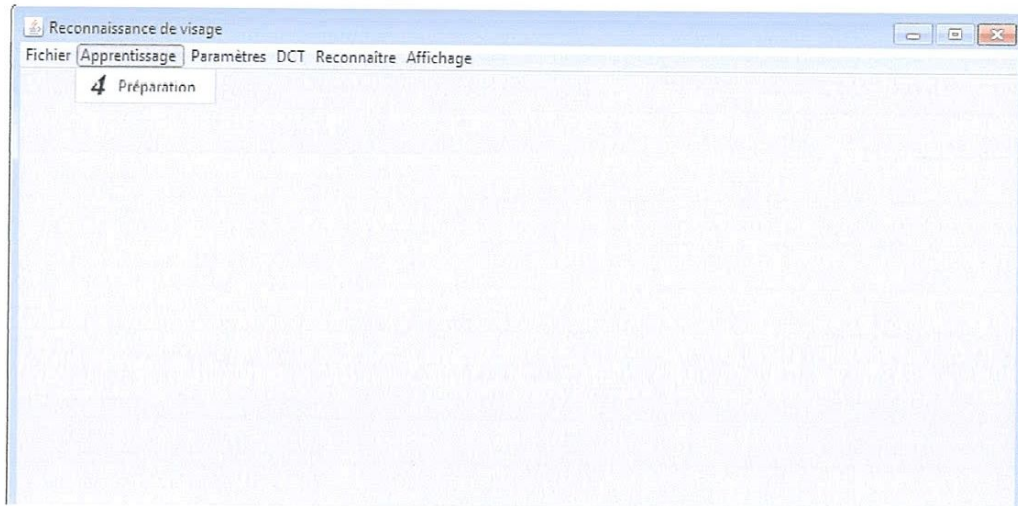


Figure 4.11 : Menu de lancement de l'apprentissage

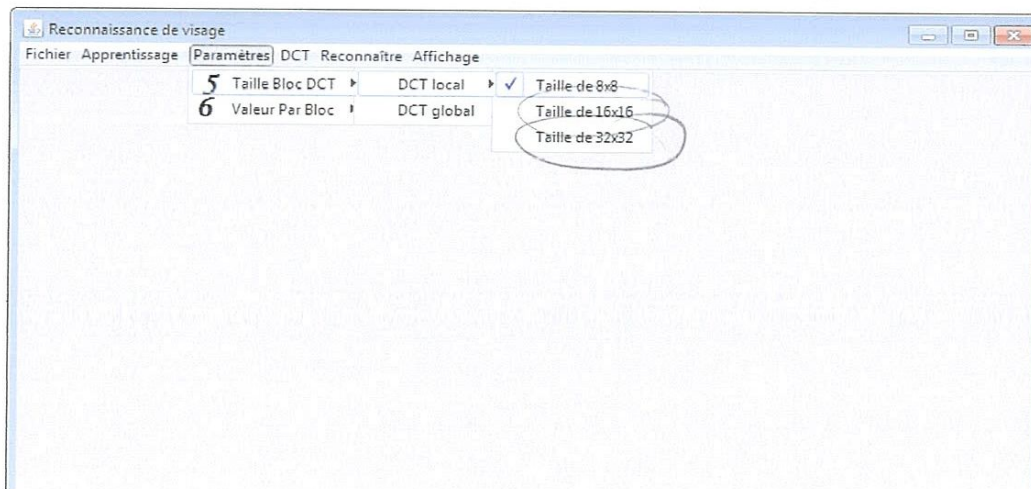


Figure 4.12 : Menu pour la manipulation des paramètres de DCT

Chapitre 4 : Implémentation

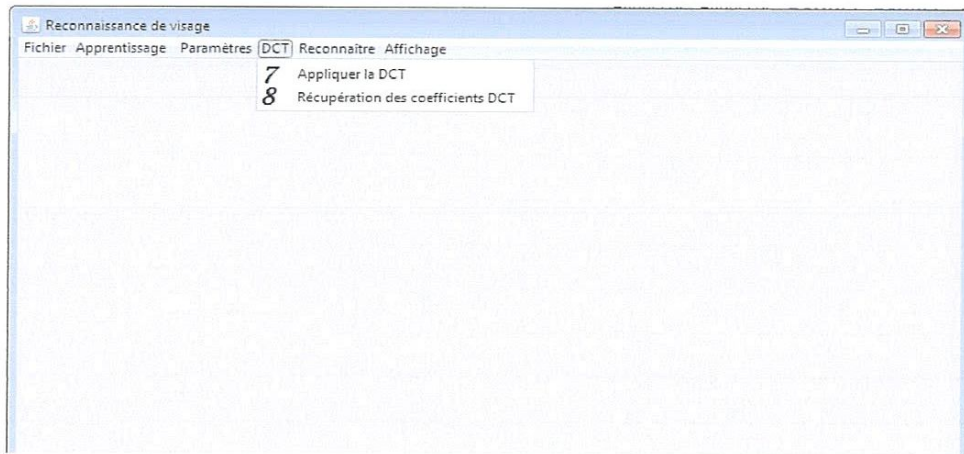


Figure 4.13 : Menu d'application et récupération des coefficients DCT

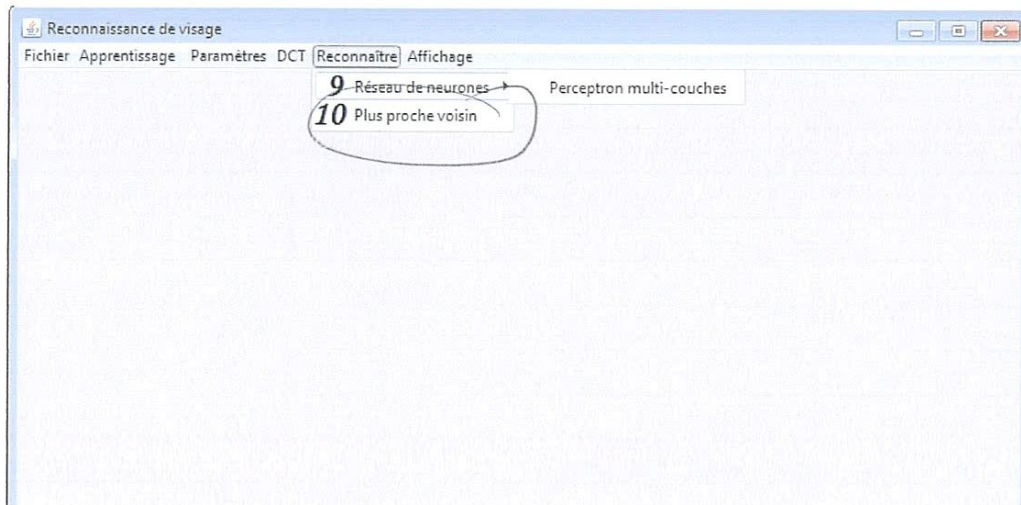


Figure 4.14 : Menu de reconnaissance des visages

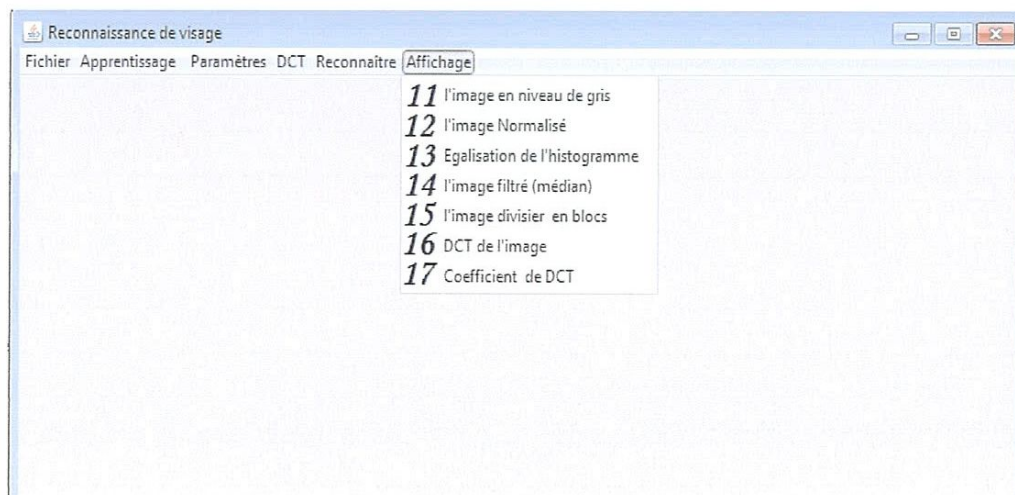


Figure 4.15 : Menu d'affichage des résultats

IV. Tests et résultats

IV.1. Test réalisé sur les paramètres de la DCT :

IV.1.1. Taille de blocs

On se fixe le nombre de coefficients en 16, et on test la taille des blocs :

Taille des blocs	Taux de reconnaissance
Image globale	46%
8 x 8	90%
16 x 16	69%
32 x 32	54%

Tableau 4.2 : Tests réalisés sur différentes tailles de blocs

