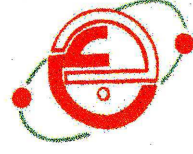


وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY  
UNIVERSITE BADJI MOKHTAR- ANNABA



جامعة باجي مختار- عنابة

Faculté des Sciences de l'Ingénieur  
Département D'Electronique

THESE

Présenté en vu de l'obtention du diplôme de DOCTORAT

# Reconnaissance de Visages Par Approche Géométrique

Option

Traitement du signal

Par

Mme BENCHERIET Chemesse Ennehar

DIRECTEUR DE THESE : Hicham TEBBIKH    Professeur    Université de Guelma

DEVANT LE JURY

PRESIDENT :	Noureddine	DOGHMANE	Professeur	Université de Annaba
EXAMINATEURS :	Mouldi	BEDDA	Professeur	Université de Annaba
	Djamel	CHIKOUCHE	Professeur	Université de Sétif
	Laamri	ZOUAOU	M.C	Université de Sétif

Soutenue le : 07-02-2007

# ABSTRACT

---

The main goal of this work consists in classifying and recognizing human faces applied on B-LAIG faces database created in LAIG (Laboratory of Automatic and Informatics of Guelma). Seen the complexity of manipulated objects (Faces), we used in classification two data types: geometric parameters defining facial features (distance between facial features) and pixel parameters defining by whole faces image.

On a comparative basis and in order to select the best classifier for our application, we used in the first time MLP (Multi Layer Perceptron) known by its performances in classification and the Principal Component Analysis (PCA) in the second time, certainly consider as an olden method but their results are the best compared to the MLP's.

At the third time and with the aim of exploiting the performances of both classification methods we propose an approach based on their combination that we called "Neuro-ACP" approach.

Another very important part of our work, consider as complement of the first one, that consists to detect human faces in images without any condition on lighting on background or on faces position and orientation. In this case, we propose method based on skin color detection and template matching. This combination permitted custom of gotten round inconveniences of both method and to minimize risks of false alert in decision face / non face. The distant purpose of this part is to create system that proceeds to classification and recognition of individuals whatever their position and lighting condition.

**Keywords** : Face recognition, faces feature, MLP, PCA, Neuro-ACP approach, face detection, skin color, color space, template matching.

# ملخص

قدمنا في هذا البحث نظام للمعرفة الآلية للوجوه في الصورة الرقمية معتمدين على استعمال الخصائص الهندسية والإجمالية للوجه. بالنسبة للخصائص الإجمالية للوجه اعتمدنا على البيكسل (pixel) - وحدة الصورة- حيث استعملنا صورة الوجه مباشرة أما بالنسبة للطريقة المبنية على الخصائص الهندسية فقد إتبعنا المرحلتين التاليتين:

• تحديد مكان العناصر المميزة للوجه.

• استخراج الابعاد والمتمثلة في المسافات بين العينين والفم و الانف.

بالنسبة لتطبيقاتنا جمعنا قاعدة للمعلومات الصورية تظم مئة شخص من الجنسين, تم اختيارهم بشكل عشوائي من الوسط الطلابي لجامعة قالمة, وذلك في مخبر الاتوماتيك و الاعلام الآلي بقالمة (LAIG).

هذه المذكرة تنقسم الى شطرين أساسيين :

الأول يتمثل في التعرف على الوجه باستعمال ثلاثة طرق تتمثل في الشبكات العصبونية الآلية المعروفة باسم (Multi Layer Perceptron : MLP) وطريقة التحليل باستعمال المركبات الاساسية (Principal Component Analysis : PCA) أما الطريقة الثالثة هي طريقة جديدة هجينة تجمع بين الطريقتين السابقتين تعمد على المعطيات الهندسية للوجه تم تسميتها طريقة Neuro-ACP.

أما الشطر الثاني فيتمثل في تحديد موقع الوجه في الصورة الرقمية بدون قيود مسبقة بمعنى أيا كانت أبعاد و وضعية الوجه و مهما كانت الشروط الضوئية في الصورة. استعملنا لهذا الغرض طريقة تمزج بين طريقة تحديد مكان لون البشرة في الصورة و طريقة تعتمد على البحث على الشكل العام للوجه تدعى « Template Matching ». الهدف الأساسي من هذه المرحلة في البحث هو وضع نظام يمكنه التعرف على الوجه أليا عن بعد و دون أن نحدد له مكان الوجه مسبقا.

**كلمات رئيسية :** التعرف على الوجه، خصائص الوجه، تحديد مكان الوجه، PCA، MLP ، Neuro-ACP ، Template Matching ، لون البشرة.

# RÉSUMÉ

---

Le travail que nous présentons dans cette thèse consiste à classifier et reconnaître les visages d'un ensemble de personnes intégrés dans une base de données visages créée au LAIG.

Vu la complexité des objets manipulés (Visages), nous avons utilisé pour la classification deux types de données : Les paramètres géométriques définissant les traits faciaux et les images entières des visages. Comme outil, nous avons utilisé en premier lieu le PMC connu par ses performances dans le domaine de la classification. La seconde méthode utilisée est l'analyse en composante Principale, certes c'est une méthode des plus anciennes mais elle nous a donné des résultats meilleurs que ceux du PMC.

Une troisième idée nous a menées à de bien meilleurs résultats, c'est l'hybridation de la première avec la seconde que nous avons appelé méthode Neuro-ACP.

Une autre partie très importante de notre travail, que nous considérons comme le complément de la première partie, consiste à détecter les visages dans une scène quelle que soit la position, l'éclairage, l'orientation,..etc. Pour cela, nous proposons une méthode basée sur l'hybridation de la méthode de détection de couleur de peau et la méthode du « template matching ». Cette hybridation nous a permis de contourner les inconvénients de chaque méthode et de minimiser les risques de fausses alertes dans la décision visage ou non-visage. Le but lointain de cette partie c'est d'arriver à créer un système qui procède à la classification et la reconnaissance des individus quelles que soient leurs positions dans la scène et dans n'importe quelle condition d'éclairage.

**Mots Clés :** Reconnaissance de visage, caractéristiques du visage, MLP, PCA, approche Neuro-ACP, detection de visages, couleur de peau, espace de couleur, template matching.

# REMERCIEMENTS

---

Ces quelques mots ne suffisent pas pour exprimer toute ma gratitude et mes remerciements les plus profonds à Monsieur Tebbikh Hicham, professeur au département d'Electronique de l'Université 08 Mai 1945–Guelma, pour avoir assuré la direction scientifique de mes travaux, pour les conseils et les critiques tout au long de ces années, et surtout pour avoir été là chaque fois que j'avais besoin d'aide à titre professionnelle ou autre... Merci du fond du cœur.

Je tiens à exprimer ma vive gratitude à Monsieur Noureddine DOGHMANE, Professeur au département d'Electronique de l'Université Badji-Mokhtar de Annaba pour l'honneur qu'il me fait en acceptant la présidence du jury de ma thèse.

J'adresse mes plus vifs remerciements à Monsieur Mouldi BEDDA, professeur au département d'Electronique de l'Université Badji-Mokhtar de Annaba et Monsieur Djamel CHIKOUCHE, Professeur au département d'Electronique de l'Université Farhat Abas de Sétif et Monsieur Laamri ZOUAOUI, Maître de Conférence au département d'Electronique de l'Université Farhat Abas de Sétif, qui ont accepté de faire partie de mon jury de thèse.

Je remercie également Boualleg Abd el Halim pour son apport personnel et scientifique.

*A*  
*Ma Petite*  
*Nour*

# LISTE DES FIGURES

## Chapitre 1 : Introduction

1.1	Vision Humaine	3
1.2	Classification des domaines de la vision	4
1.3	Outils d'un système de vision artificielle	4
1.4	Schéma d'un système de vision par ordinateur	5
1.5	Système de reconnaissance de visages	9

## Chapitre 2 : Réseaux de Neurones

2.1	Neurone biologique	16
2.2	Potentiel d'action	18
2.3	Propagation du potentiel d'action	18
2.4	Neurone Formel	19
2.5	Fonction de Heaviside	20
2.6	Fonction signe	20
2.7	Fonction linéaire	20
2.8	Fonction à Seuils	20
2.9	Fonction linéaire multi seuils	20
2.10	Fonctions sigmoïdes	21
2.11	Réseau Multicouche (non bouclé)	22
2.12	Réseau à connections locales	22
2.13	Réseau à connections récurrentes	23
2.14	Réseau à connections complètes	23
2.15	Réseau de Hopfield à quatre neurones	28
2.16	Réseaux de kohonen	29
2.17	Architecture du Perceptron	31
2.18	Illustration du problème du XOR	32
2.19	Architecture de l'ADALINE	33
2.20	Architecture du perceptron multicouche	34
2.21	Courbe de la zone de coût minimal dans l'espace de configuration des poids du réseau (représenter ici en deux dimensions)	43

## Chapitre 3 : Analyse en Composantes Principales

3.1	Diagrammes des dispersions	53
3.2	Cercle de corrélation	53

## Chapitre 4 : Classification des Visages

4.1	Exemples de résultats de localisation des yeux, du nez et de la bouche	62
4.2	Exemple de résultats de l'extraction des éléments caractéristiques de visages	63

4.3	Schéma représentatif des points caractérisant le visage	63
4.4	Exemple de résultats de l'extraction des distances de visages	63
4.5	Base d'apprentissage constituée de 40 personnes pour chaque personne nous avons pris 5 images dans différentes prises de vues	66
4.6	Base de test constituée des 40 personnes précédentes mais dans des prises de vues différentes de celles de la base apprentissage.	66
4.7	Evolution de l'erreur en fonction du nombre d'itération pour l'approche pixel	67
4.8	Organigramme de l'apprentissage pour l'approche pixel	68
4.9	Evolution de l'erreur en fonction du nombre d'itération pour l'approche géométrique	69
4.10	Organigramme de la méthode de l'ACP	71
4.11	Présentation des 200 visages de 40 personnes dans un plan de dimension 2	72
4.12	Graphe de distribution des valeurs propres	73
4.13	Représentation graphique des objets visages sur cp1 et cp2	74
4.14	Représentation graphique des objets visages sur cp1 et cp3	75
4.15	Résultats de la reconnaissance de deux personnes différentes	76
4.16	Représentation graphique d'une personne à reconnaître sur cp1 et cp2	76
4.17	Représentation graphique de personnes à reconnaître sur cp1 et cp3	77
4.18	Représentation graphique de personnes à reconnaître sur cp1 et cp2	77
4.19	Représentation graphique de personnes à reconnaître sur cp1 et cp3	78
4.20	Hybridation du PMC avec l'ACP	79
4.21	Evolution de l'erreur en fonction du nombre d'itérations pour l'approche neuro-ACP	80

## Chapitre 5 : Détection de Visages

5.1	Exemple d'extraction des régions de peau	85
5.2	Images utilisées pour l'apprentissage et visages propres correspondants.	87
5.3	Résultat d'une détection de mouvement réalisée par une différence entre deux images	89
5.4	Visages détectés dans le plan (x,y) et en profondeur z.	90
5.5	Système de détection et suivi de visages	90
5.6	Présentation des couleurs dans l'espace RGB	93
5.7	Exemple de normalisation d'une image RGB	93
5.8	Présentation des couleurs dans l'espace HSV	94
5.9	Application des différents opérateurs morphologiques	97
5.10	Exemples de squelettisation	98
5.11	Illustration du processus de détection de peau	99
5.12	Quelques exemples extraits de la base d'échantillons de peau	102
5.13	Distribution (en haut) et model gaussien (en bas) de peau selon les espaces RGBN, HSV et YCbCr.	103
5.14	Application du model Gaussien a) Image Originale. b) Image Probabilité " <b>likelihood</b> ". c) Image Binaire de peau.	104
5.15	Illustration du processus de filtrage des régions de peau	104
5.16	Exemples d'application des opérateurs morphologiques (Érosion et dilatation)	105
5.17	Exemple d'étiquetage des régions peau	105
5.18	Exemple de séparation d'une région	106
5.19	Estimation de la position du segment	106
5.20	Calcul des dimensions du segment	106
5.21	Exemple de filtrage avec la contrainte de ratio	108
5.22	Exemple de filtrage avec les contraintes de la surface	108



5.23	Exemples de modèles, (c) est le modèle utilisé	109
5.24	Quelques exemples extraits de la base de modèles de visages B-LAIG	109
5.25	Schéma illustratif de l'application du ``Template matching``	110
5.26	Extraction du segment visage	112
5.27	Extrait de la base test utilisée	112
5.28	Performance du système de détection de visages	114
5.29	Illustration de quelques résultats de détection de visages	115
5.30	Démonstration de quelques résultats sur notre interface utilisateur.	116

# LISTE DES TABLEAUX

---

## Chapitre 2 : Réseaux de Neurones

2.1	Historique des RNA	15
-----	--------------------	----

## Chapitre 4 : Classification des Visages

4.1	Exemples de distances de 6 personnes	64
4.2	Les différents taux de l'approche pixel	67
4.3	Les différents taux de l'approche géométrique	70
4.4	Résultats de la classification par ACP	75
4.5	Résultats de la méthode neuro-ACP	79

## Chapitre 5 : Détection de Visages

5.1	Résultats de détection de peau (Tests réalisés sur 100 images)	113
5.2	Résultats de détection de visages	114

# TABLE DES MATIÈRES

<i>Résumé en Anglais</i> .....	i
<i>Résumé en Arabe</i> .....	ii
<i>Résumé en Français</i> .....	iii
<i>Remerciement</i> .....	iv
<i>Dédicace</i> .....	v
<i>Liste des Figures</i> .....	vi
<i>Liste des Tableaux</i> .....	ix
<i>Table des Matières</i> .....	x

## Chapitre 1 : Introduction

1. Généralités.....	2
2. Vision et Intelligence Humaines.....	2
3. Vision par ordinateur.....	3
3.1 Vision artificielle.....	4
3.2 Traitement d'images.....	5
3.3 Infographie.....	5
4. Biométrie et reconnaissance de visages.....	6
4.1 Biométrie.....	6
4.2 Reconnaissance de visages.....	8
5. Etat de l'art sur la reconnaissance de visages.....	10
6. Principales Contributions.....	11
7. Production Scientifique.....	11
7.1 Publications.....	11
7.2 Communications.....	11
8. Le plan de l'ouvrage.....	12

## Chapitre 2 : Réseaux de Neurones

1. Introduction.....	14
2. Brefs rappels historiques.....	15
3. Neurone biologique.....	16
3.1 Constitution.....	16
3.2 Communication entre neurones.....	16
3.3 Remarque.....	17
4. Neurone formel.....	18
4.1 Principe de fonctionnement.....	18
4.2 Modélisation générale du neurone formel.....	19
5. Les Réseaux de neurones formels.....	21
5.1 Topologie des réseaux de neurones formels.....	21
5.2 Apprentissage des RNA.....	23
6. Quelques réseaux les plus célèbres.....	27
6.1 Le modèle de Hopfield.....	27
6.2 Les cartes auto organisatrice de Kohonen.....	29
6.3 Perceptron de ROSENBLATT.....	30
6.4 ADALINE : Adaptative Linear Elément.....	32
6.5 Le perceptron multicouche (PMC).....	33

7. Mise en œuvre des réseaux de neurones.....	39
8. Domaine d'application des réseaux de neurones.....	40
9. Limites et avantages des réseaux de neurones .....	41
9.1 Limites.....	41
9.2 Avantages.....	42
10. Aspects pratiques.....	42
10.1 Les minima locaux .....	43
11. Conclusion .....	44

### **Chapitre 3 : Analyse en Composantes Principales**

1. Introduction .....	46
2. Analyse de données .....	46
2.1 Objectifs de l'analyse de données .....	46
2.2 Nature des données analysables .....	47
2.3 Méthodes d'analyses de données multidimensionnelles .....	48
3. Analyse en Composantes Principales .....	49
3.1 Objectif .....	49
3.2 Principe .....	49
3.3 Propriétés des Composantes Principales.....	50
3.4 Formulation de l'ACP.....	51
3.5 Interprétation des résultats d'une ACP.....	53
3.6 Avantages et inconvénients de l'ACP .....	55
3.7 Domaines d'application de l'ACP .....	56
4. Classification des visages par ACP .....	57
4.1 Approche pixel .....	57
4.2 Approche géométrique .....	58
5. Identification d'un visage inconnu .....	58
6. Conclusion .....	59

### **Chapitre 4 : Classification des Visages**

1. Introduction .....	61
2. Extraction des primitives .....	61
2.1 Localisation des éléments caractéristiques du visage .....	61
2.2 Extraction des paramètres du visage .....	62
3. Base de données .....	64
4. Classification des visages .....	64
4.1 Résultat de l'approche pixel .....	65
4.2 Résultats L'approche géométrique .....	69
4.3 Résultats de la méthode d'analyse en composantes principales (ACP).....	70
4.4 Classification par méthode hybride Neuro-ACP .....	78
5. Conclusion .....	80

### **Chapitre 5 : Détection de Visages**

1. Introduction .....	82
2. Détection de visages : Etat de l'art .....	82
2.1 Approches basées sur les connaissances acquises.....	82
2.2 Approches basées sur les caractéristiques invariables .....	83

2.3	Approches basées sur la mise en correspondance .....	86
2.4	Approches basées sur l'apparence globale: Eigenface .....	86
3.	Problèmes de la détection de visages.....	88
4.	Principe de la détection de visages .....	89
4.1	Le suivi de personnes .....	89
4.2	La détection de visages .....	89
4.3	Mesure des performances d'un système de détection de visages.....	90
5.	Notion de couleur.....	91
5.1	Représentation fréquentielle .....	91
5.2	Espace de couleurs .....	92
6.	Traitement des images binaires .....	95
6.1	Images binaires .....	95
6.2	Méthodes de traitement des images binaires .....	95
7.	Méthode adoptée .....	98
7.1	Détection de la peau .....	98
7.2	Filtrage et séparation des régions .....	104
7.3	Template Matching .....	108
7.4	Mise en œuvre du modèle .....	111
8.	Résultats et interprétations .....	112
8.1	Test de détection de couleur de peau .....	113
8.2	Test de détection de visages .....	114
8.3	Discussion des résultats .....	117
9.	Conclusion .....	117
	<i>Conclusion générale</i> .....	119
	<i>Bibliographie</i> .....	122

# INTRODUCTION

---

# CHAPITRE 1

1. Généralités.....	2
2. Vision et Intelligence Humaines.....	2
3. Vision par ordinateur.....	3
3.1 Vision artificielle.....	4
3.2 Traitement d'images.....	5
3.3 Infographie.....	5
4. Biométrie et reconnaissance de visages.....	6
4.1 Biométrie.....	6
4.2 Reconnaissance de visages.....	8
5. Etat de l'art sur la reconnaissance de visages.....	10
6. Principales Contributions.....	11
7. Production Scientifique.....	11
7.1 Publications.....	11
7.2 Communications.....	11
8. Le plan de l'ouvrage.....	12

## 1. Généralités

L'homme a toujours rêvé de créer une machine capable de le remplacer dans certaines tâches tels que : L'analyse et les traitements des données (les sons, les images et autres types de signaux) où sa présence était indispensable. Son ambition était trop grande pour qu'il se contente de ces tâches qu'il trouvait primaires ; Il est alors allé jusqu'à chercher à humaniser la machine au point où elle n'est plus un simple moyen de calcul et de traitement mais elle est désormais capable de résoudre des problèmes, donner des décisions et des diagnostics. Mais la machine n'est toujours pas capable de faire tout ce dont l'homme est capable ; détecter et reconnaître des visages (individus), par exemple, est restée une tâche que la machine aussi performante soit elle n'est pas capable de la faire avec une précision comparable à celle de l'homme. Cette faculté de reconnaître des visages même vus depuis des années, et qui nous paraît si facile et si simple, reste un mystère et un énigme que l'homme cherche à comprendre et résoudre en essayant de créer une machine capable de reconnaître des visages déjà vus ou plus précisément déjà présentés à la machine par le biais d'un capteur optique (caméra, appareil photos, scanner,...etc).

Une question pertinente s'impose : Pourquoi l'homme cherche-t-il à créer une machine capable de reconnaître l'identité des personnes ?

Le sentiment de l'insécurité est à l'origine de cet axe de recherche : sécuriser ses biens (maison, commerce, banque, laboratoire,...etc) était son premier souci dont il ne peut s'en débarrasser que si la machine arrive à reconnaître les visages avec une précision comparable à celle de l'homme.

Reconnaître des visages veut dire que la machine doit être capable d'accomplir, dans l'ordre, les trois tâches suivantes :

- ✓ Détecter les personnes ainsi que leurs visages dans la scène.
- ✓ Extraire les primitives caractérisant chaque visage.
- ✓ Classifier et reconnaître l'identité de chaque personne.

## 2. Vision et Intelligence Humaines

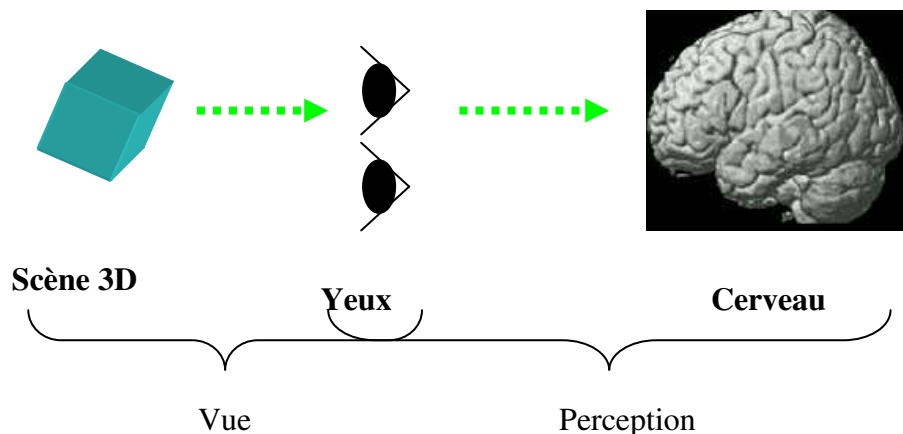
Chaque jour nous faisons des millions de gestes pour parler, manger, marcher, travailler...etc , des gestes qui sont parfois difficiles et complexes pourtant à notre naissance le seul geste que nous connaissions est la tété. Alors, comment avons nous appris tout cela ?

En effet, l'homme est doté d'une intelligence incomparable qu'il doit à de toutes petites cellules (microscopiques) appelées « neurones » que notre cerveau en contiendrait plusieurs

centaines de milliards. Cet élément minuscule n'est pas à l'origine de la gigantesque taille mémoire uniquement mais aussi c'est grâce à lui que nous pouvons reconnaître des formes et identifier des personnes en décrivant les moindres détails caractérisant leurs visages. En réalité, si nous arrivons à reconnaître nos semblables ce n'est pas grâce aux neurones (au cerveau) seulement mais aussi nous devons cela à notre système visuel qui est responsable de la mise en correspondance entre les objets 3D du monde réel et les images emmagasinées dans notre mémoire. De façon plus générale, le cerveau, en collaboration avec les cinq sens, emmagasine les odeurs, les sons et les voix, les goûts et les saveurs et surtout les images parmi lesquelles les visages vus depuis notre naissance et le long de notre existence [He00].

En effet, pour récolter des informations sur le monde qui l'entour, l'être humain dépend à 99% de sa vision qui est défini comme étant la perception des objets ou des scènes 3D par l'organe de la vue « les yeux ». Elle englobe les quatre fonctions suivantes [Gb05] :

- ❖ Vision des couleurs (peut être perçue par 1 ou 2 yeux).
- ❖ Vision des formes (peut être perçu par 1 ou 2 yeux).
- ❖ Vision de la profondeur et des distances (doit être perçue par les 2 yeux).
- ❖ Vision des mouvements (doit être perçue par les 2 yeux).

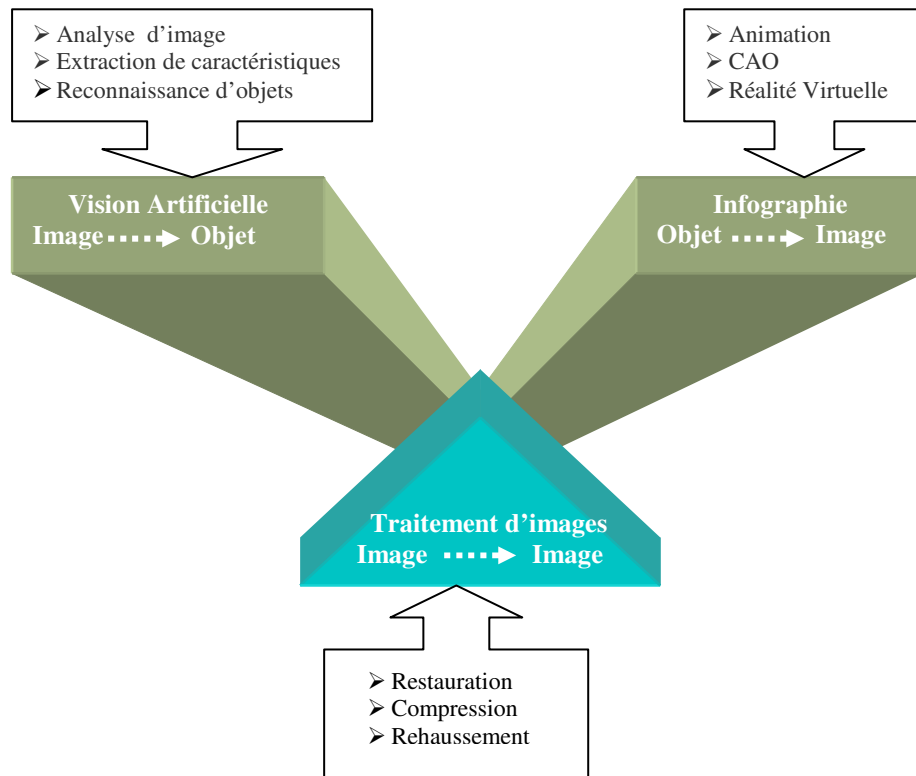


**Figure 1.1 : Vision Humaine**

### 3. Vision par ordinateur

On peut définir le domaine de la vision comme le regroupement de trois disciplines : la vision artificielle (computer vision), le traitement d'image (image processing) et l'infographie (computer graphics) [LM99].





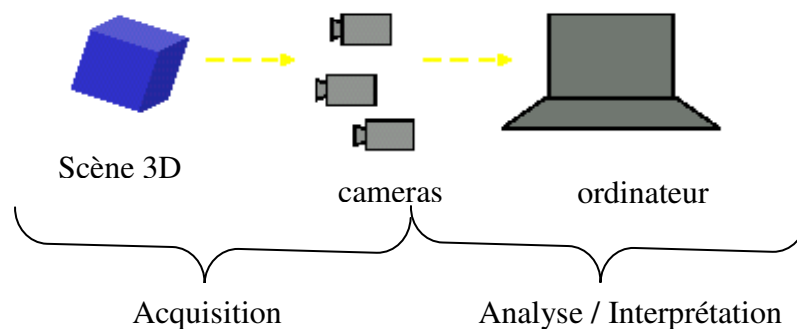
**Figure 1.2 :** Classification des domaines de la vision

### 3.1 Vision artificielle

La vision humaine est à l’origine des motivations des chercheurs en vision par ordinateur. Elle consiste à doter un ordinateur d’outils (matériels et logiciels) qui lui permettent de reproduire autant que possible les mécanismes (physiques et cognitifs) de la vision humaine (voir figure 1.3).