D'après ces résultats on constate que :

- Le fait d'appliquer la DCT localement est mieux que l'appliquer globalement
- Le meilleur taux de reconnaissance est obtenue avec des blocs de tailles 8x8.

IV.1.2. Nombre de coefficients

On se fixe la taille des blocs en 8 x 8, et on test le nombre de coefficients DCT .

Nombre de coefficients	Taux de reconnaissance
5	48%
8	66%
16	89%
20	80%

Tableau 4.3 : Tests réalisés sur le nombre de coefficients DCT

Le taux de reconnaissance augmente en augmentant le nombre de coefficients DCT et ce entre 5 et 16 coefficients. On constate aussi que le taux de reconnaissance diminue légèrement entre 16 et 20 coefficients, et à partir de 20 coefficients il y a une nette baisse du taux de reconnaissance. On déduit que le nombre de coefficient DCT égal à 16 donnera les meilleurs résultats.

IV.2. Test réalisées sur les paramètres du PMC

La performance d'un perceptron multicouche dépend directement du choix de ses paramètres. Pour les déterminer, nous avons effectué plusieurs tests :

- **Nombres des itérations**

Nous avons fixé le nombre des neurones de la couche cachée en 40. Le tableau suivant illustre les résultats obtenu :

Nombre des itérations	1000	2000	3000
Taux de reconnaissance	90,5%	96%	100%

Tableau 4.5: Taux de reconnaissance obtenus en fonction de nombre des itérations

- **Nombre de neurones sur la couche cachée**

Nombre de neurones	40	75	100
Taux de reconnaissance	90,5%	86%	64%

Tableau 4.6 : Taux de reconnaissance en fonction de nombre de neurones de la couche cachée

Le meilleur taux que nous avons obtenu est égal à 90, 5%, il est donné pour un nombre de neurones égal à 40.

V. Interprétation

- le meilleur taux de reconnaissance est obtenue avec des blocs de tailles 8x8. On peut expliquer ce résultat par le fait que la DCT est appliquée localement sur des images décomposées en blocs, donc plus la taille des blocs est petite plus on accorde d'importance aux détails de l'image mais cela peut aussi avoir l'effet inverse, c'est-à-dire que si une image est bruitée, les résultats peuvent être faussés par le bruit.
- Concernant le nombre de coefficients DCT, prendre peu de coefficients voudrait dire qu'on a une perte d'information et en prendre trop inclurait du bruit qui pourrait influencer négativement sur le taux de reconnaissance, donc le but est de trouver le nombre de coefficients DCT qui résumerait l'information en écartant le bruit.
- La DCT fournit une robustesse contre les variations de l'apparence dans les régions locales causées par l'occlusion partielle du visage ou l'expression du visage, mais reste sensible contre les variations d'éclairage, ce qui explique les résultats modestes obtenus dans le cas des images de la base YALE.
- La réussite de l'apprentissage dépend le choix des différents paramètres : nombre de neurones de la couche cachée, valeurs initiales des poids, nombre des itérations, c'est-à-dire un ensemble d'essais.

V. Conclusion

Dans ce chapitre nous avons abordé l'aspect implémentation de notre logiciel, qui consiste à mettre au point les concepts développés précédemment, l'environnement de développement matériel et logiciel ainsi que les différents résultats obtenus et leur interprétation.

Conclusion générale et perspectives

La recherche dans le domaine de la reconnaissance des visages connaissait des progrès considérables au cours de ces dernières années. Parmi les techniques les plus évoqués, nous trouvons celles qui utilisent l'optimisation de la taille des données afin d'obtenir une représentation qui permet d'effectuer la reconnaissance. Pour ces méthodes, les images de visages sont vues comme des points dans un espace de très grandes dimensions. L'idée de base consiste à encoder les données initiales de passer à un autre espace de dimensions beaucoup plus réduites, tout en préservant autant d'informations utiles.

L'objectif de notre travail est destiné à réaliser un système de reconnaissance de visage basé sur l'hybridation de deux méthodes globale, à savoir : la transformée en cosinus discrète pour l'extraction du vecteur caractéristique de l'image de visage, et les réseaux de neurones pour la classification. Pour validé ce travail nous avons testé cette technique sur deux bases de données : ORL, YALE.

La DCT est utilisée afin d'extraire les caractéristique du visage, il permet de décarreler les images et réduire leur dimension en choisissant les plus grand coefficients, le but de la transformée en cosinus discrète est de passer dans un domaine fréquentiel, ce qui nous permettra de trier efficacement l'ensemble des données de l'image et ainsi supprimer certaines données où l'œil humain ne verra que très peu de différences, ce qui revient à supprimer les hautes fréquences de l'image tout en gardant les données majeures qui sont représentées par les basses fréquences.

L'algorithme proposé a prouvé sa performance sous différentes conditions d'acquisitions. On addition, c'est le premier algorithme qui est capable de gérer les variations qui apparaissent au visage sans avoir effectué des modifications spécifiques à chaque variation.

Il est possible d'adapter cet algorithme pour effectuer :

- une reconnaissance faciale dans le cas des images 3D
- une reconnaissance multimodale des personnes, par la combinaison du système proposé avec un système d'identification du locuteur.