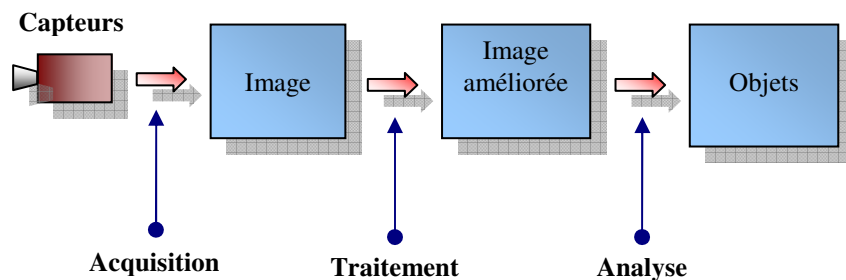
La vision artificielle [LM99] [Gb05] est alors une discipline qui consiste à convertir une image en données objets ou, plus explicitement, à identifier les objets contenus dans l’image par l’extraction et l’analyse de caractéristiques abstraites (features) à partir des pixels, suivant un processus de reconnaissance de forme similaire à celui opéré par l’humain où :

- L’œil est une caméra très performante.
- Le cerveau est un ordinateur très performant.



**Figure 1.3 :** Outils d’un système de vision artificielle

Un système de vision peut être considéré comme un processus de traitement d'information : un stimulus est présenté à l'entrée du système. L'information ainsi reçue est traitée afin d'en extraire une représentation symbolique. Cette dernière pourra, par la suite, être utilisée comme entrée à un processus de traitement de plus haut niveau ou actionner des mécanismes de navigation ou d'adaptation à l'environnement [He00].



**Figure 1.4 :** Schéma d'un système de vision par ordinateur

Parmi les applications de la vision par ordinateur, nous citons : le contrôle de la qualité sur une chaîne de production, l'identification d'un individu par biométrie (par ex., l'analyse automatisée des empreintes digitales), le diagnostic médical et la classification de terrains à partir d'images satellites.

### 3.2 Traitement d'images

Le traitement d'image ne fait pas appel, a priori, à des techniques d'intelligence artificielle comme la vision par ordinateur. Cette discipline consiste plutôt à reconstituer ou transformer une image en une autre, plus adaptée à un traitement ultérieur. Il englobe les domaines de la restauration, de l'accentuation et de la compression des « données images ».

### 3.3 Infographie

L'infographie est la discipline la plus moderne des trois. Il s'agit de la génération d'images artificielles à partir de données objets. Cette discipline touche, entre autres, les domaines de l'animation par ordinateur, de la conception et du dessin assistés par ordinateur (CAO-DAO), et de la réalité virtuelle. Des applications se retrouvent par exemple dans l'industrie du divertissement ainsi que dans les domaines de l'architecture 3D et de la simulation robotique.

En pratique, il est impossible de tirer une ligne nette entre les trois disciplines; des techniques et outils de l'une sont souvent utilisés dans une autre. Cela rend le domaine de la vision très vaste et complexe [LM99].

## 4. Biométrie et reconnaissance de visages

La croissance des communications internationales, tant en volume qu'en diversité (déplacements physiques, transactions financières, accès aux services...), implique le besoin de s'assurer de l'identité des individus. En effet, l'importance des enjeux peut motiver les fraudeurs à mettre en échec les systèmes de sécurité existants qui utilisent tous un identifiant externe tel que : badge/carte, clé, code,...etc. Le défaut commun à tous les systèmes d'authentification est que l'on identifie un objet (code, carte...) et non la personne elle-même. Face à la contrainte de l'authentification par « objets », la biométrie apporte simplicité et confort aux utilisateurs. Il existe donc un intérêt grandissant pour les systèmes électroniques d'identification et de reconnaissance. Leur dénominateur commun est le besoin d'un moyen simple, pratique, fiable et peu onéreux de vérifier l'identité d'une personne sans l'assistance d'un tiers [LM99].

### 4.1 Biométrie

Le mot Anglais « **Biometric** », utilisé pour définir « La mesure des éléments morphologiques des humains », est fréquemment traduit en français par « **Biométrie** ». D'après le Petit Robert : La Biométrie est une « science qui étudie à l'aide de mathématiques (statistiques, probabilités) les variations biologiques à l'intérieur d'un groupe déterminé ». La biométrie est donc une discipline qui s'intéresse à la mesure des caractéristiques physiques d'êtres vivants et à leur traitement statistique et les termes "biométrie" et "biométrique" se rapportent à des dispositifs destinés à reconnaître des êtres humains à partir de mesures effectuées automatiquement [Pf00].

Notons aussi que, lorsqu'il est question de mesurer des organismes humains, le terme anthropométrie est également utilisé.

#### 4.1.1 Techniques d'analyse

Il existe 2 catégories de technologies biométriques :

- Les techniques d'analyse du comportement qui englobent surtout :
  - La dynamique de la signature (la vitesse de déplacement du stylo, les accélérations, la pression exercée, l'inclinaison),
  - La façon d'utiliser un clavier d'ordinateur (la pression exercée, la vitesse de frappe).
- Les techniques d'analyse de la morphologie humaine qui s'intéressent aux empreintes digitales, à la forme de la main, à la voix, au dessin du réseau veineux de l'œil, et aux traits du visage. Ces éléments ont l'avantage d'être stables dans la vie d'un individu et ne subissent pas autant les effets du stress par exemple, que l'on retrouve dans l'identification comportementale [Pf00].

### 4.1.2 Domaines d'utilisation de la biométrie

La liste des applications pouvant utiliser la biométrie pour contrôler un accès (physique ou logique) peut être très longue. La taille de cette liste n'est limitée que par l'imagination de chacun dans son domaine d'activité. A titre indicatif, nous pouvons citer [Pf00] :

- a. Contrôle d'accès aux locaux,
  - salle informatique,
  - site sensible (service de recherche, site nucléaire).
- b. Systèmes d'informations,
  - lancement du système d'exploitation,
  - accès au réseau,
  - commerce électronique,
  - transaction (financière pour les banques, de données entre entreprises),
  - tous les logiciels utilisant un mot de passe.
- c. Equipements de communication,
  - terminaux d'accès à Internet,
  - téléphones portables.
- d. Machines et équipements divers,
  - coffre-fort avec serrure électronique,
  - distributeur automatique de billets,
  - casier sensible (club de tir, police),
  - cantine d'entreprise (pour éviter l'utilisation d'un badge par une personne extérieure),
  - casier de piscine (plus d'objet à porter sur soi),
  - contrôle des adhérents dans un club, carte de fidélité,
  - contrôle des temps de présence,
  - voiture (anti-démarrage).
- e. Etat / administration,
  - fichier judiciaire,
  - titres d'identité (carte nationale d'identité, passeport, permis de conduire),
  - services sociaux (sécurisation des règlements),
  - services municipaux (sécurisation des accès aux écoles),
  - système de vote électronique.

### 4.1.3 Moyens actuels à l'étude biométrique

Actuellement la recherche biométrique s'intéresse aux moyens suivants :

- les empreintes digitales,
- la forme de la main,

- **le visage,**
- la voix,
- l'iris,
- la rétine,
- la signature dynamique,
- le code génétique (ADN).

Dans notre travail, nous nous intéressons à l'un des moyens biométriques les plus complexes : le visage. Sa complexité réside dans : la géométrie et la forme de ses caractéristiques qui peuvent changer suivant l'humeur ou l'état de la personne ainsi que sa chrominance qui a tendance à changer selon l'intensité de la source lumineuse et la position du visage par rapport à cette dernière ou le changement climatique qui engendre un changement considérable de la couleur de peau [Pf00].

## 4.2 Reconnaissance de visages

Sans aucune hésitation, l'homme est capable de distinguer les visages de tout autre « objet ». Mieux, il est capable d'en reconnaître des centaines, malgré de très faibles différences structurelles [GB05].

Mais qu'est ce qu'un visage ? Ceci pourra constituer un exemple de définition :

*« Un visage est une structure tridimensionnelle possédant une configuration « externe » - le contour de la face modelé par les saillies osseuses et souligné par les cheveux- où s'inscrit une configuration interne formé par l'assemblage des traits. Certaines particularités s'y ajoutent comme la pilosité, les lunettes, la texture, la coloration de la peau....etc ».*

Sauf exception, tous les visages ont une structure voisine ; ne diffèrent que sur d'infimes modifications de la configuration externe ainsi que de la forme des éléments internes et de leurs rapports. Malgré ces faibles différences, et qu'un visage donné ne se présente jamais de la même façon, le système de perception humain arrive à reconnaître facilement des milliers de visages. Pour comprendre le système de perception humain, deux groupes de chercheurs se sont réunis ; les neuropsychologues et les cognitivistes, leur but était de répondre aux questions suivantes :

- ❖ Comment le système visuel humain traite-t-il un visage ?
- ❖ Comment est-il stocké en mémoire et comment un visage devient familier ?
- ❖ Comment se généralise le processus de reconnaissance pour de nouvelles images ?
- ❖ y a-t-il, dans le cerveau, un mécanisme spécifique pour la reconnaissance des visages ou est-ce qu'il s'agit d'un seul et d'un même mécanisme pour tout type de formes.

Bien que certaines questions attendent toujours des réponses cette initiative a ouvert les portes aux chercheurs de vision et intelligence artificielles pour concevoir une machine capable de reconnaître des individus par le biais de leurs visages «Reconnaissance automatique des visages ».

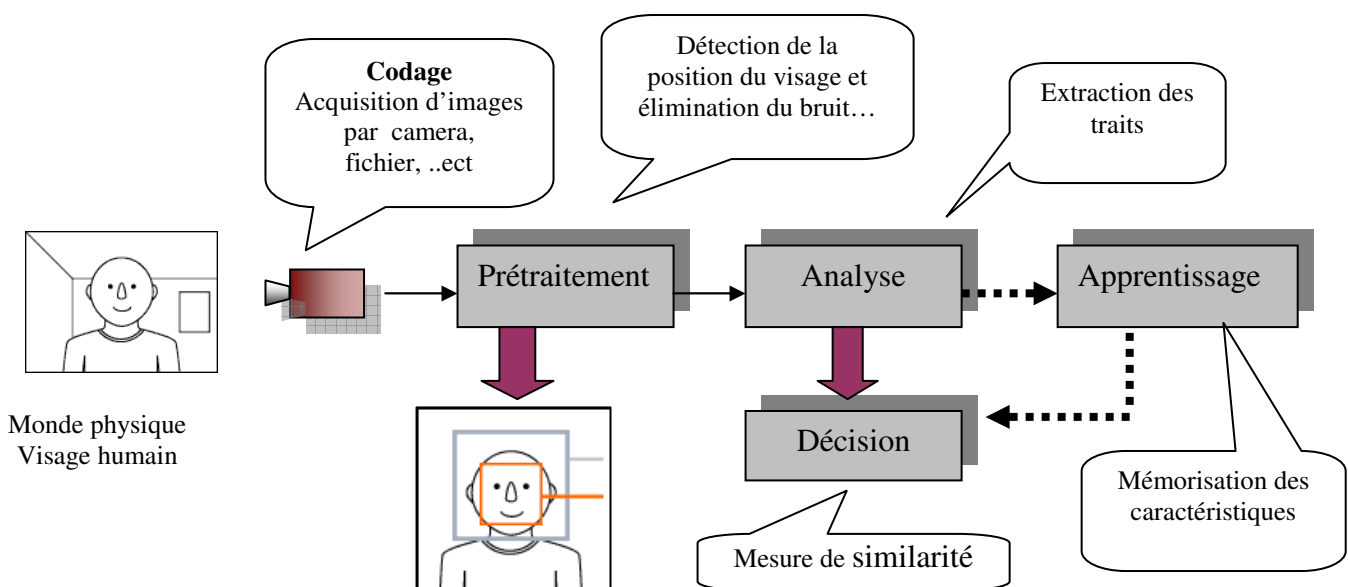
Tout processus automatique de reconnaissance des visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage est une entité dynamique qui change constamment sous l'influence de plusieurs facteurs. La figure (I.5) illustre la démarche générale adoptée pour réaliser de tels systèmes [Fe97].

Dans **le monde physique**, il y a trois paramètres à considérer : L'éclairage, la variation de posture et l'échelle. La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents.

**Le codage** consiste en l'acquisition de l'image et sa digitalisation, il comporte un risque de bruit et donne lieu à une représentation 2-D (la matrice des niveaux de gris) pour un objet 3-D (le visage). Peu de méthodes 3-D ont été proposées [NU92] [LC02], mais cela devrait changer dans le futur proche avec l'invention des Range-Cameras qui peuvent mesurer la topologie 3-D du visage.

Dans **le prétraitement**, il faut éliminer le bruit par des techniques de traitement et de restauration d'images et procéder à une localisation de la tête dans l'image. C'est une opération très complexe car nous devons procéder dans toute image ou vidéo acquise à la recherche d'un ou plusieurs visages suivant la forme et/ou la couleur de peau, pour cela il faudra faire une synthèse des couleurs de peau.

La restauration d'images ou l'élimination du bruit consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image [CF97].



**Figure 1.5 :** Système de reconnaissance de visages

Dans l'**étape Analyse** (appelée aussi indexation, représentation, modélisation ou extraction de caractéristiques), il faut extraire de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase « décision ». Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes.

**L'apprentissage** consiste à mémoriser les représentations calculées dans la phase Analyse pour les individus connus.

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

## 5. Etat de l'art sur la reconnaissance de visages

La complexité du visage en tant qu'objet renforce et encourage les chercheurs à vouloir mieux modéliser cet objet si riche en caractéristiques. Ceci à créer une diversité dans les points d'intérêts des chercheurs. Certains ont concentrés leurs recherches sur la détection de la tête (visage) dans la scène. Dans ce contexte, les travaux les plus connus se sont basés sur la détection du mouvement dans la scène suivi de la recherche de la tête (objet visage). L'objet en question peut exister en plusieurs exemplaires (plusieurs visages dans la scène) comme il peut se présenter de profil ou dans une position quelconque. Par exemple [JL02] et [He02] ont basé leurs recherches sur la géométrie de la tête et du visage. Ils trouvent que la tête se présente sous une forme ellipsoïdale, ce qui leur a conduit à utiliser des transformées telle que la transformée de Hough pour la détection de toutes les ellipses qui peuvent exister dans la scène. D'autres, tels que [GP04] et [JS99] basent leurs recherches sur la chrominance du visage et ont procédé plutôt à la recherche de la couleur de peau dans la scène.

La difficulté de la tâche de détection du visage a poussé certains chercheurs à supposer l'existence ainsi que la position du visage dans la scène ce qui leurs a permis de se concentrer sur l'extraction des traits du visage. Ici la majorité des chercheurs se sont mis d'accords sur le fait que le visage est une association de caractéristiques géométriques. Certains comme [LC96] se sont intéressés aux distances entre les éléments caractéristiques du visage (yeux, nez, bouche,...etc), d'autres [BP93], pour augmenter la précision, ont cherché à extraire les contours de ces éléments. Une autre catégorie de méthodes été basée sur l'aspect globale du visage : [BT00] par exemple utilisent une approche connexionniste qui donne de bon résultats mais son inconvénient majeur été le temps de calcul très lent, d'autre comme [SH94] adoptant une approche stochastique avance, que quand les images frontales sont balayées de haut en bas il y a un ordre naturel dans lequel apparaissent les caractéristiques, pour justifier leurs utilisation du modèle de Markov caché (HMM).

Les traitements de haut niveau c.à.d la classification et la reconnaissance de visages demeure la tâche la plus difficile vu le nombre considérable de méthodes de classifications disponibles et le choix reste critique car choisir la meilleure méthode de classification ne garantie

pas les meilleurs résultats puisque cette étape dépend en grande partie des étapes précédentes (extraction des traits). Malgré les travaux considérables effectués dans ce contexte, jusqu'à ce jour la machine n'est pas encore arrivée à reconnaître les personnes ou plus précisément distinguer les visages, avec un degré de confiance proche à celui de l'homme.

## 6. Principales Contributions

Ce travail, réalisé dans le cadre d'un travail d'équipe, a porté sur les différentes étapes d'un système de reconnaissance de visages. Notre contribution a porté essentiellement sur :

- ✓ L'extraction des primitives géométriques caractérisant chaque visage de la base de données réalisée au LAIG.
- ✓ La classification et la reconnaissance de l'identité de chaque personne par réseaux de neurones, ACP et par une nouvelle méthode hybride, proposée par notre équipe, qui est la méthode neuro-ACP.
- La détection des visages dans une scène, qui est une phase nécessaire et importante avant toute opération de reconnaissance de visages, fixes ou dynamiques, représente une première étape des futures perspectives de ce travail.

## 7. Production Scientifique

### 7.1 Publications

- 1) C. Bencheriet et H. Tebbikh, « Extraction des éléments caractéristiques du visage par approche géométrique », Revue des sciences et technologie : Synthèse, Publication de l'Université Badji-Mokhtar, Vol. 11, pages 60 à 70, juin 2002.
- 2) Ch. Bencheriet, A.H. Boualleg and H. Tebbikh, « Hybrid PMC-PCA Approach For Automatic Faces Recognition », WSEAS TRANSACTIONS on SIGNAL PROCESSING Journal, Issue N°1, Volume N°2, pp 37-44, January 2006.

### 7.2 Communications

- 1) C/E Bencheriet et H. Tebbikh « Approche géométrique pour la classification et la reconnaissance de visages », Conférence sur la Vision Artificielle CVA' 2000, 18-20 Novembre 2000, Université Mouloud Mammeri-Tizi-Ouzou, Faculté des Sciences de l'Ingénieur.
- 2) C/E Bencheriet et H. Tebbikh « Application des descripteurs de Fourier à la détection de contours », Conférence sur le Génie Electrique « CGE'01 », Ecole Militaire Polytechnique, Bordj El bahri, 25-26 Décembre 2001.
- 3) A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, « Classification et Reconnaissance de Visages par Réseaux de Neurones », 1<sup>ères</sup> Journées sur les Scientifiques et Technologies Avancées (JSTA'2003), 24-25 Mai 2003, Guelma.



- 4) A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, «Reconnaissance des visages par la méthode d'analyse en composante principale », Journée sur l'Informatique Industrielle et l'Imagerie, Guelma, le 18 Avril 2005.
- 5) A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, «Reconnaissance automatique de visages par réseaux de neurones et analyse en composante principale », , Colloque National sur le Traitement du Signal et ses Applications (CNTSA'2005) , 19 & 20 septembre 2005, Université 08 mai 45 Guelma, Algérie.
- 6) A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, «Approche Géométrique pour la Reconnaissance Automatique de Visages par une Approche Neuro-ACP », International Symposium on Signal and its Applications (IS3PA'005), Mostaganem, 28 septembre 2005.
- 7) Ch. Bencheriet, A.H. Boualleg and H. Tebbikh, « Automatic Recognition of Faces by Neural Networks and Principal Component Analysis», WSEAS, Tenerife, Spain, December 16-19, 2005.
- 8) A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, « Reconnaissance de Visage Par Une Approche Hybride Neuro-ACP », ICCMD'06 International Conference on Control Modelling and Diagnosis Annaba Algeria, May 22-24, 2006.

## 8. Le plan de la thèse

Dans cette thèse nous traitons les différentes étapes d'un système de reconnaissance de visages. Les chapitres 2 et 3 seront consacrés respectivement aux réseaux de neurones et à l'analyse en composante principale. Le chapitre 4 sera réservé à nos principales contributions, qui sont (entre autres) l'extraction des primitives (caractéristiques géométriques de l'objet « visage ») et la classification par réseaux de neurones, par ACP et par Neuro-ACP. Cette dernière approche que nous proposons est une hybridation du PMC avec l'ACP. Nous ferons aussi une comparaison avec la classification des données pixels. Comme introduction aux perspectives projetées pour ce travail, nous présentons, dans le chapitre 5, une étude sur l'image et son pré-traitement ainsi que les espaces de couleurs suivie de la description d'une méthode, basée sur la couleur de la peau et la méthode du « template matching », que nous proposons pour la détection du visage dans une scène.

# RÉSEAUX DE NEURONES

## CHAPITRE 2

1. Introduction.....	14
2. Brefs rappels historiques.....	15
3. Neurone biologique.....	16
3.1 Constitution.....	16
3.2 Communication entre neurones.....	16
3.3 Remarque.....	17
4. Neurone formel.....	18
4.1 Principe de fonctionnement.....	18
4.2 Modélisation générale du neurone formel.....	19
5. Les Réseaux de neurones formels.....	21
5.1 Topologie des réseaux de neurones formels.....	21
5.2 Apprentissage des RNA.....	23
6. Quelques réseaux les plus célèbres.....	27
6.1 Le modèle de Hopfield.....	27
6.2 Les cartes auto organisatrice de Kohonen.....	29
6.3 Perceptron de ROSENBLATT.....	30
6.4 ADALINE : Adaptative Linear Elément.....	32
6.5 Le perceptron multicouche (PMC).....	33
7. Mise en œuvre des réseaux de neurones.....	39
8. Domaine d'application des réseaux de neurones.....	40
9. Limites et avantages des réseaux de neurones.....	41
9.1 Limites.....	41
9.2 Avantages.....	42
10. Aspects pratiques.....	42
10.1 Les minima locaux.....	43
11. Conclusion.....	44

## 1. Introduction

Le cerveau humain est considéré comme le siège de l'intelligence, de la créativité, de l'émotivité, de la conscience et de la mémoire. Il est doté d'un nombre de caractéristiques absentes dans une architecture de type Von Neumann voire des machines parallèles ; en particulier notre cerveau possède :

- Une architecture massivement parallèle,
- Un mode de calcul et une mémoire distribués,
- Une capacité d'apprentissage,
- Une capacité de généralisation,
- Une capacité d'adaptation,
- Une résistance aux pannes,
- Une faible consommation énergétique.

Il est donc tentant d'étudier des systèmes basés sur les réseaux de neurones pour voir si certaines des caractéristiques de notre cerveau peuvent être « récupérées ».

Les ordinateurs actuels outrepassent les capacités humaines dans le calcul numérique et la manipulation de symboles. Cependant, les humains peuvent, sans effort, résoudre des problèmes complexes liés à la perception (reconnaissance d'un visage dans une foule) à une telle vitesse et avec une telle efficacité que le plus rapide ordinateur passe pour un "nain" dans ces domaines. Quelle peut-être la raison d'une telle différence ? L'architecture neuronale humaine n'a rien à voir avec une architecture du type Von Neumann, et cette différence affecte significativement la nature des problèmes que résolvent ces deux modèles avec efficacité.

Les efforts réalisés pour développer des programmes « intelligents » basés sur une architecture Von Neumann n'ont pas été couronnés de succès au sens où aucun programme intelligent généraliste n'a vu le jour. Inspirés des réseaux de neurones biologiques, les RNA sont des systèmes massivement parallèles, constitués de processeurs simples et interconnectés. Le modèle RNA cherche à utiliser des principes organisationnels que l'on pense être mis en oeuvre dans le cerveau. De plus, modéliser un système nerveux biologique à l'aide d'un RNA peut augmenter la connaissance des fonctions biologiques. Une étude extensive des RNA nécessite des compétences dans des domaines aussi variés que la neurophysiologie, les sciences cognitives, la psychologie cognitive, la physique et plus particulièrement la mécanique statistique, la théorie du contrôle, l'informatique, l'intelligence artificielle, les mathématiques et statistiques, la reconnaissance de formes, la vision artificielle, le calcul parallèle, le « hardware » (digital, analogique, VLSI, optique). Chaque découverte dans l'un de ces domaines alimente le champ d'études des RNA. D'un autre côté,

les RNA créent une dynamique dans ces différentes disciplines en apportant de nouveaux outils ou de nouvelles approches.

## 2. Brefs rappels historiques

La recherche dans le domaine des RNA a vécu trois périodes d'intense activité. La première, dans les années (19)40 est liée aux travaux de McCulloch et Pitts [MP43]. La seconde, durant les années (19)60 avec les travaux de Rosenblatt sur le perceptron et le théorème de convergence [Ro62] et les travaux de Minsky et Papert [MP69] démontrant les limites du perceptron mono-couche. Les résultats de Minsky et Papert eurent pour effet de refroidir l'enthousiasme de la plupart des chercheurs, et notamment des informaticiens. Le creux de la vague dura environ 20 ans, jusqu'au début des années (19)80, avec l'approche proposée par Hopfield [Ho82] et l'algorithme de rétro-propagation dans le cadre des Perceptron Multi-Couches (PMC) proposé pour la première fois par Werbos [We74], réinventé par la suite plusieurs fois avant d'être popularisé par Rumelhart et al. [RM86] en 1986. Le livre d'Anderson et Rosenfeld [AR88] fournit avec force détails l'historique des développements dans le domaine des RNA.

1890	W. James	Introduit le concept de mémoire associative
1942	Norbert Wiener	Boucle de rétroaction cybernétique
1943	McCulloch et Pitts	Modèle formel de neurone biologique
1949	Hebb	Explique le conditionnement chez l'animal par les propriétés des neurones eux-mêmes. Loi pouvant expliquer le phénomène d'apprentissage
1957	Rosenblatt	Perceptron, premier neuro-ordinateur, première tentative de reproduire le cerveau de l'homme
1960	B. Widrow	Développe le modèle Adaline à l'origine de l'algorithme de rétropropagation de gradient très utilisé aujourd'hui avec les Perceptrons multicouches.
1969	Minsky et Papert	Démontrent qu'une classe de problème n'est pas résolvable par le perceptron. Ils donnent les limitations théoriques du perceptron et de la neuronique. Les financements se dirigent vers les systèmes experts.
1982	John Hopfield	On lui doit le renouveau d'intérêt pour les réseaux de neurones artificiels: perceptron multicouche
1983	Boltzmann	La Machine de Boltzmann contourne les limitations du perceptron
1986	Werbos puis Rumelhart, Hinton et Williams	Algorithme d'apprentissage applicable au perceptron de type récuratif (back propagation). La classe de problème de Minsky et Papert est résolvable
1989	kohonen	Auto-organisation et mémoire associatives
1992	Boughanem	Réseaux de neurones appliqués à la recherche d'information

**Tableau 2.1** : Historique des RNA

### 3. Neurone biologique

#### 3.1. Constitution

Un neurone (cellule nerveuse) est une cellule biologiquement spécialisée pour le traitement de l'information. Il se compose en 4 parties principales [To92] (figure 2.1) :

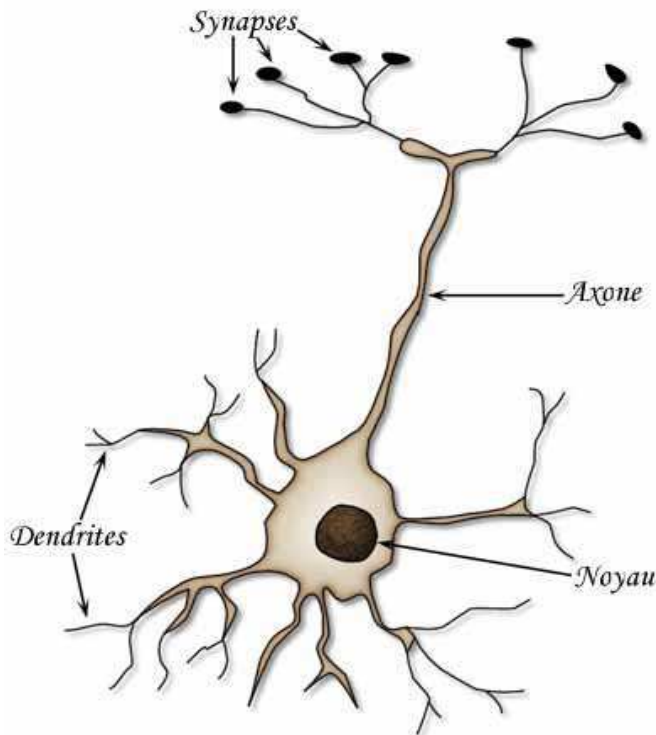


Figure 2.1 : Neurone biologique

- ❖ Les **dendrites**, sur lesquelles les autres cellules entrent en contact synaptique : c'est par les dendrites que se fait la réception des signaux.
- ❖ Le **corps de la cellule**, c'est l'unité de traitement.
- ❖ L'**axone**, où passent les messages accumulés dans le corps de la cellule : l'envoi de l'information se fait par l'axone.
- ❖ Les **synapses** par lesquelles la cellule communique avec d'autres cellules, ce sont des points de connexion par où passent les signaux de la cellule.