Bibliographie

- [1] I. B. Cyganek, P. Siebert, « An introduction to 3D Computer Vision Techniques and Algorithms ». Wiley, 2008.
- [2] R. Ouafek, M.R Meghalsin, « la detection du visage dans une image », mémoire de fin d'études, école supérieure en informatique, Alger, 2009.
- [3] J.M. Munguakonkwa, « la liaison automatique de plusieurs images perçues sur un scanner ».Wiley, 2008.
- [4] C. Diltsm, « Numérisez vos collections », RCIP, octobre 2001.
- [5]C. Taouche, « implémentation d'un environnement parallèle pour la compression d'image à l'aide des fractales », mémoire pour l'obtention du diplôme en magister en informatique, université Mantouri Constantine, 2005.
- [6] A. Hardancourt, « Fou du multimédia », Sybex, 1995.
- [7] R.C. Gonzales, P. Wintz, « Digital Image Processing », Wiley, 1997.
- [8] T. Samamba, « Reconnaissance des formes comme outil d'aide aux traitements d'image », mémoire ISP/Bukavu, 2006.
- [9] S.Thon, « codage et compression des images », 2013.
- [10] P. Nerzic, « codage et compression d'images et de vidéo », Rennes1, 2012.
- [11] B.Patrick, « compression d'images fixes et des séquences vidéo », laboratoire des images et des signaux de Grenoble, 2008.
- [12] R. Sylvain, « La Compression des Données », Conférence, Club Photoshop de Nantes, 1999.
- [13] C. Benoit, A. Dusson, « La compression des données informatiques », Dusson, 1999.
- [14] R. Souadnia, K. Benmahamed, « Compression des images numériques fixes par les fractales », mémoire de fin d'études, Université Mantouri Constantine, Algérie, 2000.
- [15] A. Bacha, N. Hadj Said, « Compression des images Fixes Par Fractale: Partitionnement Quadtree », Université Mantouri Constantine, Algérie, 2000.
- [16] D. John, J. Woodward, H. Cristopher, J. Gatune, A. Thomas, « biometrics a look at facial recognition», commission de droits et sécurité Virginia, 2003.
- [17] International Biometric Group, <http://www.biometricgroup.com/>, Avril 2013
- [18] : N. Morizet, « reconnaissance biométrique par fusion multimodale du visage et de l'iris», Thèse de doctorat, Ecole Doctorale d'Informatique Télécommunications et Electronique de Paris, 2009.

- [19] commission technique de sécurité physique, « techniques de contrôle d'accès par biométrie », CLUSIF, juin 2003.
- [20] F. Perronnin, J. L. Dugelay, « Introduction à la biométrie, Authentification des individus par traitement audio-vidéo », *Revue de traitement de signal*, Vol. 19, No. 4, 2002.
- [21] R. Beveridge, M. Kirby, « Biometrics and Face Recognition », *IS&T Colloquium*, p. 25, 2005.
- [22] S. Boudjellal, « détection et identification de personne par une méthode biométrique », mémoire de magister, université Mouloud Mammeri de Tizi Ouzou, 2008.
- [23] W. Hizem, « capteur intelligent pour la reconnaissance de visage », thèse de doctorat, université de Pierre et Marie curie paris, 2009
- [24] S.ABABSA, « authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visage 2D/3D », thèse de doctorat, université d'Evry d' Essonne, 2008.
- [25] Z. Hammal, N. Evène, A. Cappiler, P. Coulon, « extraction des traits faciaux à l'aide d'un modèle paramétrique adapté », 2003.
- [26] N. BESBES, « indexation en intervenant d'un document vidéo par identification de visage », mémoire en vue de l'obtention du master de recherche, université de Paul Sabatier Toulouse III, 2007.
- [27] R. JOURANI, « reconnaissance de visage », mémoire pour l'obtention du diplôme des études supérieures approfondies, université de Mohammad V, Rabat, 2006
- [28] M.A. Hazem, « reconnaissance de visage », Mémoire de fin d'études Universités d'Avignon et du pays du Vaucluse, 2007.
- [29] H.K Ekernel, « a robust face recognition algorithm for real world application », these de doctorat, université Fridericiana zu Karlsruhe, 2009.
- [30] Y. Adini, Y. Moses, S. Ullman, « Face recognition: The problem of compensating for changes in illumination direction », *IEEE Xplore P* 721–732, 1997.
- [31] D. Blackburn, M. Bone, P. J Phillips, « Face recognition vendor test », *Vendors*, 2001.
- [32] R. Gross, J. Shi, J. Cohn. « Face Recognition: Third Workshop on Empirical Evaluation Methods in Computer Vision », 2001.
- [33] D.Khebab, « Système de détection du visage humain base sur la méthode SVM », mémoire de fin d'études, l'école supérieure d'informatique (INI), Algérie, 2007.
- [34] M.H. Yang, D. J. Kriegman, N. Ahuja, « Detecting faces in images », *IEEE Transactions on Pattern Analysis and Machine Intelligence* volume 24(1), pages 34–58, 2002.