#### 3.2. Communication entre neurones

Un neurone stimulé envoie des impulsions électriques ou **potentiels d'action**, à d'autres neurones. Ces impulsions se propagent le long de l'**axone** unique de la cellule. Au point de contact entre neurones, les synapses, ces impulsions sont converties en signaux chimiques. Quand l'accumulation des excitations atteint un certain seuil, le neurone engendre un potentiel d'action, d'une amplitude d'environ 100mV et pendant une durée de 1ms (figure 2.2). Le neurone émettant le

signal est appelé **neurone pré-synaptique** et celui recevant ce signal, **neurone post-synaptique** (figure 2.3).

En fait, le fonctionnement synaptique va s'ajuster en fonction des signaux reçus et médiés, ce qui permet de dire que les synapses « apprennent » leur « comportement » à partir des activités dans lesquelles elles sont impliquées (ou à partir des signaux qui leur sont imposés). Ce comportement synaptique façonné par « l'histoire » de la synapse fonctionne donc de manière similaire à une « mémoire » et pourrait même être éventuellement responsable de la mémoire humaine. Chez l'homme, le cortex cérébral est une pellicule de neurones d'environ 2 à 3 millimètres d'épaisseur et d'une surface de l'ordre de  $2200 \text{ cm}^2$ , c'est-à-dire environ deux fois la surface d'un clavier. Le cortex contient environ  $10^{11}$  neurones [BL90]. Les neurones sont interconnectés entre eux de manière bien plus complexe et bien plus dense que le réseau téléphonique ne l'est. On évalue le nombre de connexions entre un neurone et ses voisins entre  $10^3$  et  $10^4$ , ce qui fait que le cortex humain contient entre  $10^{14}$  et  $10^{15}$  connexions. Les neurones communiquent en émettant des trains de potentiels rapides et très courts de l'ordre de quelques millisecondes. Le message est modulé en terme de fréquences qui peuvent varier de quelques hertz à quelques centaines de hertz, soit un million de fois plus lentement que la plus rapide bascule dans un circuit électronique actuel. Pourtant les décisions perceptives complexes, telles que la reconnaissance de visages, se font en quelques centaines de millisecondes, sur un réseau dont la vitesse est de l'ordre de la milliseconde. Ce qui indique que notre cerveau, dans de telles tâches perceptives, exécute des programmes qui n'excèdent pas la centaine d'instructions parallèles. C'est ce que l'on appelle la *règle des 100 pas* [FF88]. Des considérations temporelles du même ordre montrent que la quantité d'informations transmise d'un neurone à l'autre doit être très petite (quelques bits), et donc que l'information critique n'est pas transmise directement, mais qu'elle est contenue et répartie dans les interconnexions, d'où le nom de modèle connexionniste utilisé pour décrire des RNA.

### 3.3 Remarque

Les propriétés physiologiques des neurones qui viennent d'être exposées ne reflètent pas la complexité de la réalité : seules les propriétés de base utilisées dans le cadre du connexionnisme de base sont décrites. Ces dernières sont exploitées dans le modèle formel proposé par MacCulloch et Pitts. Le neurone est considéré comme un automate à seuil, qui suivant une excitation et un certain seuil, répond ou non par un potentiel d'action.

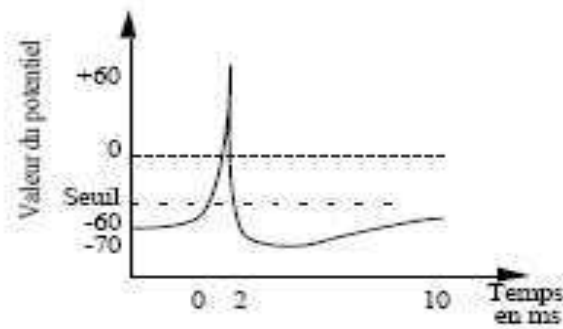


Figure 2.2 : Potentiel d'action

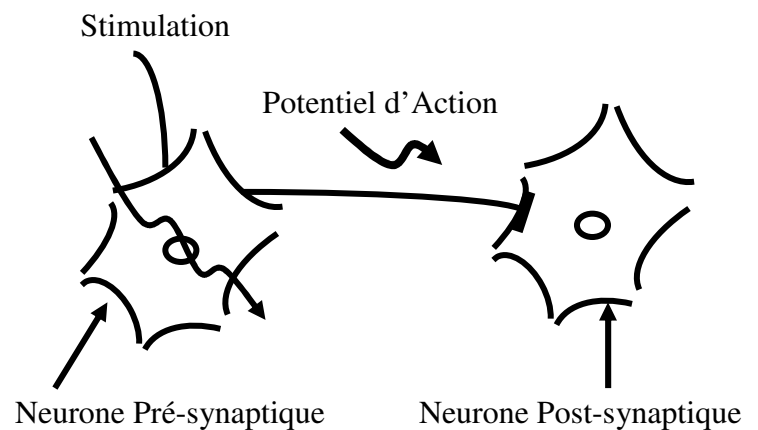


Figure 2.3 : Propagation du potentiel

#### 4. Neurone formel

C'est dans l'article " A Logical Calculus of the Ideas Immanent in Nervous Activity " paru dans le journal «Bulletin of mathematical » en 1943 qu'est apparu pour la première fois la notion de neurone formel [MP43].

A l'époque, Mc Culloch étudie le fonctionnement du cerveau des chimpanzés. Il se convainc que les circuits d'influx nerveux étaient analogues à la propagation logique d'un raisonnement.

Avec l'aide de Pitts, il va présenter le neurone formel :

- Il est binaire : il est soit actif soit inactif.
- Chaque neurone possède un seuil de déclenchement fixé.
- Le neurone reçoit des entrées de synapse excitatrices, ayant des poids identiques.
- Le neurone peut aussi recevoir des entrées de synapses inhibitrices, dont l'action est absolue ; c'est à dire que, si la synapse est active, le neurone ne peut pas être actif.
- Un quantum de temps est nécessaire à l'intégration des entrées synaptiques, basé sur le délai observé physiologiquement.

Le neurone formel est un modèle mathématique simplifié du neurone biologique, il présente un certain nombre d'entrées, les dendrites, un corps traitant les entrées suivant la méthode du tout ou rien, et un axone véhiculant la réponse du neurone.

##### 4.1. Principe de fonctionnement

Chaque entrée  $X_n$  est affectée d'un poids dit poids synaptique  $W_n$ . Le passage des entrées dans le corps du neurone se fait en deux étapes. La première étape consiste à faire une somme pondérée des entrées par les poids respectifs des connexions sur lesquelles ces entrées se propagent. La seconde étape consiste à calculer l'image de cette somme pondérée par une fonction de transfert

binaire. Le résultat obtenu provoque ou non le déclenchement d'un potentiel d'action suivant le dépassement d'un seuil, et sert à son tour d'entrée à d'autres neurones (figure 2.4).

$$U = \sum_{j=0}^N W_j X_j - \theta$$

On notant par  $X = [-1 \ X_1 \ X_2, \dots, X_n]^T$  le vecteur des entrées augmenté et par  $W = [\theta, W_1, W_2, \dots, W_n]$  le vecteur des poids augmentés, il vient  $S = f(W^T, X)$  [Go00].

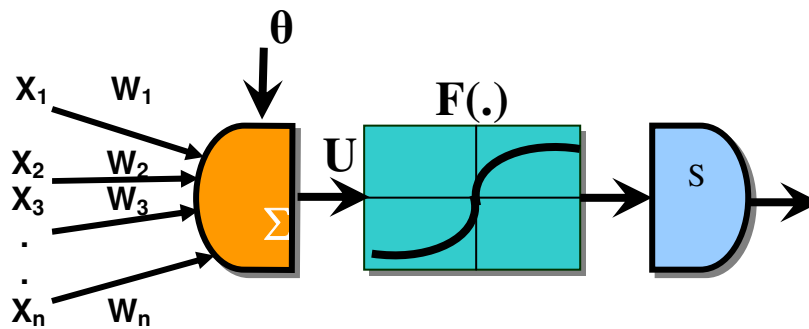


Figure 2.4 : Neurone Formel

Un neurone formel peut être défini par les cinq éléments suivants :

- La nature de ses entrées.
- La fonction d'entrée totale  $U$  définissant le prétraitement effectué sur les entrées.
- La fonction d'activation, ou d'état,  $F$  définissant l'état interne du neurone en fonction de son entrée totale.
- La fonction de sortie  $S$  calculant la sortie du neurone en fonction de son état d'activation.
- La nature de la sortie du neurone.

#### 4.2 Modélisation générale du neurone formel

- ❖ Les entrées et les sorties peuvent être : binaires (-1, +1) ou (0, 1) ou réelles.
- ❖ La fonction d'entrée totale  $U$  peut être: booléenne, linéaire  $\sum_{j=0}^N W_j X_j$ , affine  $\sum_{j=0}^N W_j X_j - \theta$  ou polynomiale de degré supérieur à deux.
- ❖ La fonction de sortie est en général considérée comme la fonction **identité**, la sortie du neurone est généralement considérée comme l'activation.
- ❖ La fonction d'activation (de transfert) : Peut prendre plusieurs formes.

##### a- Fonction binaire à seuil :

La fonction d'activation (de transfert) utilisée dans le modèle de McCulloch & Pitts est la fonction *échelon* (figure 2.5). Elle fait passer l'activation du neurone d'une valeur à une autre dès que l'entrée résultante dépasse un certain seuil. L'inconvénient de cette fonction est



qu'elle n'est pas différentiable, ce qui pose un problème pour les algorithmes basés sur le gradient [MP43].

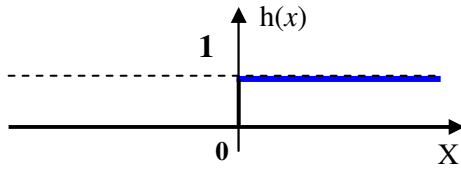


Figure 2.5 : Fonction de Heaviside

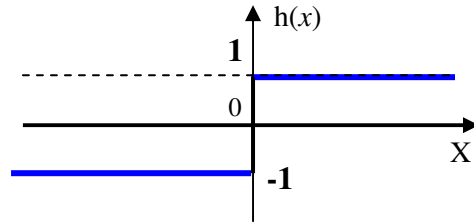


Figure 2.6 : Fonction signe

**b- Fonction Linéaire :**

Dans ce modèle, la fonction d'activation est la fonction identité. S'il a l'avantage d'être simple, il n'a que peu de rapport avec la réalité (le signal de sortie est non borné, linéaire par rapport aux signaux d'entrée, ce qui ne correspond pas du tout au fonctionnement des neurones biologiques).

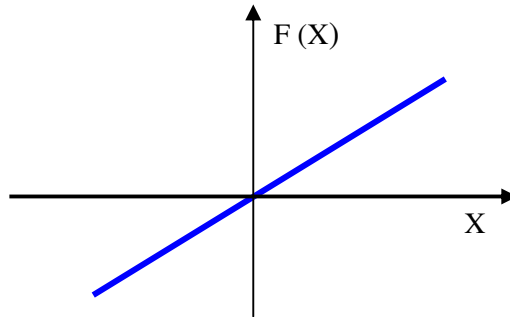


Figure 2.7 : Fonction linéaire

**c- Fonction Linéaire à seuil ou à multi seuils:**

Elle rend déjà mieux compte de la réalité, en faisant apparaître des phénomènes de saturation

$$Saturation(X) = \begin{cases} X & \text{si } X \in [U, V] \\ U & \text{si } X \leq U \\ V & \text{si } X > V \end{cases}$$

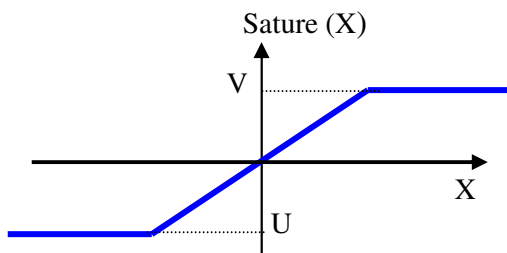


Figure 2.8 : Fonction à Seuils

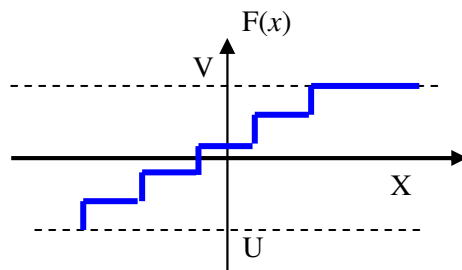


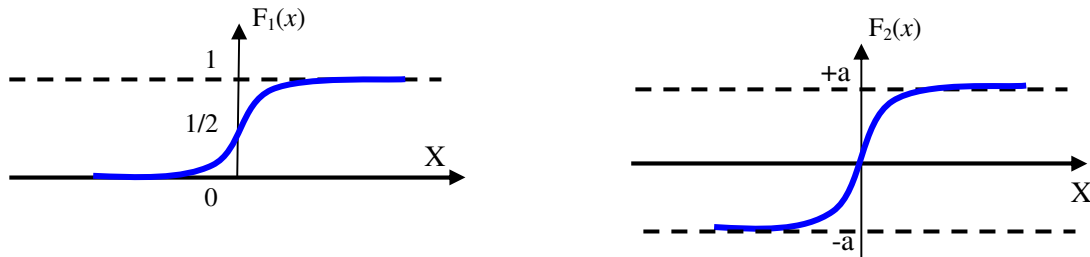
Figure 2.9 : Fonction Linéaire multi seuils

**d- La Fonction sigmoïde:**

C'est la fonction la plus utilisée dans la pratique

$$F_1(X) = 1/(e^{-X} + 1)$$

$$F_2(X) = a(e^{kx} - 1)/e^{kx} + 1$$



**Figure 2.10** : Fonctions sigmoïdes

Les fonctions sigmoïdes ont la propriété d'être différentiables, ce qui est nécessaire pour les algorithmes basés sur le gradient. Une autre propriété intéressante est le fait que les fonctions dérivées peuvent s'exprimer facilement à l'aide des fonctions elles-mêmes, ce qui permet un gain significatif de temps de calcul.

**5. Les Réseaux de neurones formels**

Les réseaux de neurones artificiels sont les regroupements en réseaux d'un certain nombre de neurones formels connectés entre eux de diverses manières [DN93]. Un réseau est défini par :

- Sa topologie, qui représente le type de connexion existant entre les divers neurones du réseau.
- La fonction de transfert qui caractérise le neurone (étudiée au paragraphe 4.2).
- Les méthodes d'apprentissage utilisées pour faire apprendre au réseau des couples d'entrées-sorties.

**5.1 Topologie des réseaux de neurones formels**

Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Elle peut être quelconque, mais le plus souvent on utilise des réseaux de structure régulière pour faciliter leurs utilisations. L'architecture des réseaux de neurones formels peut aller d'une connectivité totale (tous les neurones sont reliés les uns aux autres), à une connectivité locale où les neurones ne sont reliés qu'à leurs plus proches voisins.

Il existe de nombreuses topologies possibles, mais celles qui ont eu à ce jour plus de notoriété sont les suivantes [DN93] :

### 5.1.1 Réseau multicouche "feed-forward"

Les neurones sont rangés par couches, Il n'y a pas de connexion entre les neurones d'une même couche et les connexions ne se font qu'avec les neurones des couches avales (figure 2.9). Habituellement, chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement. Ceci nous permet d'introduire la notion de sens de parcours de l'information (de l'activation) au sein d'un réseau et donc définir les concepts de neurone d'entrée, neurone de sortie. Par extension, on appelle couche d'entrée l'ensemble des neurones d'entrées et couche de sortie l'ensemble des neurones de sortie. Les couches intermédiaires n'ayant aucun contact avec l'extérieur sont appelées couches cachées [DN93].

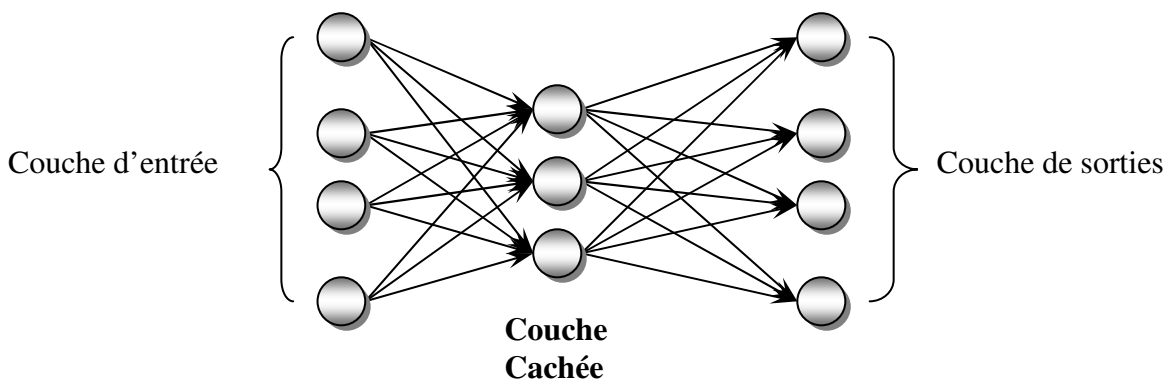


Figure 2.11 : Réseau Multicouche (non bouclé)

### 5.1.2 Réseau à connexions locales

Il s'agit d'une structure multicouche, mais qui à l'image de la rétine, conserve une certaine topologie. Chaque neurone entretient des relations avec un nombre réduit et localisé de neurones de la couche avale (figure 2.10). Les connexions sont donc moins nombreuses que dans le cas d'un réseau multicouche classique [ Jo94][DN93].

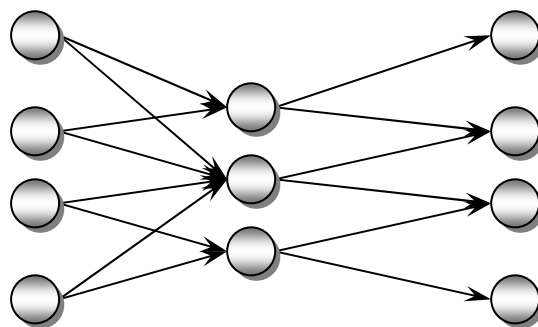


Figure 2.12: Réseau à connexions locales

### 5.1.3 Réseau à connexions récurrentes

Les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales (figure 2.13) [Jo94][DN93].

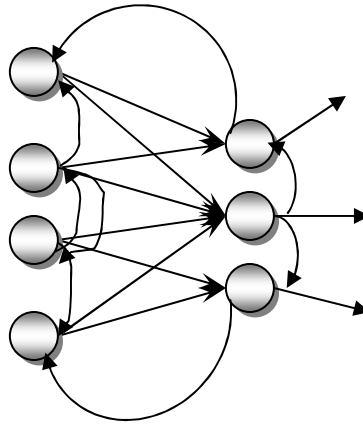


Figure 2.13 : Réseau à connexions récurrentes

### 5.1.4 Réseau à connexions complètes " Feedback "

C'est la structure d'interconnexion la plus générale (figure 2.14). Chaque neurone est connecté à tous les neurones du réseau (et à lui-même) [Jo94][DN93].

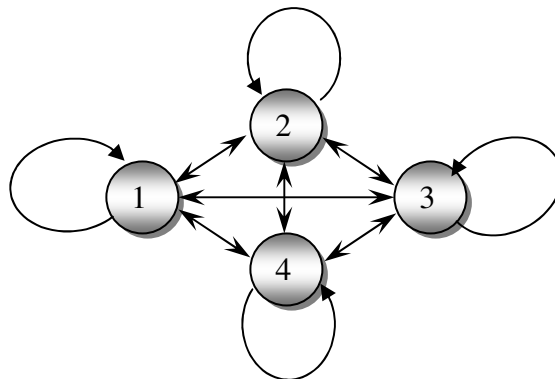


Figure 2.14 : Réseau à connexions complètes

## 5.2 Apprentissage des RNA

### 5.2.1 Définition

La capacité à apprendre est une des caractéristiques fondamentales de l'intelligence. Bien qu'une définition précise de ce qu'est l'apprentissage soit difficile à fournir, un processus d'apprentissage dans le cadre des RNA peut être vu comme le problème de la mise à jour des poids des connexions mises en œuvre dans un réseau afin de réussir la tâche qui leur est assignée. En général, un réseau doit apprendre les poids de ses connexions à partir d'un ensemble de données d'entraînement. L'efficacité du réseau est améliorée en modifiant de façon itérative les poids du réseau. La capacité pour les RNA à apprendre automatiquement à partir d'exemples les rend

extrêmement attractifs. Au lieu de suivre un ensemble de règles spécifiées par un expert humain, les RNA apprennent les règles sous-jacentes (telle que la relation entre entrées et sorties) à partir d'un ensemble représentatif d'exemples. C'est là l'un des avantages majeurs de cette approche par rapport aux systèmes experts. Pour apprendre/comprendre ou définir un processus d'apprentissage, il est nécessaire de modéliser l'environnement dans lequel opèrera le RNA, c'est-à-dire qu'il faut savoir quel type d'information sera fournie au réseau. Cette modélisation est le *paradigme d'apprentissage*. Ensuite, il faut comprendre comment le réseau met à jour les poids de ses connexions, c'est-à-dire quelles sont les *règles d'apprentissage* qui gouvernent la mise à jour. Un *algorithme d'apprentissage* est la procédure dans laquelle les règles d'apprentissage sont utilisées en vue de l'ajustement des poids [Jo94].

On peut considérer qu'il existe trois principaux paradigmes d'apprentissage : supervisé, non supervisé et hybride [DN93].

#### 5.2.1.1 Apprentissage supervisé

Un superviseur, ou professeur, fournit au réseau des couples d'entrées-sorties. Il fait apprendre au réseau l'ensemble de ces couples, par une méthode d'apprentissage, comme la rétro-propagation du gradient de l'erreur, en comparant pour chacun d'entre eux la sortie effective du réseau et la sortie désirée. L'apprentissage est terminé lorsque tous les couples entrée-sortie sont reconnus par le réseau. Ce type d'apprentissage se retrouve, entre autres, dans le perceptron.

#### 5.2.1.2 Apprentissage non supervisé

Cet apprentissage consiste à détecter automatiquement des régularités qui figurent dans les exemples présentés et à modifier les poids des connexions pour que les exemples ayant les mêmes caractéristiques de régularité provoquent la même sortie. Les réseaux auto-organiseurs de Kohonen sont les réseaux à apprentissage non supervisé les plus connus.

#### 5.2.1.3 Apprentissage hybride (semi-supervisé)

L'apprentissage hybride, quant à lui, combine les deux approches ; une partie des poids étant déterminée via un apprentissage supervisé et une autre à l'aide d'un apprentissage non supervisé. L'utilisateur ne possède que des indications imprécises (par exemple : échec/succès du réseau) sur le comportement final désiré.

L'apprentissage nécessite en général une grande quantité de données, que l'on regroupe dans un corpus d'apprentissage. Selon la technique d'apprentissage utilisée, d'autres corpus sont aussi employés, notamment pour mesurer la validité de la solution trouvée par le réseau. On appelle ces corpus supplémentaires des corpus de test ou de généralisation.

Dans la grande majorité des réseaux de neurones actuels, l'apprentissage du réseau s'effectue lors d'une période « dite d'apprentissage » préliminaire à son utilisation. Lors de son

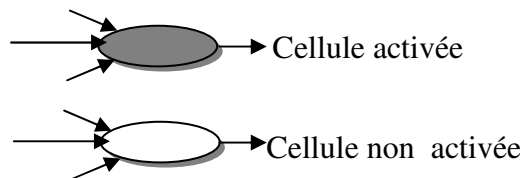
utilisation subséquente, les poids synaptiques son figés, et l'apprentissage interdit. On appel généralisation la capacité du réseau à réagir correctement lors de son utilisation à des entrées qui n'ont pas été vues lors de l'apprentissage. Ce sont les capacités de généralisation d'un réseau neuromimétique qui son la clef de son utilisation en pratique [Jo94].

### 5.2.2 Règles d'apprentissages des RNA

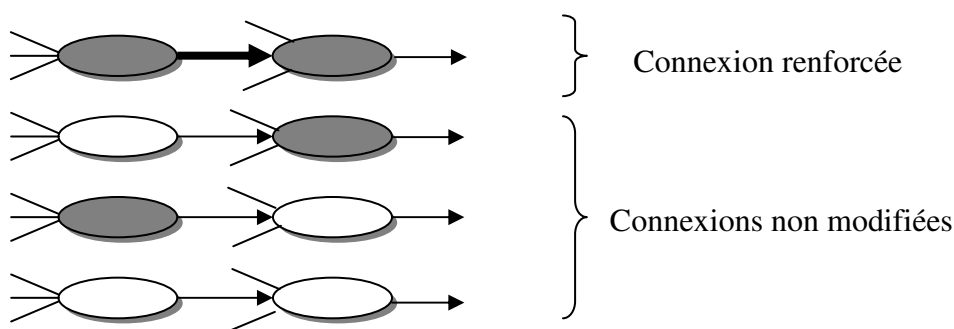
Dans les systèmes experts, les connaissances de l'expert ont une forme énumérée : elles sont exprimées sous forme de règles. Dans le cas des réseaux de neurones, les connaissances ont une forme distribuée : elles sont codées dans les poids des connexions, la topologie du réseau, les fonctions de transfert de chaque neurone, le seuil de ces fonctions et la méthode d'apprentissage utilisée. Il existe un certain nombre de règles d'apprentissage :

#### 5.2.2.1 La règle de Hebb

C'est la méthode d'apprentissage la plus ancienne (1949), elle est inspirée de la biologie. Elle traduit le renforcement des connexions liant deux neurones activés. Si un des deux neurones au moins n'est pas activé, le poids de la connexion n'est pas modifié [DN93]. C'est-à-dire que le poids  $w_{ij}$  d'une connexion entre un neurone  $i$  et un neurone  $j$  augmente quand les deux neurones sont activés en même temps et il n'est pas modifié dans le cas contraire. Si nous prenons, les conventions suivantes :



La règle de Hebb donne alors :



En se basant sur ce principe, Hebb a donné la règle d'apprentissage suivante :

$$w_{ij}(t + \partial t) = w_{ij}(t) + \mu \cdot A_i A_j$$

Avec :

$w_{ij}(t)$  et  $w_{ij}(t + \partial t)$  : Les poids de la connexion entre le neurone  $i$  et le neurone  $j$  aux instants " $t$ " et " $t + \partial t$ ".

$A_i$  et  $A_j$  : l'activation du neurone  $i$  et l'activation de neurone  $j$ .

$\mu$  ( $\mu > 0$ ) : est un paramètre de l'intensité de l'apprentissage.

### 5.2.2.2 Règle de Widrow Hoff

La règle de Widrow Hoff ou règle delta proposée en 1960, consiste à modifier, à chaque pas, les poids et les biais afin de minimiser la somme des carrés des erreurs en sortie en utilisant la règle suivante [MB01] :

$$w(k+1) = w(k) + \eta(t_k - y_k)x_k^T \quad (2.1)$$

A chaque pas d'apprentissage  $k$ , l'erreur en sortie est calculée comme la différence entre la cible recherchée  $t$  et la sortie  $y$  du réseau.

$$E_K(w) = \frac{1}{2} \sum_{k=1}^m (t_k - y_k)^2 \quad (2.2)$$

Le gradient de cette quantité par rapport à la matrice de poids  $w$  est donné par :

$$\frac{\nabla E_k}{w} = \frac{1}{2} \cdot \frac{\nabla(y_k y_k^T - 2y_k^T t_k)}{w} \quad (2.3)$$

Le gradient se calcul comme suit :

$$\frac{\nabla E_k}{w} = \frac{\partial E_k}{\partial w} = \frac{\partial E_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial w} \quad (2.4)$$

D'après l'expression de  $E_k$  et avec  $y_k = wx_k + b$  les dérivées partielles sont :

$$\frac{\partial E_k}{\partial y_k} = y_k - t_k \quad (2.5)$$

$$\frac{\partial (wx_k + b)}{\partial w} = x_k^T \quad (2.6)$$

La mise à jour des poids se fait par l'équation :

$$w(k+1) = w(k) - \eta \cdot \frac{\nabla E_k}{w} \quad (2.7)$$

avec  $\eta$  : le gain d'apprentissage ( $0 < \eta < 1$ ).

De même, on obtient l'expression de la modification du biais :

$$b(k+1) = b(k) + \eta(t_k - y_k) = b(k) - \eta \cdot \frac{\nabla E_k}{w} \quad (2.8)$$

### 5.2.2.3 La rétro-propagation du gradient de l'erreur

Cet algorithme est utilisé dans les réseaux de type feed-forward, ce sont des réseaux de neurones à couches, ayant une couche d'entrée, une couche de sortie et au moins une couche cachée. Il n'y a pas de récursivité dans les connexions et pas de connexions entre neurones de la même couche. Le principe de la rétro-propagation consiste à présenter au réseau un vecteur d'entrées et de procéder au calcul de la sortie par propagation à travers les couches, de la couche d'entrée vers la couche de sortie en passant par les couches cachées. Cette sortie obtenue est comparée à la sortie désirée, une erreur est alors obtenue. A partir de cette erreur, est calculé le gradient de l'erreur qui est à son tour propagé de la couche de sortie vers la couche d'entrée, d'où le terme de rétro-propagation. Cela permet la modification des poids du réseau et donc l'apprentissage. L'opération est répétée pour chaque vecteur d'entrée et cela jusqu'à ce que le critère d'arrêt soit vérifié [To92].

### 5.2.2.4 Les algorithmes génétiques

Ils représentent une modélisation de la sélection naturelle. Une population d'individus est générée aléatoirement. Un certain nombre d'individus répondant le mieux aux critères de sélection est choisi. A partir de cette population d'élites, une nouvelle population est générée par reproduction, mutation, ou crossover sur les individus de départ, ou parents. L'opération est recommencée jusqu'à la vérification du critère d'arrêt [To92].

### 5.2.2.5 Algorithme de Sollis et Wets

C'est une méthode stochastique d'optimisation. Soit  $M$  un vecteur contenant l'ensemble des paramètres à optimiser. Soit  $G$  un vecteur de bruit gaussien. Ce dernier sera chargé de faire évoluer les paramètres à optimiser par somme ou différence avec le vecteur  $M$ . Soit  $B$  un vecteur de biais contenant la moyenne de  $G$ , il mémorisera les réussites de la minimisation de la fonction de coût. Cette méthode converge avec une probabilité de 1 vers le minimum global [To92].

## 6. Quelques réseaux les plus célèbres

### 6.1 Le modèle de Hopfield :

Le modèle de Hopfield a été proposé en 1982 afin de mémoriser des formes et des motifs. Il est surtout utilisé en tant que mémoire associative (en anglais, Content Addressable Memory par analogie avec le pointeur qui permet de récupérer le contenu d'une case mémoire). C'est un réseau récurrent dont les cellules sont complètement interconnectées ; il n'y a plus la notion de couche. Les



neurones sont à sorties binaires (à deux états -1 et 1 ou 0 et 1), dont la loi d'apprentissage est la règle de Hebb [DN93].

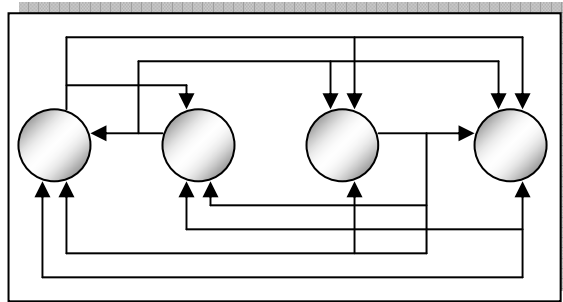


Figure 2.15 Réseau de Hopfield à quatre neurones

### 6.1.1 Architecture du réseau

L'architecture d'un réseau de Hopfield est semblable à l'auto associeur linéaire. Deux différences importantes, toutefois, le caractérisent :

- La réponse binaire, c'est ce qui le rend non linéaire et la mise à jour des cellules est asynchrone [HJ94],
- L'activation  $a_j$  de la  $j^{\text{eme}}$  cellule d'un réseau de Hopfield se calcule comme :

$$a_j = \sum_i^N x_i w_{ij} \quad (2.9)$$

avec :

$a_j$  : activation de  $j^{\text{eme}}$  cellule du réseau,

$x_N$  : valeur de sortie +1 ou -1 de la  $i^{\text{eme}}$  cellule du réseau,

$w_{ij}$  : intensité de la connexion entre la  $i^{\text{eme}}$  cellule et la  $j^{\text{eme}}$  cellule, notons que  $w_{ij}=0$ ,

La fonction d'activation du réseau est la fonction signe, notée  $\text{sgn}$  :

$$a_j = x_j = \text{sgn } a_j = \begin{cases} -1 & \text{pour } a_j < \theta \\ +1 & \text{pour } a_j \geq \theta \end{cases} \quad (2.10)$$

où  $\theta$  est le seuil.

### 6.1.2 Règle d'apprentissage

La règle d'apprentissage proposée par Hopfield pour son modèle n'est autre que la règle de Hebb.

## 6.2 Les cartes auto organisatrice de Kohonen

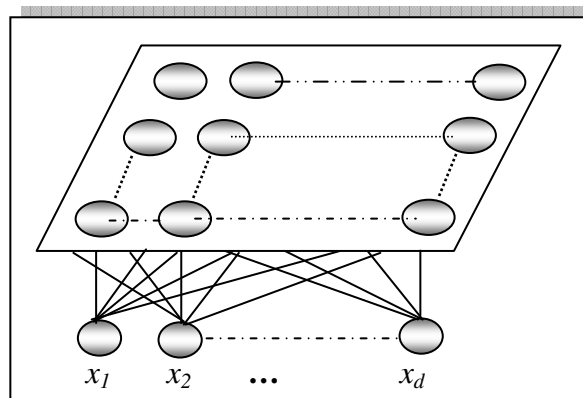
Le modèle a été proposé dès les années 70 par Von des Marsberg, puis par Kohonen. Ce modèle a pour objectif de représenter des données complexes. Il est réalisé à partir d'un réseau de  $K$  neurone à  $n$  entrées et donc  $k$  sorties.

Les entrées sont des vecteurs à  $n$  composantes toutes connectées totalement aux  $k$  neurones du réseau par  $nk$  connexions modifiables. Les neurones du réseau sont placés en général dans un espace de 1 ou 2 dimensions.

### 6.2.1 Architecture du réseau

Les neurones d'une carte auto organisatrice sont disposés en une seule couche, à partir de laquelle est amenée une topologie définie par une notion de voisinages des cellules.

La figure 2.16 illustre un tel réseau dans sa configuration la plus courante, le cas bidimensionnel.



**Figure 2.16 :** Réseaux de Kohonen

L'entrée unique du réseau est commune à tous les neurones et le vecteur de caractéristiques  $x$  représente un objet. A chaque neurone sont associées des coordonnées  $(x, y)$  indiquant sa position sur la carte ainsi qu'un vecteur de poids synaptiques [Go00].

### 6.2.2 Règle d'apprentissage

L'apprentissage de la carte auto organisatrice est basé sur deux constructions biologiques essentielles dans le cerveau, à savoir que :

1. Chaque cellule nerveuse correspond à un stimulus spécifique,
2. Il existe une région d'intense activité (la plus stimulée).

Le modèle d'enchaînement de la carte consiste donc à :

- ❖ Sélectionner le neurone correspondant le mieux à un signal d'entrée donné ;
- ❖ Induire dans un voisinage de l'élue de la région d'intense activité.

Dans un premier temps, l'apprentissage est non supervisé et consiste à répéter les étapes suivantes :

- Présenter un objet  $x$  à l'entrée du réseau, sans préciser la classe à laquelle il appartient ;
- Rechercher le neurone  $q'$  dont le vecteur de poids est plus proche, au sens de la distance euclidienne, du vecteur d'entrée :

$$q' = \min_q \frac{1}{2} (x - w_q)^T (x - w_q) \quad (2.11)$$

- Adapter le vecteur de poids de ce neurone ainsi que ceux de ses voisins topologiques de manière à ce qu'ils se rapprochent d'avantage du vecteur d'entrée :

$$w_q(k+1) = w_q(k) + \eta (x - w_q(t)) \quad \text{si } q \in V(q') \quad (2.12)$$

$$w_q(k+1) = w_q(k) \quad \text{sinon}$$

où  $V(q')$  désigne le voisinage du neurone ( $q'$ ) dans lequel les vecteurs de poids cellules sont adaptés et  $\eta$  ( $\eta > 0$ ) est le taux d'apprentissage .

Les cartes de Kohonen peuvent être utilisées dans le cadre de la projection de données multi-variées, d'approximation de densité ou de classification. Elles ont été utilisées avec succès en reconnaissance de la parole, traitement d'images, robotique, contrôle de processus. Les paramètres à définir dans ce cas sont :

- La dimension des tableaux de cellules,
- Le nombre de cellules dans chaque dimension,
- La forme du voisinage,
- L'échéancier de réduction de la taille du voisinage,
- Le coefficient d'apprentissage.

### 6.3 Perceptron de ROSENBLATT :

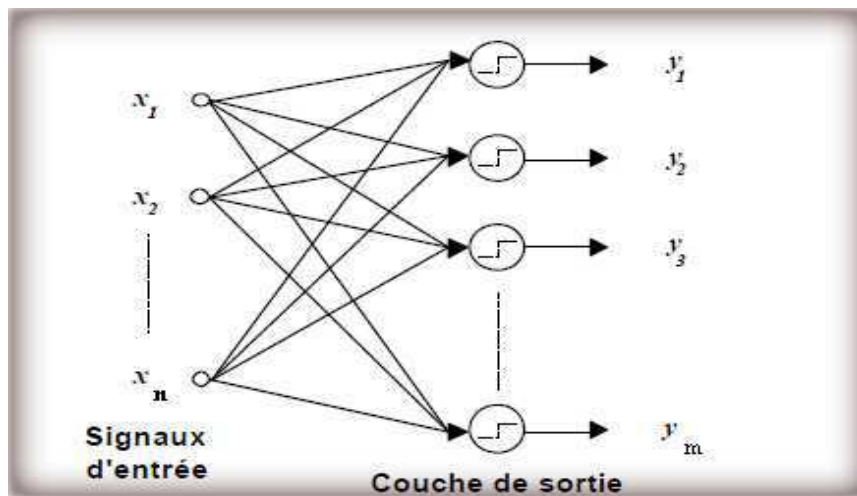
Présenté originellement par Rosenblatt, en 1958, le perceptron est la forme la plus simple du réseau de neurones. Il permet de classifier correctement des objets appartenant à deux classes linéairement séparables. C'est un réseau mono couche, inspiré du système visuel [To92][DN93].

#### 6.3.1 Architecture du réseau

Le perceptron se compose de deux couches de neurones (figure 2.17), la première était appelée à l'origine la rétine du perceptron, quant à la deuxième couche donne la réponse du perceptron correspondant à la stimulation présentée en entrée.

Les neurones de la première couche répondent par oui ou non, la réponse oui correspond à une valeur 1 en sortie pour le neurone, la réponse non correspond à une valeur 0. Les

cellules de la rétine sont reliées aux cellules de sortie grâce à des synapses d'intensité variable.



**Figure 2.17** : Architecture du Perceptron

L'apprentissage pour le perceptron se fait en modifiant l'intensité de ses synapses. Les cellules de sortie évoluent l'intensité de la simulation en provenance des cellules de la rétine en effectuant la somme des intensités des cellules actives, avec la formule suivante :

$$a_j = \sum_i x_i w_{ij} \quad (2.13)$$

avec :

$a_j$  : activation de la  $j^{\text{ème}}$  cellule de la couche de sortie.

$x_j$  : valeur de sortie (0 ou 1) de la  $i^{\text{ème}}$  cellule de la rétine

$w_{ij}$  : intensité de la connexion entre  $i^{\text{ème}}$  cellule de la rétine et la  $j^{\text{ème}}$  cellule de la couche de sortie.

Les cellules de sortie deviennent actives si leur degré d'activation dépassé un seuil noté  $\theta_j$  généralement fixé à 0, mais sans que cela soit nécessaire.

### 6.3.2 Règle d'apprentissage

Les paramètres du perceptron, c'est à dire les points synaptiques des neurones, peuvent être déterminés grâce à un entraînement supervisé, effectué sur un ensemble de formes pré classifiées. La règle d'apprentissage du perceptron prend le nom de « règle de Widrow Hoff ». Pour apprendre, le perceptron doit savoir qu'il a commis une erreur, et il doit connaître la réponse qu'il aurait du donner. De ce fait, on parle d'apprentissage supervisé. L'apprentissage se produit jusqu'à ce que le perceptron soit capable de donner toutes les réponses correctes [To92][DN93].

La règle d'apprentissage de Widrow Hoff s'écrit de la manière suivante :

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \eta(t_j - o_j)x_i = w_{ij}^t + \Delta w_{ij} \tag{2.14}$$

avec :

$\Delta w_{ij}$  : changement à effectuer pour la valeur  $w_{ij}$ ,

$x_i$  : valeur de sortie (0 ou 1) de la  $i^{\text{ème}}$  cellule de la rétine,

$o_j$  : réponse de la  $j^{\text{ème}}$  cellule de sortie (0 ou 1),

$t_j$  : réponse technique (ou désirée) de la  $j^{\text{ème}}$  cellule de la couche de sortie (0 ou 1),

$w_{ij}^{(t)}$  : intensité de la connexion entre la cellule de la rétine et la  $j^{\text{ème}}$  cellule de sortie au temps  $t$

$\eta$  : une constante positive généralement comprise entre 0 et 1.

### 6.3.3 Limite du PERCEPTRON

- Les sorties du réseau seront limitées à deux valeurs (vues ses fonctions de transfert).
- Il ne peut traiter que les problèmes linéairement séparables cette propriété des neurones à seuil constitue une très forte limite à leurs possibilités de classification. En effet, il est clair qu'un neurone à seuil ne peut réaliser la fonction booléenne 'OU Exclusif' (XOR), et ce, quels que soit les poids de ses connexions (figure 2.18) [DN93]. Cependant, il a été démontré que si le problème est linéairement séparable, le perceptron donne toujours des résultats satisfaisants.

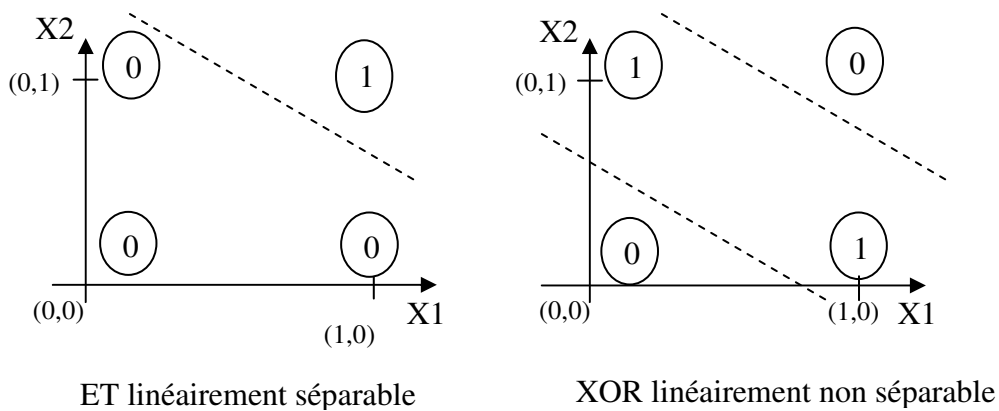


Figure 2.18 : Illustration du problème du XOR

### 6.4 ADALINE : Adaptive Linear Element :

Ce modèle est proposé par Widrow et Hoff, est un perceptron linéaire sans couche cachée donc à un seul neurone, qui fonctionne comme un séparateur linéaire, l'ADALINE est

utilisé dans les décisions binaires et dans l'association d'un modèle d'entrée à une sortie : il est capable de reconnaître la lettre Y par exemple quelle que soit sa taille et son orientation.

### 6.4.1 Architecture du réseau

L'ADALINE est un neurone qui possède des valeurs d'activation continues et une fonction d'activation linéaire.

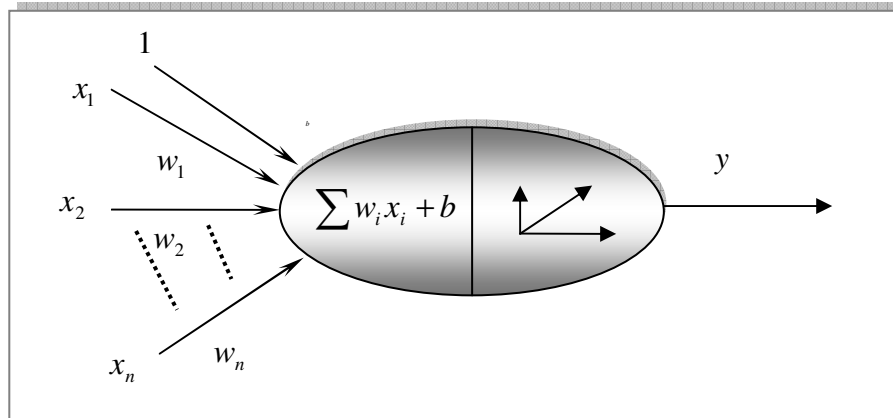


Figure 2.19 : Architecture de l'ADALINE

### 6.4.2 Règle d'apprentissage

La règle d'apprentissage de l'ADALINE, appelée règle de Widrow Hoff (règle des Moindres carrées) utilise la descente de gradient de la fonction de la somme des erreurs quadratiques, donnée par :

$$E(w) = \frac{1}{2} \sum_{k=1}^m (t_k - y_k)^2 \quad (2.15)$$

avec

$$y_k = w^t x_k \quad (2.16)$$

La descente du gradient donne la direction qui minimise  $E(w)$  :

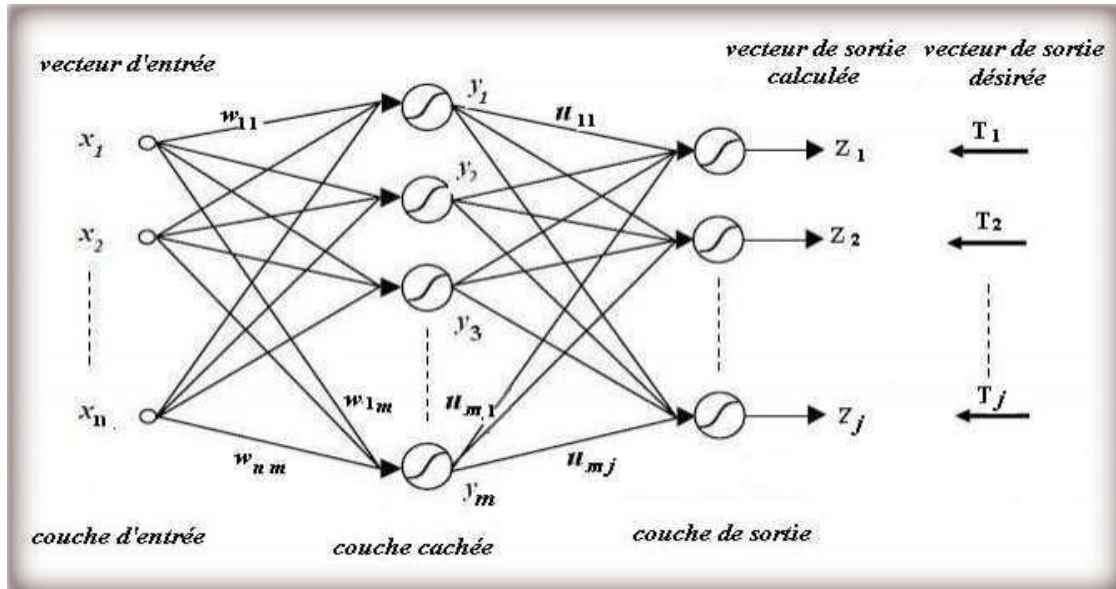
$$w(t+1) = w(t) - \eta \nabla E(w) \quad (2.17)$$

### 6.5 Le perceptron multicouche (PMC)

Le perceptron multicouche est une amélioration du perceptron comprenant une ou plusieurs couches intermédiaires dites cachées. Il utilise, pour modifier ses poids, un algorithme de rétro propagation du gradient, qui est une généralisation de la règle de Widrow Hoff. Il s'agit toujours de minimiser l'erreur quadratique, ce qui est assez simple quand on utilise une fonction  $f$  dérivable (la sigmoïde par exemple), et consiste à propager la modification des poids de la couche de sortie jusqu'à la couche d'entrée.

### 6.5.1 Architecture du réseau

L'architecture est obtenue par la mise en cascade de perceptrons (figure 2.20).



**Figure 2.20** : Architecture du perceptron multicouche

Lorsque le vecteur de caractéristiques d'un objet est présenté à l'entrée du réseau, il est communiqué à tous les neurones de la première couche. Les sorties des neurones de cette couche sont alors communiquées aux neurones de la couche suivante et ainsi de suite. La dernière couche du réseau est appelée couche de sortie, les autres étant désignées sous le terme de couches cachées car les valeurs de sortie de leurs neurones ne sont pas accessibles de l'extérieur [Go00].

La sortie du réseau aura une expression de la forme suivante :

$$y = f(\sum Z(f(\sum wx + w_0)) + Z_0) \quad (2.18)$$

La fonction d'activation des neurones doit absolument être non linéaire, car sans elle, le perceptron multicouche ne ferait qu'implanter une série de transformations linéaires consécutives, qui pourraient dès lors se réduire à une seule. L'algorithme d'apprentissage des perceptrons multicouches, connu sous le nom d'algorithme de rétro propagation, nécessite toute fois que les fonctions d'activations des neurones soient continues et dérivables.

### 6.5.2 La règle d'apprentissage

L'apprentissage du perceptron multicouche est supervisé et consiste à adopter les poids des neurones de manière à ce que le réseau soit capable de réaliser une transformation donnée, présentée par un ensemble d'exemples constitué d'une suite de  $N$  vecteurs d'entrées

$$x_k = [x_{k1} \ x_{k2} \ \dots \ x_{kd}]^T \text{ associés à une autre suite de vecteurs de sorties désirées } T^k = [t_1^k \ t_2^k \ \dots \ t_{hl}^k]^T .$$

#### Formalisation :

Nous donnons ici une formalisation assez sommaire de cet algorithme. Notons :

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

$x_k$  : stimulus n°  $k$  (colonne n°  $k$  de la matrice  $x$ ).

$T$  : matrice cible ou réponse théorique de dimensions  $(1, J)$ .

$t_k$  : cible n°  $k$  (colonne 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  $J$  composantes.

$w$  : matrice de poids des connexions liant la couche d'entrée à la couche cachée, de dimension  $(L, N)$ .

$w_0$  : 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 :

- 1- passage du stimulus de l'entrée vers la sortie du réseau ;
- 2- la rétro propagation de l'erreur dans le réseau pour la mise à jour des matrices des poids et de biais afin de réduire l'erreur entre la sortie obtenue et la cible recherchée.

Le principe de passage de stimulus à travers le réseau est le suivant :

Chaque cellule de la couche cachée somme toutes ces entrées et fait passer le signal obtenu à travers sa fonction de transfert. De même pour les cellules de la couche de sortie. Le vecteur de sortie est comparé à celui désiré  $t_k$ .

Sous forme matricielle, le vecteur de sortie n°  $k$  à pour expression :

$$O_k = g[Z(f(wx_k + w_0)) + Z_0] \quad (2.19)$$

Où  $f$  et  $g$  sont les fonctions de transfert (activation), respectivement des couches cachées et de sortie.



Le vecteur « erreur » s'écrit :

$$e_k = t_k - o_k \quad (2.20)$$

### Rétro propagation de l'erreur :

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

$$E_k = e_k^t e_k = \frac{1}{2} (t_k^t t_k + o_k^t o_k - 2o_k^t t_k) \quad (2.21)$$

$$\nabla E_k / Z = \frac{1}{2} [\nabla (o_k^t o_k - 2o_k^t t_k)] / Z \quad (2.22)$$

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

$$o_k = f[Z h_k + Z_0] \quad (2.23)$$

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

$$\nabla E_k / Z = \frac{\partial E_k}{Z} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial (Z h_k + Z_0)} \frac{\partial (Z h_k + Z_0)}{\partial Z} \quad (2.24)$$

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

$$\frac{\partial E_k}{\partial o_k} = o_k - t_k \quad (2.25)$$

La deuxième dérivée partielle dépend du type de la fonction de transfert utilisée.