- [35] A. Chaari, « Reconnaissance de visages par réseaux d'ondelettes de Gabor », thèse de doctorat, université de Lille 1, 2009.
- [36] I. mark, « face recognition algorithms », thèse de doctorat, 2010.
- [37] R. Rioux, T. Simon, « Redondance de classes pour la reconnaissance de visages par SVM », 14ème Col loque National de la Recherche en IUT, Campus de la Doua , Lyon Villeurbanne, France, mai 2008.
- [38] K. Etemad, R. Chellappa, « Discriminant analysis for recognition of human face images », Journal of the Optical Society of America A, 1997.
- [39] N. Morizet, F. Amiel, I. D. Hamed, T. Ea, «A comparative implementation of PCA face recognition algorithm». Conference international d'électronique, volume 2, pages 865 – 868.décembre 2007.
- [40] J. F. Cardoso, « Analyse en composantes indépendantes». Journées de Statistique, JSBL 2002, Bruxelles Belgique, mai 2002.
- [41] C. Havran, L. Hupet, J. Czyz, J. Lee, L Vandendorpe, M. Verleysen, « Independent Component Analysis for face authentication », Proc. Knowledge-Based Intelligent Information & Engineering Systems, 2002 .
- [42] Y. B. Jemaa, S. Khanfir. «Automatic Gabor features extraction for face recognition using neural networks». Computing research repository, 2008.
- [43] M. S. Charifa, A. Suliman, M. Bikdash , « Face recognition using a hybrid general backpropagation neural network ».Conference international Granular Computing, USA, novembre 2007.
- [44] W. Wang, « Face recognition based on radial basis function neural networks». Séminaire international, future d'information et technologie, pages 41 - 44, Leicestershire, UK, novembre 2008
- [45] S.C. Chen, J. Liu, Z. H. Zhou, M. Fld. « face recognition with one sample per person». Pattern Recognition, 2004.
- [46] X. Tan, S.C. Chen, Z.-H. Zhou, F. Zhang. « Recognizing partially occluded, expression variant faces from single training image», Pattern Recognition 2005.

- [47] V. Perlibakas. « Face recognition using Principal Component Analysis and Log Gabor Filters», Computing Research Repository, 2005.
- [48] X. Tana, C. Songcan. «Face recognition from a single image per person: A survey Pattern Recognition » ,2006.
- [49] N. Ahmed, T.Natarajam, K.Rao, « Discret cosine transform ».IEEE Transforms on computers, Javier, 1974.
- [50] A.F. Abate, M. Nappi, D.Riccio, G.Sabatino, «2D and 3D face recognition ».Elsevier, 2007
- [51] Z.Hafed,M.D.Levine, « Face recognition transform». Internationnal journal of computer vision, 2001
- [52] A.Samara, S.E.T.G. Alla,R. Ibrahim, «Face recognition using Wavelet transform, fast fourier transform and discret cosine transform». IEEE Internationnal symposium on micro mechatronic and human science, 2005.
- [53] H.Yin, P.Fu, J.Qiao, « face recognition based on DCT and 2DLDA». second International conference in innovation computing, information and control, 2007.
- [54] P.Abouzar, S.Yousef,S.Setarehdon, «Hybrid WT based-DCT face recognition».Dans IEEE Internationnal conference on signal processing and communication, Dubai,United Arab Emirates, 2007.
- [55] M. Bengherabi, L. Mezai, F. Harizi, A. Guessoum, M. Cheriet, «fusion de la DCT PCA et la DCT RLDA pour la reconnaissance de visage ».Thèse de doctorat, Université de Saad Dahlab, Blida, Alger, 2011.
- [56] E.Hervet, « réseaux de neurones ». Conférence, Canada, 2009.
- [57] C.Gagné, « apprentissage et reconnaissance ».université de Laval, Canada, 2011.