Dans le cas du sigmoïde unipolaire, la deuxième dérivée partielle a pour expression :

$$\frac{\partial o_k}{\partial (Z h_k + Z_0)} = f(Z h_k + Z_0) \cdot [1 - f(Z h_k + Z_0)] = o_k \cdot (1 - o_k) \quad (2.26)$$

où 1 représente un vecteur unitaire de même taille que le vecteur de sortie  $o_k$  et l'opérateur « $\cdot$ » définit le produit terme à terme ou produit de hadamard.

La troisième dérivée partielle est simple à calculer :

$$\frac{\partial (Z h_k + Z_0)}{\partial Z} = h_k^T \quad (2.27)$$

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

$$Z(k+1) = Z(k) - \eta \nabla E_k / Z = Z(k) + \eta (t_k - o_k) \cdot o_k \cdot (1 - o_k) h_k^T = Z(k) + \eta \partial_s h_k^T \quad (2.28)$$

avec

$$\partial_s = (t_k - o_k) \cdot o_k \cdot (1 - o_k) \quad (2.29)$$

$\eta$ : 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'entrée passant par les biais est unitaire, on obtient facilement :

$$Z_0(k+1) = Z_0(k) - \eta \nabla E_k / Z_0 = Z_0(k) + \eta (t_k - o_k) \cdot * o_k \cdot * (1 - o_k) = Z_0(k) + \eta \partial_s \quad (2.30)$$

L'erreur de sortie  $\partial_s$  est rétro propagée à la sortie de la couche cachée à travers la transposée de la matrice  $Z$ .

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

$$w(k+1) = w(k) - \eta \nabla E_k / w = w(k) + \eta Z(k)^T \partial_s \cdot * h_k \cdot * (1 - h_k) x_k^t = w(k) + \eta \partial_h x^T \quad (2.31)$$

$$w_0(k+1) = w_0(k) - \eta \nabla E_k / w_0 = w_0(k) + \eta Z(k)^T \partial_s \cdot * h_k \cdot * (1 - h_k) = w_0(k) + \partial_h \quad (2.32)$$

où

$$\partial_h = \eta Z(k)^T \partial_s \cdot * h_k \cdot * (1 - h_k) \quad (2.33)$$

### 6.5.3 Limites de l'algorithme de la rétro propagation :

- Difficulté d'évaluation des paramètres initiaux d'apprentissage.
- Initialisation aléatoire.
- Apprentissage lent, convergence incertaine, peut être piège dans un minimum local.
- Difficulté de dimensionnement du réseau, difficulté de choix du nombre de couches cachées et de neurones dans ces couches.
- Pas d'interprétation possible des poids de connexions.

### 6.5.4 Propriété du perceptron multicouche

Cybenko [Ha94][Ke89] est le premier à avoir démontré qu'un perceptron multicouche à "d" entrées et " $h_L$ " sorties, comportant une seule couche cachée composée de neurones à fonction d'activation continue non linéaire, est suffisant pour approximer, au sens des moindres carrés, avec une erreur arbitraire faible pour un ensemble donné d'objets d'apprentissage, n'importe quelle transformation continue représentée par un ensemble de vecteurs d'entrée  $x_k = [x_{k1} \ x_{k2} \ \dots \ x_{kd}]^T$  et un ensemble de vecteurs de sortie désirés  $T^k = [t_1^{(k)} \ t_2^{(k)} \ \dots \ t_{hL}^{(k)}]^T$ .

Le théorème d'approximation prouve qu'un perceptron multicouche à une seule couche cachée est en théorie toujours suffisant. Toutefois, il ne préjuge en aucun cas du nombre

d'unités cachées nécessaire pour atteindre une qualité d'approximation suffisante ; ce nombre peut être gigantesque.

Cependant, l'utilisation d'un perceptron multicouche pourrait parfois s'avérer plus avantageuse. Pour le problème de classification, il a été démontré [BW89] que si un perceptron multicouche à une seule couche cachée, composé de neurones à fonction d'activation continue non linéaire, est enchaîné sous les conditions suivantes :

1. Le réseau comporte un neurone de sortie par classe et il est enchaîné à produire 1 sur la sortie associée à la classe de l'objet présenté et 0 sur les autres sorties.
2. Le critère d'optimisation est celui des moindres carrés de l'erreur.
3. Le nombre d'unités est suffisamment grand.
4. L'apprentissage ne converge pas vers un minimum local.

Alors, les sorties du perceptron multicouche peuvent être interprétées comme de bonnes approximations [Go00].

Seulement, cette qualité n'est pas garantie et dépend de l'apprentissage du réseau. D'autre part, la présence d'une non linéarité dans la couche de sortie est à présent essentielle. Son rôle ici est de simuler une décision logique binaire, ce qui permet de minimiser le taux d'erreur de la classification. Le choix de la non linéarité se restreint par conséquent aux fonctions sigmoïdes [Go00].

### 6.5.5 Evolution du perceptron multicouche

Les réseaux de neurones multicouches enchaînés par l'algorithme de rétro propagation du gradient sont les modèles connexionnistes les plus étudiés et utilisés à ce jour. Les champs d'application de ces réseaux sont très vastes : classification [MP69][Co91][Bu91], identification de processus [Lo92][Bu91] et prédiction de séries temporelles [NP91], commande de processus [CG93][MH91] et de robot [Bu91], traitement d'images [Co91][TW88] et de paroles [LB89][Bo92]. Ce sont des modèles robustes [LW90] et dont les entrées et les sorties peuvent indépendamment être choisies binaires ou réelles, ce qui permet de traiter de très vastes classes de problèmes. Cependant, lors de la réalisation d'une application basée sur un réseau de neurones multicouche enchaîné par l'algorithme de rétro propagation, il faut prendre en compte les points suivants [Ou98] :

- La mise en œuvre de cet algorithme exige souvent des temps de calcul très longs qui rendent son application incommode pour de nombreux problèmes de tailles raisonnables.
- La rétro propagation peut trouver piège dans les minima locaux et ne peut pas donner la bonne réponse. Pour remédier à ce problème, on peut entraîner le réseau à partir de plusieurs choix

initiaux de poids pour ne garder que le meilleur ou bien ajouter un bruit aléatoire puis relancer l'apprentissage.

- Le réseau apprend à partir des exemples de la base d'apprentissage. Si cette dernière concerne un fonctionnement dans un domaine réduit, le réseau ne saura répondre en dehors de ce domaine. Si le superviseur donne des informations erronées, le réseau apprendra un modèle erroné et si le processus que l'on cherche à modéliser est non stationnaire ou change brutalement, le réseau réapprendra un bon modèle.

Aujourd'hui, les perceptrons multicouches sont les réseaux utilisés par les développeurs d'applications. Ces résultats théoriques sur le mécanisme de comportement de ces réseaux sont encore très pauvres, mais des résultats satisfaisants ont été mis en valeur dans des domaines d'applications très divers. C'est la raison pour laquelle de nombreux chercheurs étudient ce modèle et tentent de mieux le comprendre et de l'améliorer.

## 7. Mise en œuvre des réseaux de neurones

Les réseaux de neurones réalisent des fonctions non linéaires paramétrées. Ils peuvent approcher toute fonction de régression si l'on trouve un réseau de neurone qui rend minimum la fonction de coût des moindres carrés (Somme des carrés des différences entre les valeurs calculées par le réseau de neurones et les valeurs mesurées). La mise en œuvre d'un réseau de neurone nécessite donc [Jo94] :

- De déterminer les entrées pertinentes, c'est-à-dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
- De choisir l'architecture du réseau, c'est-à-dire les entrées externes, le nombre de neurones cachés et l'agencement des neurones entre eux, de telle manière que le réseau soit en mesure de reproduire ce qui est déterministe dans les données. Notons que le nombre de poids ajustables est un des facteurs fondamentaux de la réussite d'une application ; si le réseau possède un trop grand nombre de poids, c'est-à-dire si le réseau est très « souple », il risque de s'ajuster au bruit qui est présenté dans les données de l'ensemble d'apprentissage, et même en absence de bruit, il risque de présenter des oscillations non significatives entre les points d'apprentissage, donc de présenter de mauvaises propriétés de généralisation. Si ce nombre est trop petit, le réseau est trop « rigide » et ne peut conduire que la partie déterministe de la fonction. Le problème de la détermination de l'architecture optimale est restée pendant longtemps un problème ouvert, mais il existe actuellement diverses méthodes, mettant notamment en jeu des tests statistiques qui permettent :
- De déterminer cette architecture pour une vaste classe de réseaux

- De calculer les poids du réseau ou, en d'autres termes, d'estimer les paramètres de la régression non linéaire, à partir des exemples, en minimisant l'erreur d'approximation, de telle manière que le réseau réalise la tâche désirée.
- D'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage.

## 8. Domaine d'application des réseaux de neurones

Les réseaux de neurones artificiels ne sont plus de simples outils de filtrage réservés à des spécialistes de la programmation, mais deviennent des outils courant dans un grand nombre de domaines [Ja01]. Beaucoup d'applications ont vu le jour par exemple :

Dans le domaine financier :

- Comportement humain vis-à-vis du remboursement de crédit (analyse de risque, crédit scoring). Dans ce cas, on cherche à partir d'une description du client, à créer un modèle qui prédit la probabilité de "casse" du prêt.
- Comportement d'achat (ciblage de marketing, fidélisation, vente croisée), il s'agit également ici d'associer un comportement à un dossier de client (description du client et historique de la consommation). Le modèle construit permettra d'obtenir des prédictions sur le comportement de nouveaux prospects jamais rencontrés auparavant.
- Comportements de systèmes financiers (prédiction de cours, prise de décision d'achat ou de vente), ici les cours passés servent de base à une prédiction des cours futurs, par l'intermédiaire d'un réseau de neurones construit sur les premiers et appliqué aux derniers.

Dans le domaine aérospatial :

- Développement de pilotes automatiques performants.
- Simulation de trajectoires de vols.

Dans le domaine automobile :

- Systèmes de guidage automatique de véhicule.

Dans le domaine de la défense :

- Poursuite de cibles.
- Discrimination d'objets.
- Reconnaissance des traits du visage.
- Extraction de caractéristiques, suppression de bruit, compression de données dans des images radars et sonores.

Dans le domaine du traitement du signal et de la classification:

- Reconnaissance de caractères manuscrits.

- Reconnaissance de la parole.
- Classification et extraction de signaux d'électroencéphalogrammes (EEG).
- Classification des cellules sanguines.
- Inspection de cartes de circuits imprimés.

Dans le domaine du contrôle/commande:

- Contrôle de bras manipulateur.
- Contrôle de robots mobiles.
- Commande et régulation de réacteurs chimiques.
- Analyse et maintenance de machines.

La liste est loin d'être exhaustive, mais elle témoigne de la vitalité de ce secteur de recherche. C'est surtout ces deux dernières décennies que les applications ont vu le jour, et cela ne fait que continuer. En témoigne aussi le fait que les réseaux de neurones artificiels, de plus en plus formalisés, sont devenus une discipline d'enseignement en écoles d'ingénieurs.

## 9. Limites et avantages des réseaux de neurones

Les réseaux de neurones ont des limites et des avantages [Jo94][DN93][To92].

### 9.1 Limites

Comme pour tout outil, il est important de réaliser que les réseaux ont leurs limites d'applications. Pour cela, nous citons :

- ✓ Difficulté d'utilisation du parallélisme inhérent aux réseaux de neurones : la plupart des réseaux sont simulés sur des machines séquentielles, ce qui entraîne des temps de calcul importants.
- ✓ Le passage de la simulation à l'implantation des réseaux neurones sur hardware n'est qu'à ses débuts alors que c'est la seule façon d'exploiter réellement leur parallélisme.
- ✓ Les performances des réseaux neuronaux dépendent de la quantité et de la pertinence du prétraitement effectué lors de l'apprentissage.
- ✓ L'une des principales reproches faites aux réseaux neuronaux est leur incapacité à expliquer les résultats qu'ils fournissent.
- ✓ Ces réseaux se présentent comme des boîtes noires dont les règles de fonctionnement sont inconnues, la qualité de leurs performances ne peut être mesurée que par des méthodes statistiques, ce qui cause une certaine méfiance de la part des utilisateurs.

## 9.2 Avantages

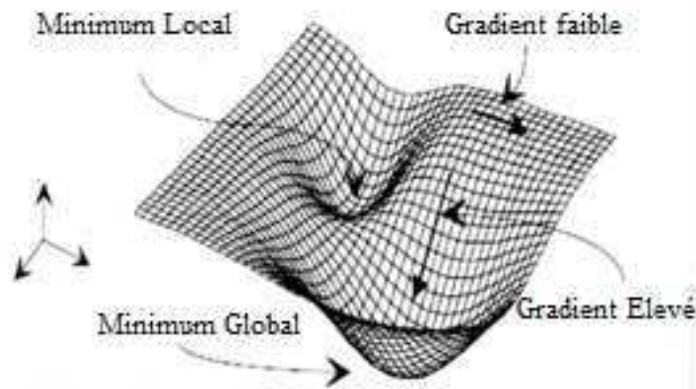
- ❖ Ils ont la capacité de représenter n'importe quelle dépendance fonctionnelle. Le réseau découvre (apprend, modélise) la dépendance lui-même sans avoir besoin qu'on lui "souffle" quoi que ce soit. Il n'a pas besoin de lui postuler un modèle.
- ❖ On passe directement des données au prédicteur, sans intermédiaire, sans recodage, sans discrétisation, sans simplification ou interprétation sujette à caution.
- ❖ Ils résistent au bruit ou au manque de fiabilité des données.
- ❖ Ils disposent d'une grande variété possible pour la variable prédite : oui/non, valeur continue, une ou plusieurs classes parmi n, etc.
- ❖ Ils sont simples à manier, beaucoup moins de travail personnel à fournir que dans l'analyse statistique classique. Aucune compétence en maths, informatique ou statistiques requise. On n'a pas besoin de détecter les colinéarités, non plus.
- ❖ En segmentation, le réseau de neurones détermine de lui-même combien de clusters distincts renferment chaque classe.
- ❖ Leur comportement est acceptable en cas de faible quantité de données.
- ❖ Pour l'utilisateur novice, l'idée d'apprentissage est plus simple à comprendre que les complexités des statistiques multivariées.
- ❖ Le modèle fourni est continu et dérivable et se prête bien à la visualisation, à une recherche de profils-type (par ascension du gradient de probabilité prédite), au calcul de sensibilités (par dérivation partielle), etc.

## 10. Aspects pratiques

Apparus en 1985, les réseaux multicouches sont aujourd'hui les modèles les plus employés. Plusieurs couches de traitement leur permettent de réaliser des associations non linéaires entre l'entrée et la sortie. Ils sont ainsi capables de résoudre le cas du "ou exclusif" (cf Perceptron). On sait depuis les années soixante que les possibilités de traitement des réseaux multicouches sont supérieures à celle du Perceptron, cependant l'algorithme d'apprentissage manquait. Pour la couche de sortie, on peut appliquer l'apprentissage du Perceptron, mais comment modifier les poids pour les connexions qui ne sont pas en relation avec un neurone de sortie ? Le problème est ramené à l'obtention d'une estimation de la valeur désirée pour chaque neurone de la couche cachée. La rétropropagation de gradient est une solution à ce problème. Cet algorithme a été proposé indépendamment par trois équipes en 1985.

Le principe utilisé par la rétropropagation ("backpropagation" en anglais) de gradient est la minimisation d'une fonction dépendante de l'erreur. Il s'agit d'une méthode générale, largement

employée dans d'autres domaines tels que la physique. Une perception intuitive de cet algorithme consiste à considérer l'apprentissage comme la recherche sur la surface de coût de la position de coût minimal. A chaque configuration de poids correspond un coût. Le gradient est une estimation locale de la pente de la surface. La minimisation du gradient permet de parcourir cette surface orthogonalement aux courbes de niveau d'un pas fixé. Les problèmes rencontrés durant l'apprentissage résultent des zones très plates et des minima locaux [Jo94].



**Figure 2.21** : Courbe de la zone de coût minimal dans l'espace de configuration des poids du réseau (représenter ici en deux dimensions)

### 10.1 Les minima locaux

A la différence de réseaux simple comme l'Adaline, la surface d'erreur d'un PMC est complexe, et comporte généralement des minima locaux où les performances du réseau sont nettement sous-optimales. La rétropropagation étant une technique de descente de gradient, elle peut donc se trouver prisonnière d'un de ces minima, et produire des réseaux peu performants.

Ainsi [KP90] a montré que les performances de deux réseaux dont les poids synaptiques ne diffèrent que de peu peuvent posséder des performances très différentes après apprentissage. Cela s'explique par le fait que les deux choix de poids peuvent se trouver de part et d'autre d'une ligne de partage entre deux bassins de minima différents. Ce genre de phénomènes explique la pratique courante d'entraîner un même réseau à partir de plusieurs choix initiaux de poids, pour ensuite ne garder que le meilleur d'entre eux. Cette technique n'est pas la seule méthode ad hoc pour éviter les minima locaux, nous pouvons aussi :

- ❖ Modifier le pas d'apprentissage du réseau pour pousser le réseau hors des minima locaux [Mo93].
- ❖ Ajouter une dose de bruit aléatoire aux poids du réseau, puis relancer l'apprentissage, dans le but de « déloger » le réseau hors de bassin peu profond.



❖ Réduire tous les poids du réseau par une petite quantité à chaque pas d'apprentissage [FW90][ Mo93] [Pe93]. L'activation d'un neurone se sature quand ses liens possèdent de trop grands poids synaptiques. Ce problème est difficile à résoudre lors de l'apprentissage car les valeurs d'activation extrêmes se traduisent souvent au niveau de la rétropropagation par de petites corrections. Le réseau se trouve alors dans un minimum local. Réduire systématiquement tous les poids par une petite valeur à chaque correction contribue à garder petits les poids synaptiques et peut résoudre ce problème.

Ces techniques heuristiques (quelques unes parmi plusieurs) témoignent d'une faiblesse de l'apprentissage par descente du gradient et de la difficulté de trouver une solution théorique au problème.

## 11. Conclusion

Dans ce chapitre, nous avons présenté un bref historique des réseaux de neurones marqué par leur développement qui a connu un grand essor dans les 30 dernières années, ce qui a amené les chercheurs à s'éloigner de plus en plus des origines de leurs travaux, fondés sur l'observation physique des neurones en fonctionnement.

Les réseaux de neurones abordés dans ce chapitre sont des systèmes paramétriques non linéaires (PMC) qui relient des variables d'entrée à des variables de sortie. Les paramètres sont déterminés statistiquement lors d'une phase d'apprentissage à partir d'une base de données d'exemples, appelée base d'apprentissage.

Les réseaux de neurones possèdent des propriétés d'apprentissage intéressantes, de reconnaissance de formes et d'approximateurs universels.

Les inconvénients ne sont pas négligeables, citons notamment le nombre de poids à optimiser, le choix de la structure du réseau, le choix des fonctions de transfert et les performances de la règle d'apprentissage.

# ANALYSE EN COMPOSANTES PRINCIPALES

## CHAPITRE 3

1. Introduction.....	46
2. Analyse de données	46
2.1 Objectifs de l'analyse de données.....	46
2.2 Nature des données analysables.....	47
2.3 Méthodes d'analyses de données multidimensionnelles	48
3. Analyse en Composantes Principales.....	49
3.1 Objectif.....	49
3.2 Principe.....	49
3.3 Propriétés des Composantes Principales.....	50
3.4 Formulation de l'ACP.....	51
3.5 Interprétation des résultats d'une ACP.....	53
3.6 Avantages et inconvénients de l'ACP.....	55
3.7 Domaines d'application de l'ACP.....	56
4. Classification des visages par ACP.....	57
4.1 Approche pixel.....	57
4.2 Approche géométrique.....	58
5. Identification d'un visage inconnu.....	58
6. Conclusion.....	59

## 1. Introduction

Dans les trente dernières années, la révolution apportée par l'informatique a (entre autres) rendu nécessaire le développement d'outils adaptés aux traitements de grandes masses d'informations. Il a ainsi fallu développer des outils permettant une appréhension rapide de l'information contenue de « grand tableaux de données ». C'est l'objectif de l'Analyse de Données Multidimensionnelles. Ces outils se sont développées d'abord à partir de méthodes préexistantes à cette période (analyse en composante principale, analyse des correspondances), mais dont l'utilisation restait confidentielle vu le manque de moyens de calcul, puis en utilisant pleinement la structure et la puissance de calcul de l'ordinateur.

De nos jours, l'étude de statistiques tient une place importante dans de nombreux domaines. Qu'il s'agisse de politique, d'économie, de science ou de sport, la possession d'informations peut être synonyme de connaissance d'indices et de phénomènes permettant d'une façon générale la prise de décisions adéquates. L'obtention de ces informations se fait par les moyens bien connus des sondages, études, relevés ou expériences.

Mais bien entendu, l'étude statistique ne s'arrête pas à la simple collecte des données. Encore faut-il pouvoir les analyser, c'est-à-dire les organiser, en extraire les informations les plus pertinentes, les présenter sous une forme exploitable et les interpréter, pour enfin en tirer les conclusions aboutissant à la prise de décision.

Cependant, la quantité parfois colossale de données collectées (par exemple dans le cas d'une étude des caractéristiques d'une population), ainsi que leur complexité, leur apparent désordre et leurs contextes rendent ces tâches souvent difficiles. D'où, la nécessité de méthodes rigoureuses d'analyse de données, permettant de structurer, faciliter et rendre plus efficace ces différentes opérations. L'Analyse en Composante Principale (ACP) est l'une de ces méthodes.

## 2. Analyse de données

### 2.1 Objectifs de l'analyse de données

L'analyse de données est apparue dans les années trente grâce à H. Hotelling, qui initialisait le principe de l'analyse en composantes principales et de l'analyse canonique en s'inspirant des travaux de C. Spearman et de K. Pearson datant du début du siècle.

Ces méthodes furent ensuite complétées et perfectionnées sans grand succès, car elles nécessitaient de très importants calculs. Le développement de l'informatique, dans les années

soixante, a remis cette approche de la statistique au goût du jour et la multiplication de micro-ordinateurs scientifiques a permis sa vulgarisation [Do03].

La statistique classique ne s'intéresse qu'à un nombre limité de caractères mesurés sur un ensemble restreint d'individus. Elle s'appuie sur des estimations ou des tests très restrictifs. Les hypothèses émises sont souvent très approximatives et les lois de distribution très théoriques. Dans la réalité, les individus sont le plus souvent décrits par de très nombreux paramètres, qui n'obéissent pas à des lois connues à priori. La différence essentielle de l'analyse de données, par rapport à la statistique classique, est qu'elle réalise une étude globale du couple individus caractères en s'appuyant sur des représentations graphiques beaucoup plus lisibles que de volumineux tableaux de mesures. Les ressemblances ou les dissemblances entre individus ou entre caractères, leurs corrélations positives ou négatives sont à la base de l'analyse, qui peut s'effectuer suivant plusieurs points de vue. Des méthodes de classification automatique viennent compléter les analyses graphiques, les classes obtenues pouvant à leur tour être visualisées sous forme d'arborescence ou de cluster.

Cette approche multidimensionnelle, depuis qu'elle est maîtrisée grâce à l'apport de l'informatique, a connu un très grand nombre d'applications pratiques pour l'observation et la compréhension des phénomènes complexes notamment en médecine, physique, sciences naturelles, sciences humaines, etc...

## 2.2 Nature des données analysables

En analyse de données on distingue deux types d'informations [Do03] :

- Les individus
- Les variables

Les **individus** peuvent représenter des entités comme des personnes, des entreprises, des pays, des journaux, etc. Sur ces individus nous allons réaliser des observations ou des mesures. Les individus peuvent représenter l'ensemble d'une population ou simplement une partie qui n'est pas nécessairement sensée être un échantillon représentatif.

Les **variables** sont des caractères relatifs à ces individus. Elles peuvent être de deux types :

- Quantitatives
- Qualitatives

Une variable sera **quantitative** si elle prend des valeurs sur une échelle numérique comme par exemple :

- L'âge
- La taille
- Le poids

- Le salaire
- Le nombre d'enfants
- Le chiffre d'affaires

Les opérations algébriques habituelles (addition, multiplication, calcul de moyenne) peuvent être utilisées pour combiner ou normaliser leurs valeurs numériques.

Une variable sera qualitative si elle peut prendre des modalités non numériques comme :

- Le sexe
- La profession
- Les diplômes
- Le niveau hiérarchique
- La situation de famille, etc.

Les valeurs possibles prises par une variable de type quantitatif sont appelées des modalités. Si ces variables peuvent être ordonnées même grossièrement, on parlera de variables qualitatives ordinales :

- Niveau hiérarchique, diplômes,
- Niveau de satisfaction, etc.

Sinon ceux sont des variables qualitatives nominales :

- Couleur,
- Entreprise,
- Région d'origine,
- Pays, etc,
- Le nombre d'employés, etc.

### 2.3 Méthodes d'analyses de données multidimensionnelles

Ces méthodes ont pour but de décrire des données consignées dans un tableau croisant des individus ( $m$  lignes) et des caractères ou variables ( $n$  colonnes) qui sont mesurés sur ces individus. L'étude s'effectue par réduction du nombre de caractères, non sous forme d'une simple sélection de certains d'entre eux, mais par construction de nouveaux caractères synthétiques obtenus en combinant les caractères initiaux (méthode factorielle). De plus la méthode de modélisation sera ici linéaire puisque chaque facteur synthétique sera combinaison linéaire des caractères initiaux. Les principales méthodes issues de cette démarche seront :

- L'analyse en composantes principales (ACP),
- L'analyse factorielle des correspondances (AFC),
- L'analyse canonique (AC),
- L'analyse factorielle discriminante (AFD).

- L'analyse Factorielle d'un Tableau de Distances (AFTD),
- L'analyse des Correspondances Multiples (ACM),
- L'analyse Factorielle Multiple (AFM).

Ces méthodes reposent toutes sur les mêmes notions théoriques, mais chacune produit un genre de résumé spécifique et s'applique sur un type de données précis. C'est pourquoi nous devons choisir la méthode la plus adaptée au type d'informations que nous possédons et aux phénomènes que nous voulons étudier.

### 3. Analyse En Composantes Principales

#### 3.1 Objectif

C'est la méthode la plus ancienne des méthodes d'analyse des données, sa mise au point par Hotelling date de 1933 [Ho33], et elle présente toujours un réel intérêt pratique.

Le but de l'analyse en composantes principales (ACP) est de condenser les données originales en de nouveaux groupements de façon à ce qu'ils ne présentent pas de corrélation entre eux et soient ordonnés en termes de pourcentage de variance apporté par chaque composant. Ainsi la première composante principale contient les informations relatives à la variance maximale, la deuxième contient les informations relatives à la variance suivante. Le processus est répété jusqu'à l'obtention de la nième et dernière composante principale [Do03].

En résumé les objectifs poursuivis par une ACP sont :

- La représentation graphique "optimale" des individus (lignes), minimisant les déformations du nuage des points,
- La représentation graphique des variables dans un sous-espace  $F_q$  en explicitant au "mieux" les liaisons initiales entre ces variables,
- La réduction de la dimension (compression).

Le dernier objectif permet d'utiliser l'ACP comme préalable à une autre technique préférant des variables orthogonales (régression linéaire) ou un nombre réduit d'entrées (réseaux neuronaux) [BB04] [La98].

#### 3.2 Principe

Nous avons vu que l'objectif principal de l'ACP était de réussir à exprimer un système complexe de données de dimension quelconque, représenté par un tableau dit Individus/Variables, dans un plus petit nombre de dimensions, tout en minimisant la perte d'informations engendrée.

Ce processus consiste à construire de nouveaux caractères synthétiques à partir de facteurs (d'où le terme d'analyse multifactorielle) obtenus par combinaisons linéaires des variables initiales. Ceux sont ces nouveaux caractères synthétiques qui sont appelés composantes principales.

La problématique consiste alors à trouver les  $n$  composantes principales qui "expliquent" (c'est-à-dire qui résument) la plus grande quantité d'information originale ( $n$  étant le nombre de dimensions auxquelles on veut réduire l'étude. Dans de nombreux cas, on se limite à 2. L'étude des données pourra alors être réduite à ces composantes principales, si elles sont suffisamment représentatives. Il faut donc pouvoir quantifier l'efficacité des composantes principales, pour ne retenir que les plus significatives.

Dans le cadre des représentations graphiques, les composantes principales sont assimilées aux axes. Ainsi, dans le cadre d'une représentation de données dans un espace à deux dimensions, l'axe des abscisses représentera la première composante principale, celle qui explique la plus grande partie de l'information d'origine, et l'axe des ordonnées sera la deuxième composante principale, celle qui explique la plus grande partie de l'information restante [BF01].

### Remarque

L'ACP ne peut s'appliquer que si les variables quantitatives qu'on étudie et qui ne sont pas indépendantes deux à deux et possèdent des coefficients de corrélation non nuls. On n'utilisera donc cette méthode que si on a déjà un à priori concernant la relation de certaines de ces variables.

## 3.3 Propriétés des Composantes Principales

Les composantes principales issues d'une ACP sont généralement caractérisées par:

### 3.3.1. Nombre

Bien que l'objectif soit en général de n'utiliser qu'un petit nombre de Composantes Principales, l'ACP en construit initialement  $p$ , autant que de variables originales. Ce n'est que par la suite que l'analyste décidera du nombre de Composantes à retenir.

### 3.3.2. Orthogonalité

Les Composantes Principales définissent des directions de l'espace des observations qui sont deux à deux orthogonales. Autrement dit, l'ACP procède à un changement de repère orthogonal, les directions originales étant remplacées par les Composantes Principales.

### 3.3.3. Décorrélation

Les Composantes Principales sont des variables qui s'avèrent être deux à deux décorrélées.

### 3.3.4. Ordre et sous-espaces optimaux

La propriété fondamentale des Composantes Principales est de pouvoir être classée par ordre décroissant d'importance dans le sens suivant :

- Si l'analyste décide de décrire ses données avec seulement  $k$  ( $k < p$ ) combinaisons linéaires de ses variables originales tout en perdant le moins possible d'information, alors ces  $k$  combinaisons linéaires sont justement les  $k$  premières Composantes Principales.

Ainsi, le meilleur sous-espace à  $k$  dimensions dans lequel projeter les observations est justement celui engendré par les  $k$  premières Composantes Principales.

### 3.4 Formulation de l'ACP

La mise en œuvre mathématique de l'ACP peut être divisée en 6 étapes principales.

#### Etape 1 : Préparation des données

Les données que l'on doit traiter par l'ACP sont stockées dans un tableau  $X$  de type individus/variables de la forme :

$$X = \begin{matrix} & \begin{matrix} \text{Variables} \\ 1 & \dots & p \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ n \end{matrix} \text{ individus} & \begin{bmatrix} X_1^1 & \cdots & X_p^1 \\ \vdots & \ddots & \vdots \\ X_1^n & \cdots & X_p^n \end{bmatrix} \end{matrix} \quad (3.1)$$

Nous avons alors :

- $p$  variables, représentées en colonnes,
- $n$  individus, représentés en lignes,
- des valeurs prises par chaque variable, pour chaque individu, notée :

$$X_j^i \quad (1 \leq i \leq n, 1 \leq j \leq p) \quad \text{avec} \quad \forall (i, j), X_j^i \in R \quad (3.2)$$

Le problème est que si nous analysons directement la matrice  $X$ , les résultats seraient faussés par les valeurs relatives des variables (Par exemple si les valeurs ont été mesurées dans des unités différentes). Préparer les données pour le traitement consiste donc à transformer le tableau de données de façon à réduire ces effets.

Nous pouvons alors travailler sur une version centrée-réduite  $X_{cr}$  de la matrice  $X$ . Pour centrer la matrice  $X$ , on soustrait à chacune des valeurs la moyenne de sa variable.

Pour réduire la matrice centrée, on divise chacune des valeurs par l'écart-type de sa variable.



**Etape 2 : Calcul de la matrice des coefficients de corrélations**

Dans cette étape, nous calculons la matrice de corrélations des données contenues dans le tableau  $X_{cr}$ , notée « *Corr* ».

**Etape 3 : Calcul des Valeurs et vecteurs propres**

Les valeurs et vecteurs propres de la matrice « *Corr* » sont les facteurs utilisés pour construire les composantes principales.

**Etape 4 : Classement des vecteurs propres**

La qualité des estimations auxquelles conduit l'ACP dépend, de façon évidente, du choix de  $q$ , c'est-à-dire du nombre de composantes retenues pour reconstituer les données, ou de la dimension du sous-espace de représentation. De nombreux critères de choix pour  $q$  ont été proposés dans la littérature. Nous présentons ici ceux, les plus courants :

**a- Classement des vecteurs propres dans l'ordre décroissant des valeurs propres associées :**

On dispose alors des facteurs dans l'ordre décroissant de la quantité d'information qu'ils expliquent. Il est également possible d'exprimer en pourcentage l'importance de chacun, afin de visualiser l'importance relative des composantes principales. Notons par « *U* » la matrice dont les colonnes sont les vecteurs propres de « *Corr* » classés par ordre décroissant de leurs valeurs propres associées.

**b- Part d'inertie (méthode du seuil)**

La qualité globale des représentations est mesurée par la part d'inertie expliquée :

$$r_q = \frac{\sum_{i=1}^q \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (3.3)$$

La valeur de  $q$  est choisie de sorte que cette part d'inertie expliquée  $r_q$  soit supérieure à une valeur seuil fixée a priori par l'utilisateur. C'est souvent le seul critère employé.

**c- La règle de kaiser**

On considère que, si tous les éléments de  $Y$  sont indépendants, les composantes principales sont toutes de variances égales (égales à 1 dans le cas de l'ACP réduite). On ne conserve alors que les valeurs propres supérieures à leur moyenne car seules jugées plus "informatives" que les variables initiales ; dans le cas d'une ACP réduite, ne sont donc retenues que celles plus grandes que 1.

**d- Méthode du coude**

Le principe consiste à rechercher un coude dans le graphe des valeurs propres et de ne conserver que les valeurs jusqu'à ce coude.

**Etape 5 : Calcul de la matrice des composantes principales**

La matrice appelée matrice des composantes principales, notée CP, est celle qui contient les coordonnées des individus dans l'espace formé par celles-ci. Elle est calculée de la façon suivante :

$$CP = X_{cr} \cdot U \quad (3.4)$$

**Etape 6 : Représentations graphiques**

Le but de l'ACP étant de résumer une situation donnée, la représentation graphique est la phase finale et la plus importante de ce processus, car elle permet d'avoir rapidement un aperçu de ce que le calcul numérique ne peut pas fournir.

L'ACP peut fournir deux types de représentations graphiques :

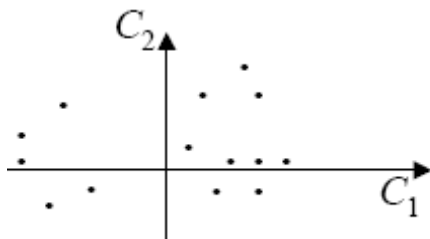
**a. Les diagrammes de dispersion**

Il s'agit d'utiliser la matrice des CP précédemment calculée pour représenter les individus sous forme de nuages de points, dans des plans ou espaces factorielles de deux ou trios composantes principales.

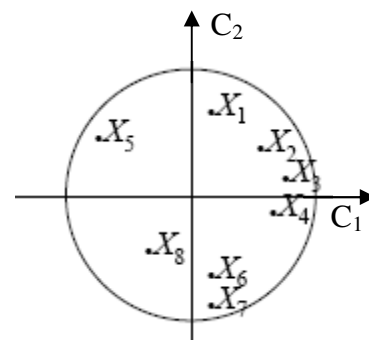
**b. Les cercles de corrélation**

Il s'agit de représenter les variables dans des systèmes d'axes factoriels à deux composantes principales, munis d'un cercle de rayon 1 pour aider à l'interprétation.

Ce type de graphique est plus utile que les diagrammes de dispersions, car il permet de visualiser facilement les relations existant entre les variables elles-mêmes, la force de ces relations et quelles composantes principales les expliquent le mieux. Il offre donc une vision plus globale de la situation, plus dans la lignée de ce qu'on veut obtenir en effectuant une ACP.



**Figure 3.1** : Diagrammes des dispersions



**Figure 3.2** : Cercle de corrélation

**3.5 Interprétation des résultats d'une ACP**

La phase de l'interprétation permet d'étudier la façon d'exploiter les résultats. Interpréter les résultats d'une ACP consiste en grande partie à donner un sens aux axes factoriels découverts, et à déterminer la nature des relations existantes entre ces axes.

L'interprétation des résultats d'une ACP peut se faire selon deux approches, la première basée sur les individus et la seconde sur les variables :

### 3.5.1 Approche basée sur les individus

Cette indication est donnée par les diagrammes de dispersion, qui expriment la contribution de chacun des individus à la variance des axes.

L'interprétation des composantes principales se fait à l'aide de la position des individus. La position est déterminé par :

$$c = x.u \quad \text{avec} \quad u = (u_1, u_2, u_3, \dots, u_q) \quad (3.5)$$

### 3.5.2 Approche basée sur les variables

Cette indication est donnée par les cercles de corrélation, qui expriment en termes de coefficients de corrélation la contribution des variables aux axes factoriels.

L'ACP construit de nouvelles variables artificielles et des représentations graphiques permettant de visualiser des relations entre variables.

Les corrélations, elles sont synthétisées dans un cercle imaginaire à deux axes ou plus. Chaque axe représente une composante principale. On regarde la position des variables par rapport à cet axe :

- Si les variables sont proches de la direction positive de l'axe alors elles sont corrélées.
- Si les variables sont proches de la direction négative de l'axe alors elles sont anti-corrélées.
- Si les variables sont proches d'un autre axe, alors elles sont non corrélées.
- Pour comparer des points entre eux on ne le fera que s'ils sont proches de la circonférence du cercle.
- Par contre, on ne peut rien dire quand les variables sont agglomérées au centre du cercle, ou de la sphère unité.
- Les corrélations avec la métrique M choisie, s'observent dans un espace sphérique.

Toutes les corrélations calculées, on les représente dans un plan formé par un couple de composantes principales  $C_1$  et  $C_2$  où elles figureront dans le cercle de corrélation (Figure 3.2).

### 3.5.3 Règles d'interprétation

Pour interpréter les résultats de l'ACP, on utilise à la fois des méthodes empiriques et des méthodes numériques.

#### a- Les méthodes empiriques

Ceux sont des règles basées seulement sur l'observation visuelle du graphique, parmi ces règles empiriques, on peut citer les deux règles principales suivantes :

- Plus la coordonnée d'une variable sur l'axe est proche de 1 ou -1, plus la composante principale explique cette variable.
- Plus deux variables sont proches sur le graphe, plus leur lien est fort.

L'avantage de cette méthode est fournit sans manipulation mathématique supplémentaire des interprétations rapides et relativement fiable de l'ACP.

Cependant, elle comporte également des contraintes, elles sont difficilement applicables lorsqu'on possède de très nombreuses variables et bien entendu s'avèrent insuffisantes si on a besoin de conclusion précise. C'est pourquoi on leur associe des méthodes numériques, principalement pour l'étude des individus qui en ont plus de variables.

### b- Les indicateurs numériques

Ceux sont des indicateurs mathématiques qui expriment des caractéristiques particulières du système axe/individus/variables qu'on a déterminé par l'ACP.

Parmi les indicateurs numériques les plus utilisés, on peut citer :

- La qualité globale de représentation : est le pourcentage de l'inertie pris en compte par les sous espaces de deux dimensions.

$$\left[ \frac{1}{\sum_{i=1}^p \lambda_i} (\lambda_1 + \lambda_2 + \dots + \lambda_q) \right] * 100 \quad (3.6)$$

- La contribution de chaque individu à chaque axe ou par rapport à tous le nuage permettent d'identifier les individus très influents qui peuvent déterminer l'orientation de certains axes.

$$CTR(i, \alpha) = \frac{1}{n} * \frac{(c_i^\alpha)^2}{\lambda_\alpha} * 100 \quad (3.7)$$

Avec : i: Indice de l'individu,  $\alpha$  : Indice de la composante principale.

## 3.6 Avantages et inconvénients de l'ACP

### 3.6.1 Avantages

- **Simplicité mathématique :**

L'ACP est une méthode factorielle car la réduction du nombre des caractères ne se fait pas par une simple sélection de certains d'entre eux, mais par la construction de nouveaux caractères synthétiques obtenus en combinant les caractères initiaux au moyen des "facteurs". Cependant, il s'agit seulement de combinaisons linéaires. Les seuls véritables outils mathématiques utilisés dans l'ACP sont le calcul des valeurs et vecteurs propres d'une matrice et les changements de base.

Sur le plan mathématique, l'ACP est donc une méthode simple à mettre en œuvre.

- **Simplicité des résultats :**

Grâce aux graphiques qu'elle fournit, l'Analyse en Composantes Principales permet d'appréhender une grande partie de ses résultats d'un simple coup d'œil.

- **Puissance:**

L'ACP offre, en quelques opérations seulement, un résumé et une vue complète des relations existantes entre les variables quantitatives d'une population d'étude, résultats qui n'auraient pas pu être obtenus autrement, ou bien uniquement au prix de manipulations fastidieuses.

- **Flexibilité :**

L'ACP est une méthode très souple, puisqu'elle s'applique sur un ensemble de données de contenus et de tailles quelconques, pour peu qu'il s'agisse de données quantitatives organisées sous forme individus/variables. Cette souplesse d'utilisation se traduit surtout par la diversité des applications de l'ACP, qui touche tous les domaines, comme exposé dans le paragraphe 3.7.

### 3.6.2 Inconvénients

En tant que méthode d'analyse de données, l'ACP n'a pas réellement d'inconvénients en soi. Elle s'applique simplement sur des cas précis et pour générer un type de résultats particuliers. Ça n'aurait donc aucun sens de dire que c'est un inconvénient de l'ACP qu'elle ne s'applique pas en dehors de ce contexte.

De même, étant donné qu'il s'agit avant tout d'une technique de résumé de données, la perte d'information forcément engendrée n'est pas un inconvénient, mais plutôt une condition d'obtention du résultat, même si elle occulte parfois des caractéristiques pourtant représentatives dans certains cas particuliers.

On peut, donc, parler tout au plus des précautions d'utilisation. La règle est, au même titre que toutes les méthodes d'aide à la décision, l'ACP n'est qu'un outil mathématique mis à disposition de l'analyste de données pour résumer la réalité. Ses résultats ne doivent, donc, pas être pris en compte de manière brute, mais nécessitent d'être interprétés et critiqués.

### 3.7 Domaines d'application de l'ACP

Dans quasiment tous les domaines où l'étude statistique est utilisée, on est amené à faire des traitements sur des variables quantitatives et à interpréter des relations entre elles, afin de prendre des décisions.

C'est pourquoi, l'ACP est l'une des méthodes d'analyse de données les plus employées.

Elle est appliquée dans des domaines aussi variés que :

- L'étude des caractéristiques d'une population, quelles qu'elles soient (sondages politiques, études de santé, etc...),

- Les indicateurs économiques (études de marchés, rentabilité d'investissements, etc...),
- L'assurance qualité et la fiabilité de produits,
- L'étude de résultats sportifs,
- L'étude des performances d'un réseau (de télécommunication, de distribution, etc...),
- L'exploitation d'expériences et de relevés scientifiques (en physique, chimie, géologie, archéologie, etc...),
- Le traitement de l'information (reconstitution de photos satellites, reconnaissance vocale, reconnaissance de visages... etc...),
- Et de nombreux autres domaines...

#### 4. Classification des visages par ACP

Pour appliquer la méthode de classification précédente, nous avons filmé des séquences vidéo pour quelques personnes. Pour chaque personne, nous avons prélevé huit séquences vidéo de dix images ; les images de chaque séquence sont extraites d'un court enregistrement vidéo.

- Ce qui différencie les êtres humains les un par rapport aux autres ; c'est évidemment leur forme est plus exactement la forme de leurs visages et leurs traits.
- Sous un contexte plutôt scientifique, nous pouvons dire que chaque visage possède une géométrie propre à lui, il peut donc être identifié soit :
- Par les textures propres de chaque élément du visage dans ce cas nous devons prendre l'image entière pour la classification (Approche pixel).
- Par les distances calculées entre les points caractéristiques des visages (Approche géométrique).
- Du fait de la mobilité de la tête, les positions des points et les distances caractérisant le visage varient lors des différentes prises de vues.

##### 4.1 Approche pixel

Cette approche consiste à présenter l'image d'un visage en terme d'un système optimal de coordonnées. Ceci est équivalent à trouver la composante principale de la distribution des faces, ou les vecteurs propres de la matrice de covariance de l'ensemble des faces. L'image est traitée comme une collection de traits qui, ensembles, caractérisent les variations dans les images des visages. Chaque visage contribue à leur définition et apparaît comme un semblant de visage, c'est pour cela qu'on l'appelle « visage propre » ou « eigenfaces » [MP93][TP90]. Chaque image de visage peut être représentée en termes de combinaisons linéaires des visages propres, ou approximée en tenant compte seulement des vecteurs (visages) propres les plus significatifs, c'est à dire ceux associés aux plus grandes valeurs propres. L'idée d'utiliser les vecteurs propres pour la reconnaissance a été

motivée par une technique développée par Kirby et Sirovich en 1987 qui consiste en la représentation de l'image de visage en utilisant l'ACP. Une phase d'initialisation constituée des étapes suivantes est nécessaire [TB00] :

1. Acquisition d'un ensemble initiale de  $M$  visages (ensemble d'apprentissage).
2. Calcul des visages propres de cet ensemble, en gardant juste les  $M'$  ( $M' < M$ ) composantes associées aux plus grandes valeurs propres. Les  $M'$  visages propres définissent l'espace des visages.
3. Calcul de la distribution correspondante des poids de chaque individu en projetant ces images sur l'espace des visages. Une mise à jour de la base des visages peut être effectuée dès la présentation d'un visage. Les étapes utilisées pour reconnaître de nouveau visage sont :
  - Projeter l'image du visage à identifier sur l'espace des visages,
  - Déterminer si l'image représente un visage ou non,
  - S'il s'agit d'un visage, classer ses poids de manière à définir s'il est connu ou pas.

#### 4.2 Approche géométrique

Chaque visage est identifié par des distances calculées entre des points caractéristiques du visage. Du fait de la mobilité de la tête, les positions des points ainsi que les distances caractérisant le visage, varient lors des différentes prises de vues. Afin de prendre en compte cette variation, un visage est décrit par des données de type intervalle. On appelle de l'objet « visage » ou chaque objet est décrit par des données complexes (intervalle, ensemble,...) [CT03].

Un outil de description des objets complexes et symboliques a été proposé par Diday [De93][DK91]. Ce formalisme permet d'une part, de décrire des objets par des variables multivaluées (à valeurs intervalles, ensemblistes, probabilistes,...), et d'autre part de prendre en compte les différents liens et contraintes liant les objets ou les parties composant l'objet [LC96].

### 5. Identification d'un visage inconnu

Une fois la procédure de classification est implémentée, qui est la dernière étape de notre système de vision, vient le rôle de l'identification pour un visage particulier inconnu et de voir son degré d'appartenance à chacune des classes de références.

L'identification d'un visage inconnu consiste alors à projeter les caractéristiques mesurées sur ce visage dans la base des composantes principales et à utiliser des mesures de similarité à fin de déterminer le taux de reconnaissance de ce visage (son degré d'appartenance à chacune des classes) en utilisant la matrice des coefficients de corrélations [Bo05].

## 6. Conclusion

Nous avons consacré ce chapitre à l'ACP (Analyse en Composante Principale) : Un grand classique des méthodes d'analyse multidimensionnelles. Le but principal de cette méthode est d'analyser des données en cherchant les axes principaux d'un nuage de points et d'analyser la représentation de ce dernier dans le repère formé par les premiers de ces axes. Ceci se fait selon les modalités suivantes :

- Fournir des outils simples et lisibles de représentation des informations traitées, permettant de faire ressortir des données brutes les éventuels liens existant entre les variables.
- Donner des indications sur la nature, la force et la pertinence de ces liens afin de faciliter leur interprétation et découvrir quelles sont les tendances dominantes de l'ensemble de données.
- Réduire efficacement le nombre de dimensions étudiées (et ainsi simplifier l'analyse), en cherchant à exprimer le plus fidèlement possible l'ensemble original de données grâce aux relations détectées entre les variables.

L'ACP est donc une bonne méthode pour déceler des liens entre les variables dans un ensemble de données multidimensionnelles afin de comprendre leurs interactions et faciliter leur étude. Pour ces raisons, nous allons exploiter cette méthode à la classification des visages (chapitre 4), cette dernière nous permet d'obtenir une représentation discriminante des visages tout en tenant compte de la mobilité de certains éléments telle que la bouche.

Les résultats de L'ACP seront exploités dans deux nouvelles approches (conçues par nos soins) que nous proposons dans le chapitre 4 :

- La première c'est une approche de décision qui permet de déterminer aisément le degré d'appartenance d'un visage à une classe en utilisant la matrice de corrélation.
- Alors que la deuxième est une nouvelle méthode née de l'hybridation des réseaux de neurones avec l'ACP que nous avons appelé Neuro-ACP, elle nous a permis d'augmenter les performances des deux méthodes de classification séparées.



# EXTRACTION DES PRIMITIVES ET CLASSIFICATION DES VISAGES

---

## CHAPITRE 4

1. <i>Introduction</i> .....	61
2. <i>Extraction des primitives</i> .....	61
2.1 <i>Localisation des éléments caractéristiques du visage</i> .....	61
2.2 <i>Extraction des paramètres du visage</i> .....	62
3. <i>Base de données</i> .....	64
4. <i>Classification des visages</i> .....	64
4.1 <i>Résultat de l'approche pixel</i> .....	65
4.2 <i>Résultats de l'approche géométrique</i> .....	69
4.3 <i>Résultats de la méthode d'analyse en composantes         principales (ACP)</i> .....	70
4.4 <i>Classification par méthode hybride Neuro-ACP</i> .....	78
5. <i>Conclusion</i> .....	80

## 1. Introduction

La classification a longtemps été un problème dont la solution était toujours relative et non définitive. Ceci est du d'une part à la complexité des objets à classer, car si autrefois la classification était limitée à des objets de formes simples les tendances aujourd'hui sont vers des objets de plus en plus complexes tels que les visages, et d'autre part et surtout à la méthode de classification adoptée parmi les innombrables méthodes existantes qui varient entre les classificateurs paramétriques : Euclidien, Quadratique, Gaussien, non paramétriques : K-means (plus proche voisin) et Estimation des probabilités à posteriori, connexionnistes : réseaux de neurones artificiels, et les méthodes d'Analyse de données : ACP, AFC, ACM,..etc. Même si toutes ces méthodes donnaient des résultats satisfaisants une autre catégorie de méthodes à vu le jour c'est le résultat de l'hybridation de deux méthodes de classifications dans le but de rassembler les performances de deux méthodes en une. Dans le cadre de ce travail, nous présentons une étude comparative de l'application des RNA seuls [BB03] puis L'ACP seule [BB05] et enfin l'hybridation des deux : méthode Neuro-ACP [BB06].

## 2. Extraction des primitives

L'extraction des primitives consiste à sélectionner les plus significatives afin de réduire la quantité d'informations à manipuler selon l'objectif du système.

### 2.1. Localisation des éléments caractéristiques du visage

L'acquisition des images se fait sous un éclairage fixe, et la prise de vue est frontale. Le processus de localisation des éléments du visage est fondé sur l'utilisation de cette particularité d'acquisition, il s'agit en effet de structures peu mobiles pouvant subir des déformations et qui présentent un contraste important par rapport à leur environnement spatial. La première tâche consiste à localiser les yeux. Le nez et la bouche seront ensuite recherchés en utilisant la géométrie du visage [Be00] [Bo05].

#### 2.1.1 Localisation des yeux

L'existence d'une tâche de réflexion spéculaire due aux sources lumineuses frontales sur les iris, nous permet de localiser l'œil en utilisant l'intensité lumineuse de chaque pixel.

En faisant un balayage de l'image de haut en bas, nous pouvons faire ressortir les positions locaux de la fonction d'intensité de niveau de gris, ce qui permet de localiser les positions des points (pixels) possédant un maximum d'énergie, c'est-à-dire dont l'intensité est maximale. D'où, la détection des positions des iris.

### 2.1.2 Localisation du nez et de la bouche

Pour localiser le nez et la bouche, nous utilisons ensuite l'information spatio-temporelle de façon à caractériser les points ayant enregistré la plus forte variation d'intensité au cours de la séquence. L'image  $I_{som}$  ainsi obtenue vérifie :

$$I_{som} = \sum_{i=0}^{i=n} \sqrt{\left(\left(\frac{\partial I(x, y, t)}{\partial x}\right)^2 + \left(\frac{\partial I(x, y, t)}{\partial y}\right)^2 + \mu \left(\frac{\partial I(x, y, t)}{\partial t}\right)^2\right)} \quad (4.1)$$

Avec :

$I(x, y, t)$  : L'intensité du niveau de gris,

$x, y$  : Les composantes spatiales,

$t$  : La composante temporelle,

$\mu$  : Un terme de normalisation dépendant de l'échantillonnage temporel de la séquence et du mouvement du visage,

$n$  : Le nombre des images acquises.

La localisation du nez et de la bouche consiste à chercher le long de la médiatrice passant par les iris, les deux régions ayant les plus fortes valeurs sur l'image de  $I_{som}$ .

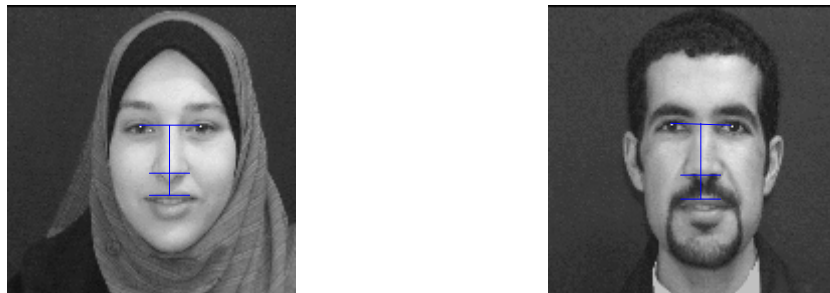
Quelques exemples de résultats de localisation du nez et de la bouche obtenus sur notre base de données sont illustrés dans la figure 4.1.



**Figure 4.1** : Exemples de résultats de localisation des yeux, du nez et de la bouche

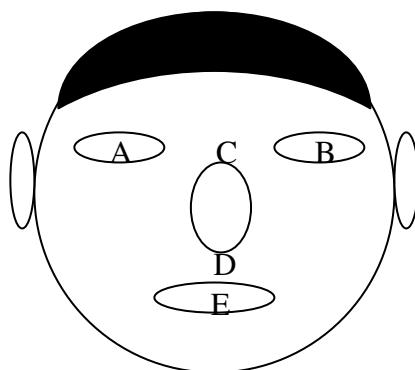
## 2.2. Extraction des paramètres du visage

Après avoir détecté les éléments caractéristiques de chaque visage, nous allons extraire les paramètres permettant de faire la distinction entre une personne et une autre. Ces paramètres sont représentés par les distances entre les éléments caractéristiques du visage.



**Figure 4.2 :** Exemple de résultats de l'extraction des éléments caractéristiques de visages

Les distances calculées sont : AB, AC, CD, DE, DB, DA, EB, EA où les points A, B, C, D, E représentent respectivement l'iris droit, l'iris gauche, le point milieu entre les deux iris, le nez et la bouche (figure 4.3). Une illustration des résultats obtenus de l'extraction de ces distances est présentée dans la figure 4.4



**Figure 4.3:** Schéma représentatif des points caractérisant le visage



**Figure 4.4 :** Exemple de résultats de l'extraction des distances de visages

Le tableau suivant regroupe les distances correspondantes à 6 personnes de notre base de données (3 hommes, 3 femmes) [Bo05].

Distance	AB	AC	CD	DE	DB	DA	EB	EA
<b>Halim</b>	30.0000	15.0000	27.0000	9.0000	30.8869	30.8869	39.0000	39.0000
<b>aidoud</b>	29.0000	14.5000	23.0000	9.0000	27.1892	27.1892	35.1319	35.1319
<b>fayçal</b>	31.0000	15.5000	29.0000	7.0000	32.8824	32.8824	39.1950	39.1950
<b>Nedjela</b>	31.0000	15.5000	28.0000	9.0000	32.0039	32.0039	40.1155	40.1155
<b>nabila</b>	30.0000	15.0000	26.0000	10.0000	30.0167	30.0167	39.0000	39.0000
<b>samiha</b>	29.0000	14.5000	35.0000	8.0000	37.8847	37.8847	45.3790	45.3790

**Tableaux 4.1:** Exemples de distances de 6 personnes

### 3. Base de données

La base donnée B-LAIG [La] que nous avons utilisée a été créée au sein du LAIG (Laboratoire d'Automatique et d'Informatique de Guelma) en 2004, ou nous avons rassemblé 8000 images faciales de 100 personnes dont 50 femmes et 50 hommes, pour chaque personne nous avons pris 8 séquences vidéo qui diffèrent l'une par rapport à l'autre par l'éclairage (meilleur, moyen, faible), l'orientation du visage (nous avons pris 4 positions) et l'émotion (Sourire) [Bo05].

Pour notre application nous avons utilisé les séquences vidéo, de 40 personnes de sexes différents dont les prises de vue sont frontales avec un éclairage meilleur, pour créer deux autres bases : une base d'apprentissage (figure 4.5) et une autre pour le test (figure 4.6) qui nous permettra de tester les performances de notre système.

### 4. Classification des visages

Les travaux portant sur la reconnaissance de visages à partir de vues frontales 2D prises dans des conditions d'illumination constantes peuvent être répertoriés en deux catégories distinctes, suivant qu'ils se rapportent à une approche géométrique ou une approche pixel.

*L'approche géométrique* consiste à extraire la position relative des différents éléments du visage tels que le nez, la bouche et les yeux afin de restreindre l'espace des données initiales. Différentes méthodes statistiques, inspirées des algorithmes classiques de reconnaissance de formes, sont alors utilisées pour discriminer les visages à partir de ces mesures [Be00].

*L'approche pixel* est une approche globale qui nécessite peu de reconnaissances préalables sur la structure du visage et se base principalement sur l'information pixel. Certains travaux s'inspirent de méthodes de traitement d'image liées à la mise en correspondance de régions ; les

mesures utilisées pour identifier les visages sont alors obtenues par des calculs de corrélation entre les régions. D'autres méthodes utilisent une analyse en composantes principales réalisée à partir de l'information pixel, pour rechercher une base de représentation permettant de discriminer les visages [Tu91].

#### 4.1 Résultats de l'approche pixel

Notre application consiste à classifier un ensemble de visages de 40 personnes dont 20 femmes et 20 hommes dans le but de les reconnaître. Les informations extraites de chaque visage sont représentées par l'ensemble des pixels constituant la matrice image correspondante. Pour cela nous avons utilisé deux bases de données :

La première représente la base d'apprentissage (figure 4.5) représentant 40 personnes de sexes différents et pour chaque personne nous avons pris 5 images de prises de vues différentes. Donc, pour l'apprentissage nous avons utilisé une base constituée de 200 images.

La seconde est la base de test qui est l'ensemble des fichiers images des visages des mêmes personnes précédentes mais sur lesquels nous n'avons pas appliqué l'apprentissage (prises de vues différentes) (figure 4.6).

Nous avons appliqué le PMC (figure 4.8) pour la classification d'un ensemble de visages correspondants à quarante personnes différentes où nous avons pris 10 images pour chaque personne, ce qui donne un total de 400 images à manipuler.

##### 4.1.1 Architecture du PMC

Le perceptron multicouche utilisé est constitué de trois couches : une couche d'entrée constituée de 2808 neurones, chaque neurone reçoit un pixel de chacune des images des visages à reconnaître, une couche de sortie dont le nombre de neurones est égale au nombre de classes des visages 40 classes dans notre cas et une couche cachée de 60 neurones et dont le nombre a été fixé après plusieurs essais sur les performances et la rapidité de sa convergence vers une erreur minimale en fonction du nombre de neurones de cette couche.

##### 4.1.2 Apprentissage

Après avoir effectué plusieurs tests pour fixer les caractéristiques adéquates du réseau, nous avons procédé à la phase d'apprentissage en utilisant un nombre d'itérations considérable pour minimiser l'erreur, ce qui nous a mené au graphe de la figure 4.7 où l'erreur a atteint la valeur 0.01.



**Figure 4.5.** : Base d'apprentissage constituée de 40 personnes pour chaque personne nous avons pris 5 images dans différentes prises de vues



**Figure 4.6** : Base de test constitué des 40 personnes précédentes mais dans des prises de vues différentes de celles de la base apprentissage

### 4.1.3 Reconnaissance

Pour tester la fiabilité de notre système, nous avons effectué des tests sur la base d'apprentissage elle-même, mais avec des prises de vues différentes, cela a conduit à des résultats très satisfaisants où le taux d'apprentissage a atteint les 99 %. Ensuite, nous avons effectué les tests de reconnaissance sur la base de tests. Les résultats sont illustrés aussi dans le tableau 4.2, où nous constatons que le système de reconnaissance que nous venons de mettre en œuvre n'est pas très satisfaisant puisque le taux de reconnaissance ne dépasse pas les 85% soit 15 % des visages non reconnus ceci est dû à la sensibilité de cette approche au changement d'intensité aussi minime soit il.

	Taux de Reconnaissance (en %)	Taux de Rejet (en %)
Base d'apprentissage	99%	1%
Base de test	85%	15%

Tableau 4.2: Les différents taux de l'approche pixel

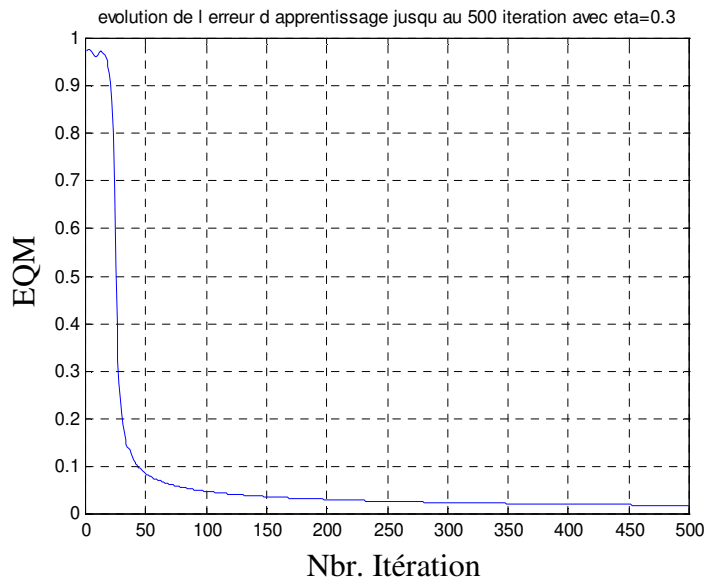


Figure 4.7 : Evolution de l'erreur en fonction du nombre d'itérations pour l'approche pixel



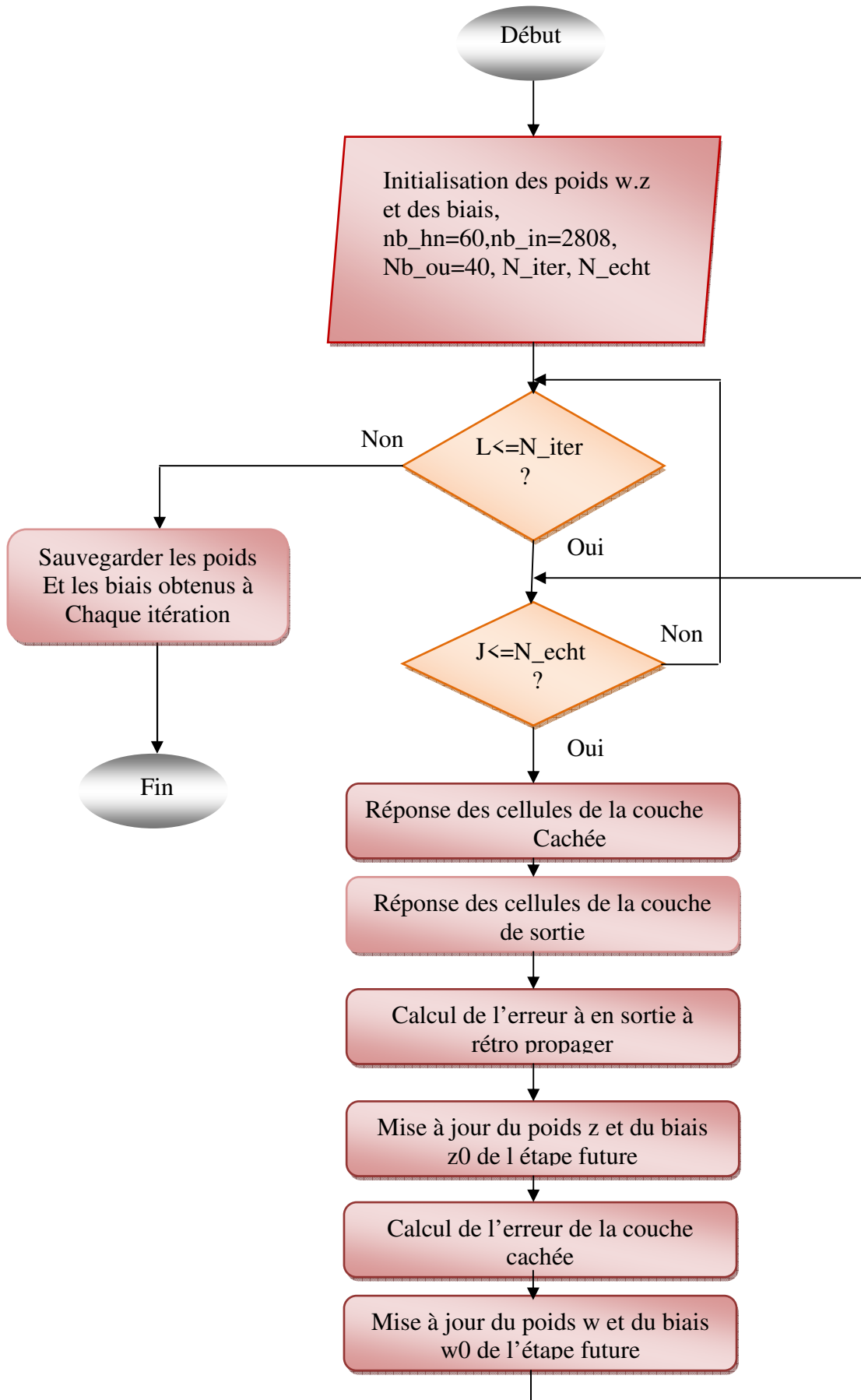


Figure 4.8 : Organigramme de l'apprentissage pour l'approche pixel

## 4.2. Résultats de l'approche géométrique

Les données à classer dans cette approche ne sont plus des images (pixels) mais se sont les distances, correspondant à l'écart entre les éléments caractéristiques des visages, extraites précédemment. Ceci nous permet un gain en quantité de données analysable (classifiable) et en temps de calcul (cf 2.2).

### 4.2.1 Architecture du PMC

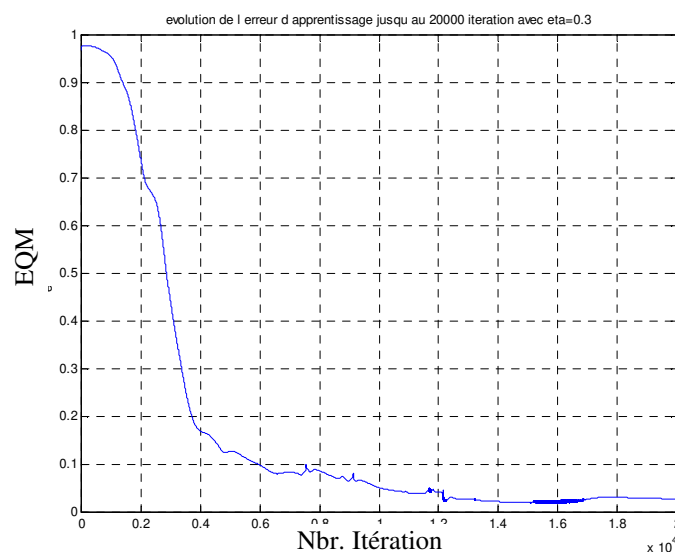
Le PMC utilisé dans cette approche est constitué d'une couche d'entrée de 8 neurones (correspondantes aux 8 distances caractéristiques de chaque visage), d'une couche de sortie de 40 neurones représentant les différentes classes que nous devons séparer et une couche intermédiaire dont le nombre de neurones a été déduit en faisant plusieurs essais pour une meilleure reconnaissance ce qui nous a mené à prendre 30 neurones dans cette couche.

Nous avons appliqué le PMC pour la classification d'un ensemble de distances correspondantes à 40 personnes différentes, pour chaque personne nous avons pris 10 mesures déduites de 2 séquences de 5 images chacune (5 pour l'apprentissage et 5 pour le test). L'ensemble des images utilisées est alors 400 (200 pour l'apprentissage et 200 pour le test).

### 4.2.2 Apprentissage

En utilisant l'architecture du réseau, décrite précédemment, nous avons effectué une première phase d'apprentissage sur les éléments (distances) caractéristiques de 40 visages différents. Sachant que le nombre de distances extraites de chaque image est 8, le nombre total de distances utilisées est  $40 \times 5 \times 8 = 3200$  distances. Les étapes à suivre sont les mêmes utilisées dans l'approche pixel (organigramme figure 4.8) en changeant les paramètres du réseau bien sûr.

Durant la phase d'apprentissage l'erreur n'a atteint sa valeur (estimée) minimale (0.02) qu'après 20000 itérations (figure 4.9)



**Figure 4.9** : Evolution de l'erreur en fonction du nombre d'itérations pour l'approche géométrique

### 4.2.3 Reconnaissance

Après avoir obtenu les meilleurs poids (pour lesquels l'erreur est minimale) et biais nous avons muni notre réseau de ces informations nécessaires pour la reconnaissance du moins des 40 personnes appris au réseau. Nous avons, ensuite, effectué plusieurs tests sur des personnes faisant partie de la base appris au réseau et des personnes complètement inconnus. Les résultats des tests de la base d'apprentissage pour les 40 personnes sont illustrés dans le Tableau 4.3.

Utiliser l'approche géométrique pour la reconnaissance des visages nous a certainement pris plus d'itérations que l'approche pixel pour minimiser l'erreur et fixer l'architecture du réseau mais elle nous a offert par contre une minimisation considérable du taux de rejet 9% (6% de moins que l'approche pixel).

L'avantage de cette méthode est qu'elle se base sur la géométrie propre de chaque visage qui est loin d'être sensible à l'intensité ou au changement d'éclairage (peau bronzée ou éclairage mal réglé). Le seul facteur auquel elle est sensible, est le temps (changement des traits du visage en vieillissant), s'il s'agit de dizaine voir de vingtaine d'années, cependant nous sommes loin de tomber sur ce type de problèmes car l'écart entre notre base d'apprentissage et base de test est de 3 mois seulement (une durée pendant laquelle la couleur de peau peut changer mais pas la géométrie du visage).

	Taux de Reconnaissance (en %)	Taux de Rejet (en %)
<b>Base d'apprentissage</b>	99%	1%
<b>Base de test</b>	91%	9%

**Tableau 4.3:** Les différents taux de l'approche géométrique

### 4.3 Résultats de la méthode d'analyse en composantes principales (ACP)

En se référant au chapitre 3, la mise en œuvre de l'ACP nécessite les 6 étapes décrites dans le paragraphe suivant (figure 4.10).

#### 4.3.1 Etapes de l'ACP pour la classification de visages

**Etape 1 :** Calcul du tableau de données centrées et réduites par rapport aux variables  $\tilde{X}$  :

$$x'_i = (x_i - x_m) / s_i \quad (4.2)$$

où  $x_m$  est la moyenne des  $x_i$  et  $s_i$  est l'écart type défini par :

$$s_i = \sqrt{\sum_{i=1}^n ((x_i - x_m)^2 / n)} \quad (4.3)$$

**Etape 2 :** Calcul des matrices de variances  $V_{ii}$ , de covariances  $V_{ij}$  et de corrélation  $R_{ij}$  :

$$V_{ii} = \sum_{i=1}^n (x_i - x_m)^2 / n \quad (4.4)$$

$$V_{ij} = \sum_{i=1}^n \sum_{j=1}^n (x_i - x_{mi})(x_j - x_{mj}) / n \quad (4.5)$$

$$R_{ij} = V_{ij} / (V_{ii} V_{jj})^{1/2} \quad (4.6)$$

**Etape 3 :** Calcul des valeurs propres  $\lambda_i$  et des vecteurs propres  $V_i$  où  $\lambda_i$  sont les solutions de l'équation  $\det (R - \lambda_i \cdot I) = 0$  :

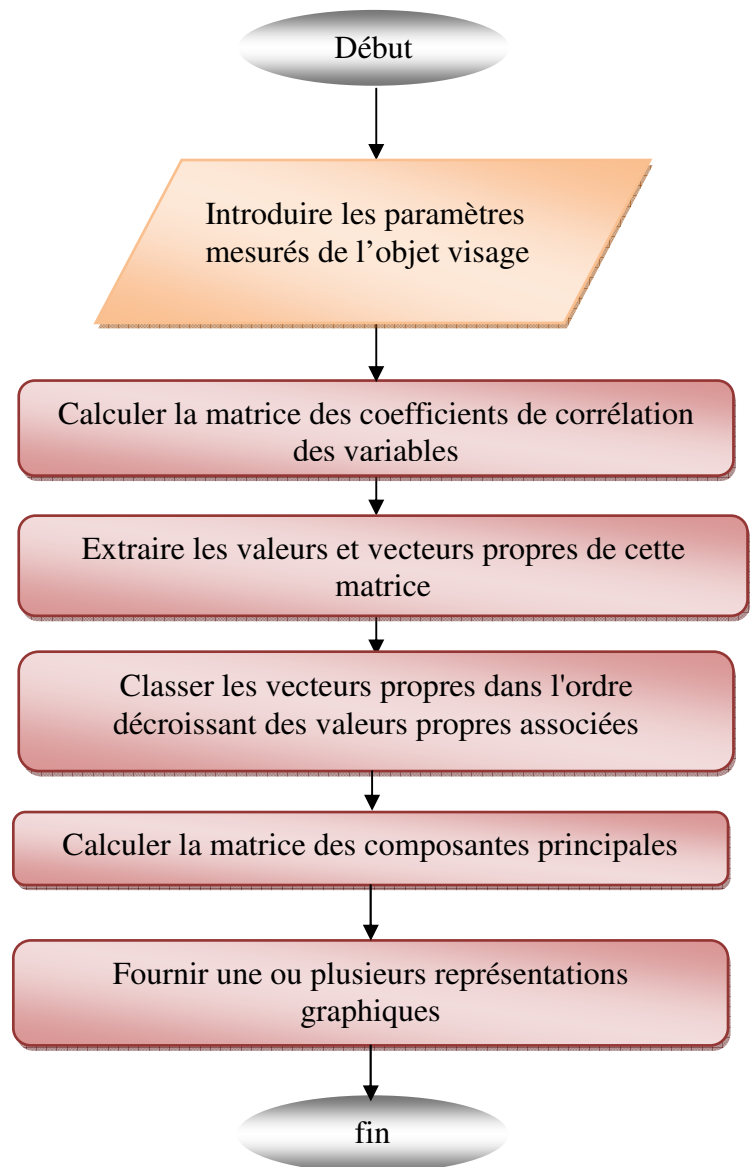
$$R * V_i = \lambda_i * V_i \quad (4.7)$$

**Etape 4 :** Classement des vecteurs propres dans l'ordre décroissant des valeurs propres associées

**Etape 5 :** Calcul des composantes principales :

$$CP_i = X_{cr} * V_i \quad (4.8)$$

**Etape 6 :** Représentations graphiques.



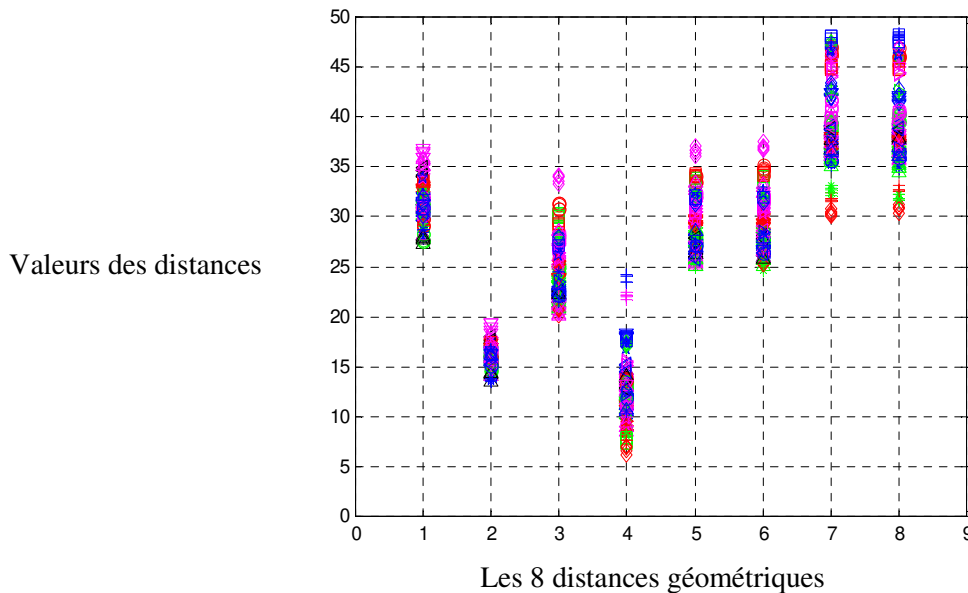
**Figure 4.10 :** Organigramme de la méthode de l'ACP

### 4.3.3 L'application

Pour notre application, nous avons suivi les étapes illustrées dans l'organigramme de la figure 4.10.

Nous avons utilisé les mêmes données et même procédure que pour l'approche géométrique.

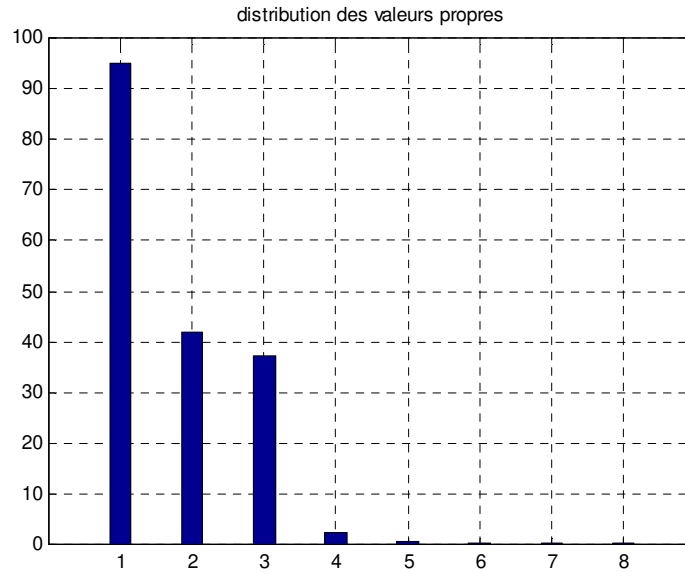
Dans la figure 4.11, nous avons présenté les 200 visages des 40 personnes (5 images par personne) à 8 dimensions (distances extraites de chaque image) dans un plan de 2 dimensions.



**Figure 4.11:** Présentation des 200 visages de 40 personnes dans un espace de dimension 2

Cette représentation ne donne aucune information et ne nous permet pas d'analyser ou de classifier les visages selon les distances associées. Ceci pour monter l'avantage d'utiliser l'analyse en composante principale pour la représentation d'une telle quantité d'information car l'ACP, comme signalé dans le chapitre 3, cherche à diminuer l'espace de représentation à deux ou trois dimensions sans perdre le maximum d'information.

En triant les valeurs propres de la plus grande à la plus petite et même chose pour les vecteurs propres associées (étape 4 de l'ACP), nous avons obtenu le graphe de distribution des valeurs propres suivant.



**Figure 4.12** : Graphe de distribution des valeurs propres

Le calcul des composantes principales permet de donner de nouvelles coordonnées pour chaque personne. Ces résultats permettent d'établir une nouvelle représentation suivant les 2 ou 3 composantes principales (axes principaux) au lieu des 8 dimensions précédentes.

La figure 4.13 donne la représentation des 40 personnes (avec 5 visages chacune) dans le plan des composantes principales 1 et 2, et la figure 4.14 donne celle des mêmes personnes suivant les composantes 1 et 3. Pour simplifier la représentation, nous avons codé chaque personne par une forme géométrique (triangle, rectangle, étoile, cercle, ...etc), une légende des différentes codifications est associée à chaque graphe.

Suivant cette représentation, nous remarquons que les classes représentant les différentes personnes sont séparées les unes des autres sauf certaines qui chevauchent certainement à cause de la ressemblance de certaines personnes (qui ont des paramètres proches les uns des autres).

Pour pouvoir distinguer entre ces classes trop proches, il suffit de combiner la représentation de composantes principales (par exemple (CP1, CP2), (CP1, CP3), (CP1, CP4),...etc)

Pour la classification d'une personne à reconnaître, nous projetons ces composantes sur cette représentation ou il peut appartenir à une classe existante ou pas.

Les figures 4.16 et 4.17 illustrent le cas de projection d'une personne inconnue dans le plan de composantes principales 1 et 2 pour la première figure et les composantes 1 et 3 pour la seconde. Les figures 4.18 et 4.19 illustrent la même opération reprise avec une autre personne. La différence entre les deux cas est que pour la première personne la projection est proche des autres classes et pour la seconde, sa projection est séparée des autres classes.

En fin, pour décider à quelle classe la personne appartient, nous calculons la matrice de corrélation du vecteur projeté avec ceux des différentes classes. Le coefficient le plus grand permet de définir la classe d'appartenance et affiche la personne reconnue.

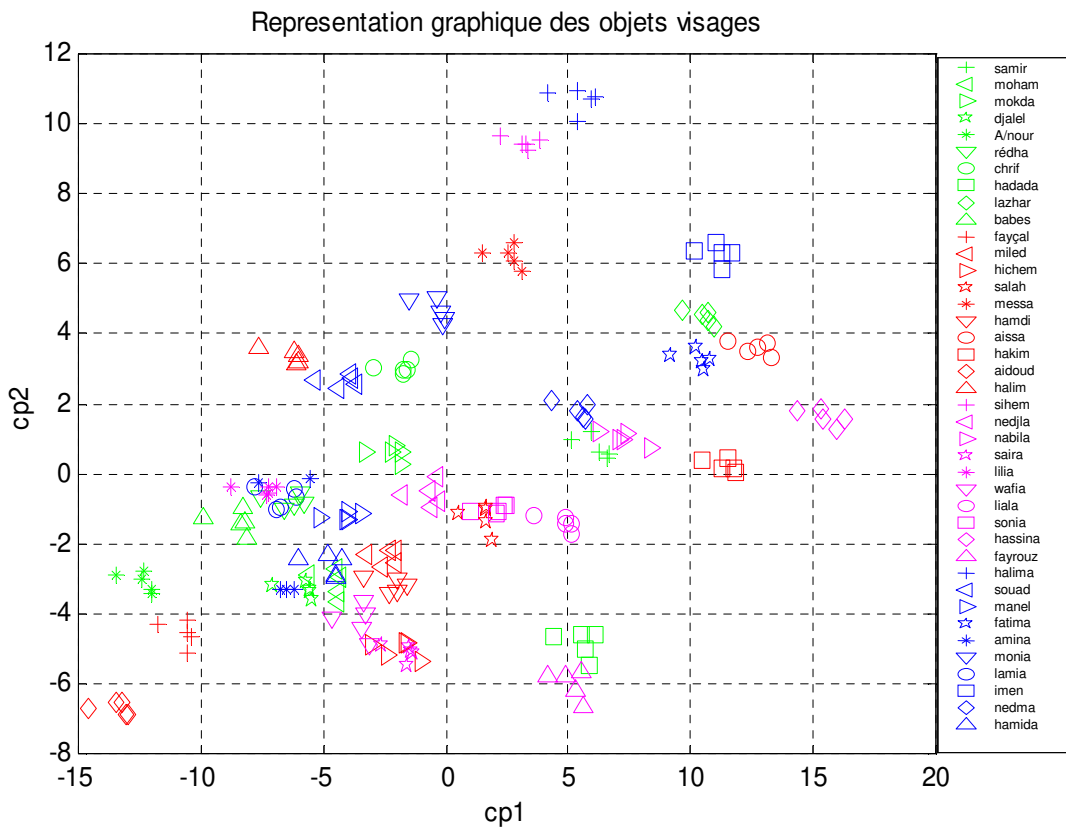
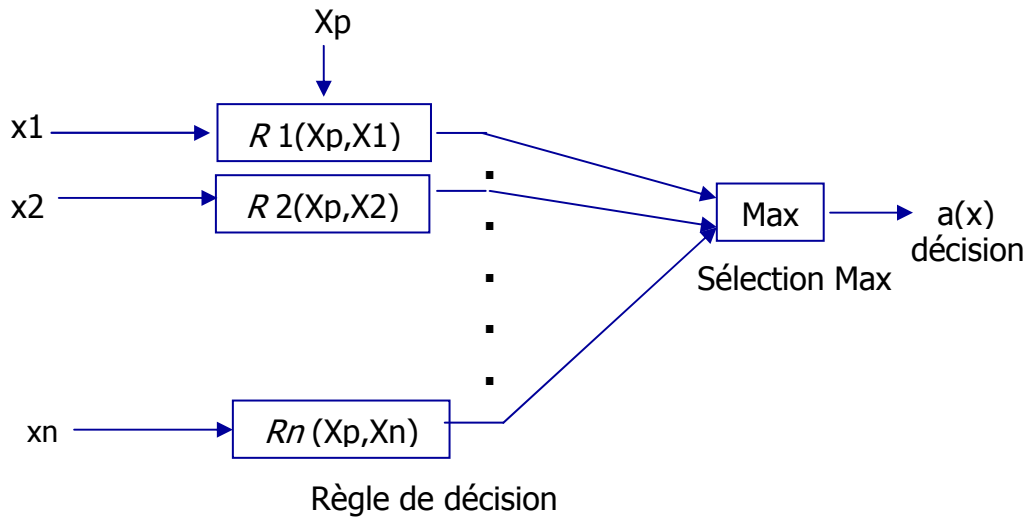
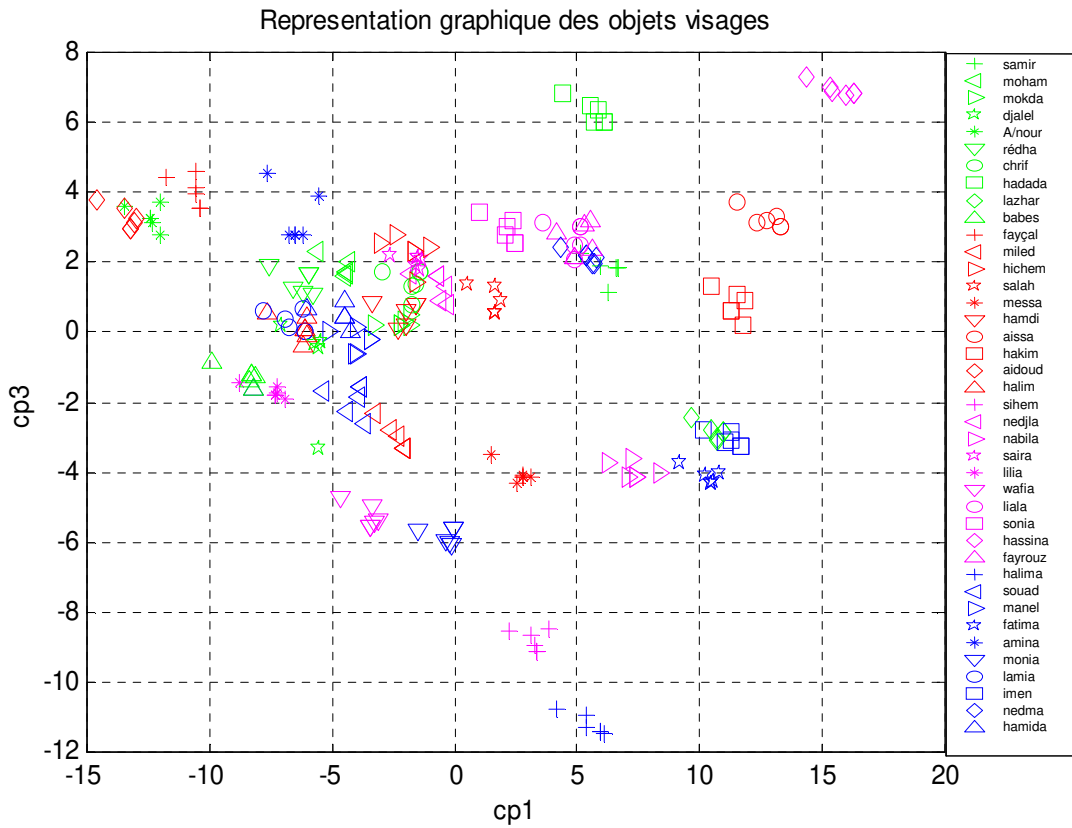


Figure 4.13: Représentation graphique des objets visages sur cp1 et cp2



**Figure 4.14:** Représentation graphique des objets visages sur cp1 et cp3

En revenant aux exemples représentés sur les figures 4.16 à 4.19, les deux personnes ont été reconnues parfaitement. La figure 4.15, illustre ces résultats.

Les tests effectués sur la base de projection (appelée base d'apprentissage précédemment) ont donné un taux de rejet de 0 % soit une bonne séparation des 40 classes de la base de projection. Les mêmes tests ont été répétés sur la base test, les résultats obtenus ne sont pas loin de ceux du PMC de l'approche géométrique 0,5 % du moins pour le taux de rejet.

Nous estimons que le taux de reconnaissance est satisfaisant, mais nous pensons pouvoir l'améliorer par la méthode que nous proposons dans le paragraphe suivant (cf 4.4).

	Taux de Reconnaissance (en %)	Taux de Rejet (en %)
<b>Base de projection</b>	100%	0%
<b>Base de test</b>	91.5%	8.5%

**Tableaux 4.4:** Résultats de la classification par ACP



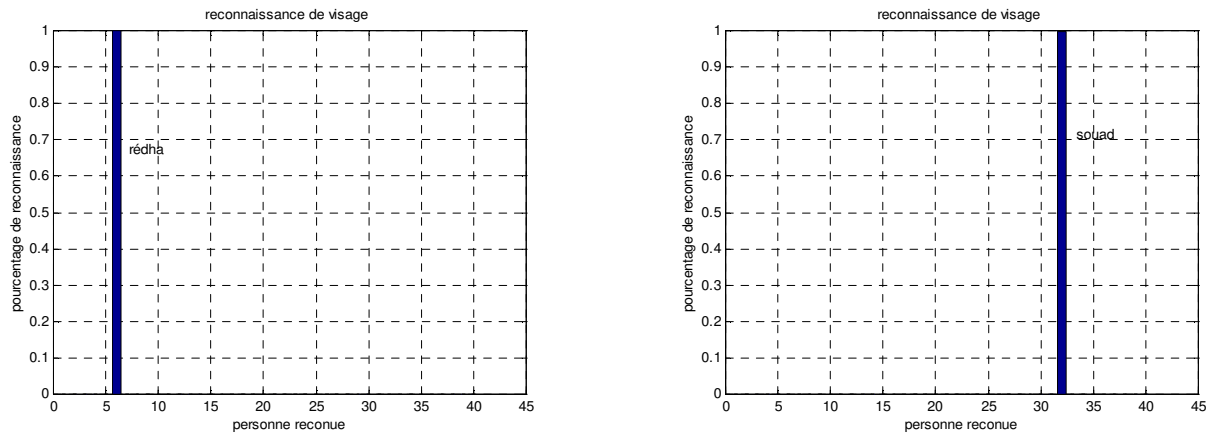


Figure 4.15 : Résultats de la reconnaissance de deux personnes différentes

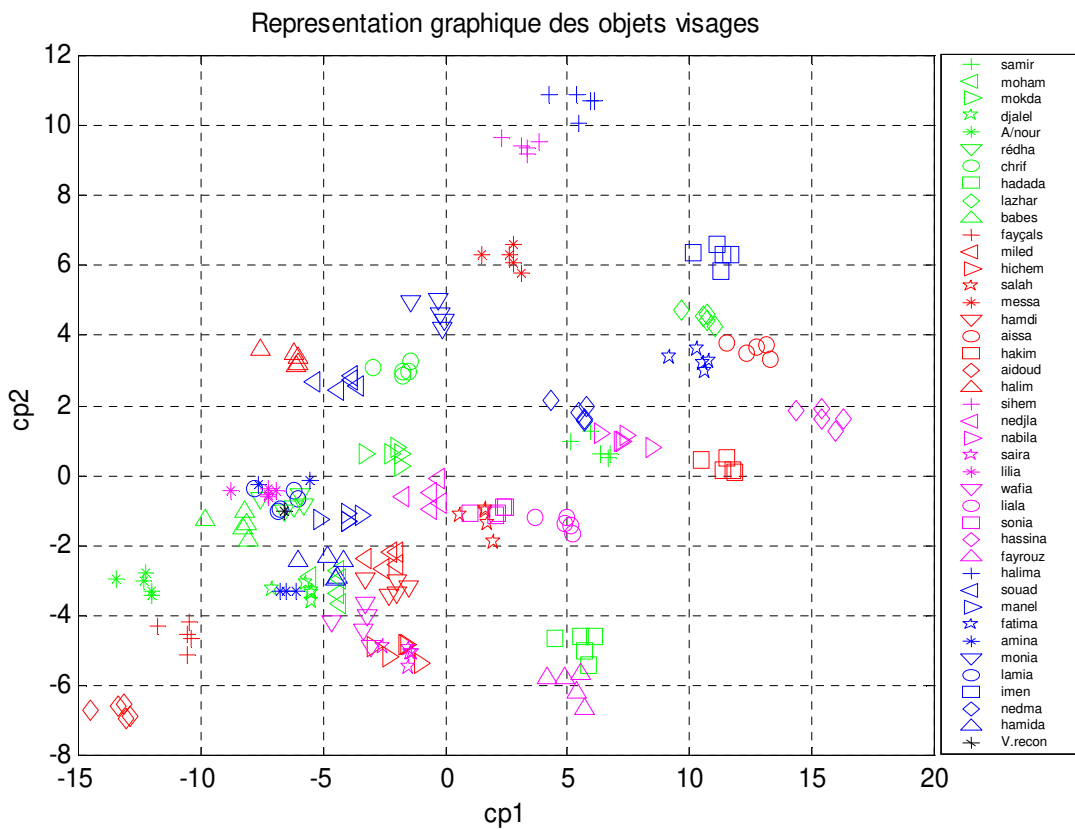
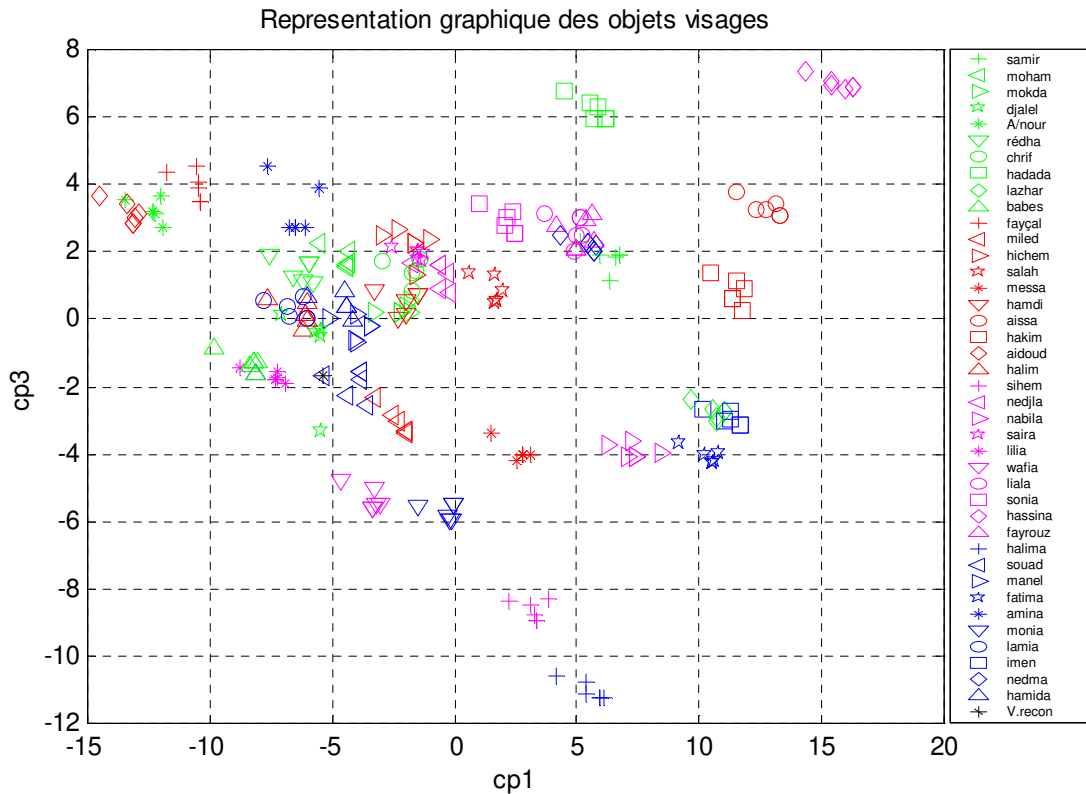


Figure 4.16: Représentation graphique d'une personne à reconnaître sur cp1 et cp2





**Figure 4.19** : Représentation graphique de personnes à reconnaître sur cp1 et cp3

## 4.4 Classification par méthode hybride Neuro-ACP

### 4.4.1 Principe

L'analyse en composante principale est une méthode qui a conduit à des résultats considérables dans le domaine de la classification et avec sa simplicité et sa fiabilité elle ne présente presque pas d'inconvénient.

Les réseaux de neurones font aussi partie des meilleures méthodes de classification, seulement leur taux de reconnaissance reste toujours à améliorer. C'est pour cette raison nous avons essayé d'hybrider ces deux méthodes [BS01].

Rappelons que les paramètres de la méthode de l'ACP, qui sont les coordonnées des personnes suivant les axes principaux, correspondent aux lignes de la matrice des composantes principales où chaque ligne regroupe 8 paramètres de l'un des visages.

L'idée [GY00], sur laquelle nous nous sommes basés pour développer cette nouvelle méthode [Bo05], est au lieu de fournir au réseau de neurones les distances, nous les remplaçons par les 8 paramètres de l'ACP (figure 4.20).

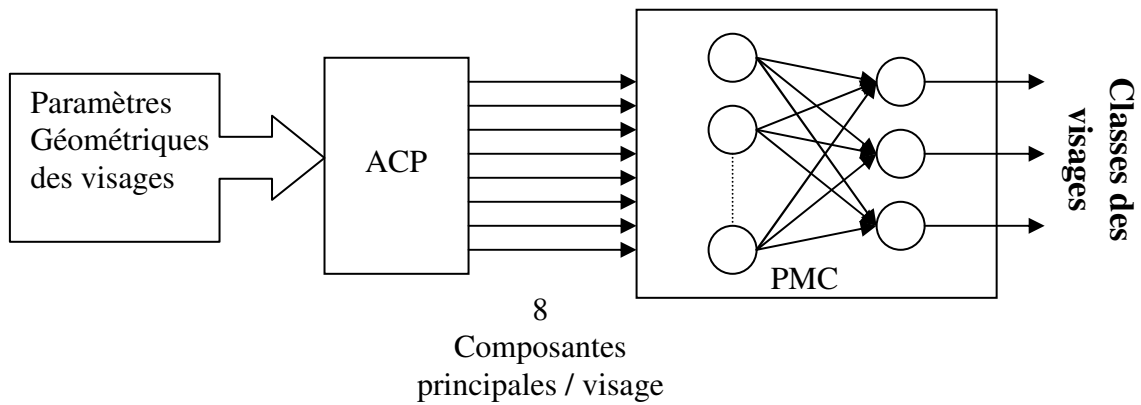


Figure 4.20 : Hybridation du PMC avec l’ACP

#### 4.4.2 Implémentation et résultats

Rappelons que les paramètres de la méthode de l'ACP, qui sont les coordonnées des personnes suivant les axes principaux, correspondent aux lignes de la matrice (200x8) des composantes principales où chaque ligne regroupe 8 paramètres de l’un des 200 visages.

L’idée sur laquelle nous nous sommes basés pour développer cette nouvelle méthode est de remplacer les 8 distances extraites du visage par ces 8 paramètres de la même personne.

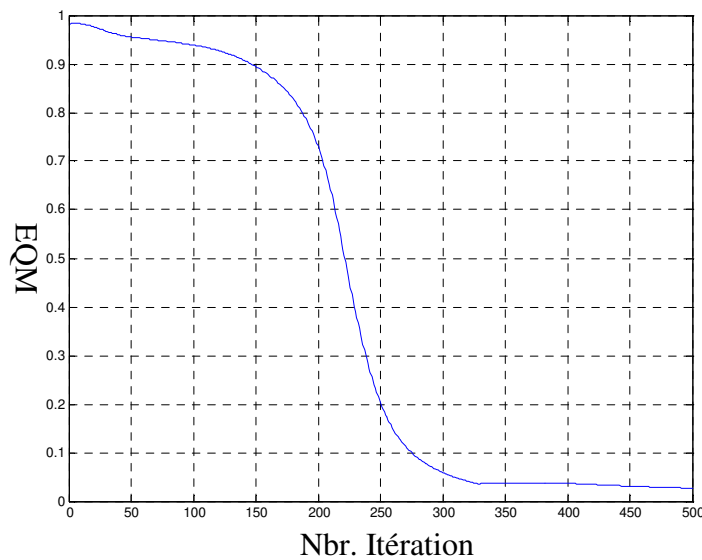
Pour cela nous avons utilisé le même PMC que celui de la méthode géométrique (avec la même structure et le même principe d’apprentissage). Rappelons que cette structure (8 neurones d’entrée, 30 neurones pour la couche cachée et 40 neurones de sortie) le pas d’apprentissage de 0,75 sont obtenus après plusieurs essais et ils ont conduit à une erreur de 0,0267 pour un nombre d’itérations égal à 500 (voir figure 4.21)

Le tableau ci-dessous illustre les résultats obtenus :

Il est clair que l’hybridation du PMC avec l’ACP a augmenté les performances des deux méthodes utilisées séparément et le taux de rejet a diminué de 3%, soit un taux de reconnaissance de 94.5 %. Cette amélioration nous permettra de distinguer entre les différentes classes et d’en reconnaître presque les 40 personnes.

	Taux de Reconnaissance (en%)	Taux de Rejet (en%)
Base d’apprentissage	99%	1%
Base de test	94.5%	5.5%

Tableau 4.5: Résultats de la méthode neuro-ACP



**Figure 4.21** : Evolution de l'erreur en fonction du nombre d'itérations pour l'approche neuro-ACP

## 5. Conclusion

On peut conclure que l'approche géométrique donne des résultats meilleurs que ceux de l'approche pixels. Cela peut être interprété par le fait que l'information pixel est très sensible aux changements des conditions d'acquisitions et de l'état de la personne.

Concernant les méthodes de classification, les réseaux de neurones conduisent à des résultats acceptables, seulement le temps de réponse est assez élevé et les conditions de mise en œuvre du système sont très lourdes.

L'ACP est une méthode directe et simple à mettre en œuvre et donne de bons résultats pour la classification reste à choisir la méthode de décision.

L'idée d'hybrider les RNA plus précisément le PMC avec l'ACP nous a permis de d'associer les performances de chaque méthode et donc d'obtenir des résultats meilleurs que ceux des méthodes séparées.

# DÉTECTION DE VISAGES

## CHAPITRE 5

1. Introduction .....	82
2. Détection de visages : Etat de l'art.....	82
2.1 Approches basées sur les connaissances acquises....	82
2.2 Approches basées sur les caractéristiques invariables	83
2.3 Approches basées sur la mise en correspondance.....	86
2.4 Approches basées sur l'apparence globale: Eigenface	86
3. Problèmes de la détection de visages.....	88
4. Principe de la détection de visages.....	89
4.1 Le suivi de personnes.....	89
4.2 La détection de visages.....	89
4.3 Mesure des performances d'un système de détection de visages.....	90
5. Notion de couleur.....	91
5.1 Représentation fréquentielle.....	91
5.2 Espace de couleurs.....	92
6. Traitement des images binaires.....	95
6.1 Images binaires.....	95
6.2 Méthodes de traitement des images binaires.....	95
7. Méthode adoptée.....	98
7.1 Détection de la peau.....	98
7.2 Filtrage et séparation des régions.....	104
7.3 Template Matching.....	108
7.4 Mise en œuvre du modèle.....	111
8. Résultats et interprétations.....	112
8.1 Test de détection de couleur de peau.....	113
8.2 Test de détection de visages.....	114
8.3 Discussion des résultats.....	117
9. Conclusion.....	117

## 1. Introduction

Les visages constituent une catégorie de stimulus importante par la richesse des informations qu'ils véhiculent. Ils sont à la fois des vecteurs visuels principaux à l'identité des individus et des vecteurs essentiels de communication (verbale et non verbale). La détection automatique de visages est un problème très important. Du point de vue application, elle est à la base de tout système de reconnaissance de visages, car avant de reconnaître n'importe quelle personne, il est indispensable de localiser son visage.

## 2. Détection de visages : Etat de l'art

Une grande variété de méthodes de détection de visages a été proposée dans les dernières années. Seulement, la plupart d'entre elles s'attardent à répondre aux questions : "Y a-t-il des visages dans cette image ? Si oui, où sont-ils ?", Sans toutefois extraire les caractéristiques du visage. Or, il est possible dans plusieurs cas de dériver ou de spécialiser la technique afin d'obtenir ces informations supplémentaires. Les prochaines sous-sections résumeront les principales méthodes de détection [Le03]. Le fonctionnement de la reconnaissance de visage, quelle que soit la technique utilisée, se base toujours sur une capture d'image (vidéo par exemple) ou sur une photo. Les caractéristiques du visage sont ensuite extraites selon la méthode choisie, puis enregistrées dans une base de données.

Il est possible de distinguer 4 approches théoriques [FL04] :

- **Approches basées sur les connaissances acquises (Knowledge-based method).** Représentation intuitive du visage : présence de la bouche, des yeux, etc.
- **Approches basées sur les caractéristiques invariables (Feature invariance).** Les traits du visage qui ne changent pas quelque soient les conditions de luminosité, etc.
- **Approches basées sur la mise en correspondance (Template matching).** Représentation paramétrique d'un visage qui constitue un modèle ; l'étude tourne autour d'une corrélation avec le visage à traiter.
- **Approches basées sur l'apparence globale (Appearance-based method).** Décomposition de l'image en éléments simples (présentation du système Eigenface).

### 2.1 Approches basées sur les connaissances acquises

Cette méthodologie s'intéresse aux parties caractéristiques du visage comme le nez, la bouche et les yeux. Les positions relatives des différentes composantes du visage sont étudiées après avoir été détectées. La difficulté dans cette approche est de traduire par des règles strictes, à définir, la manière dont le chercheur représente le visage. Si ces règles sont trop précises, elles

n'identifient pas certains visages (false negative). Dans le cas contraire, elles provoquent de fausses alertes (false positive). Il est évident qu'il est impossible de définir des règles qui tiennent compte de toutes les variabilités comme la position du sujet, par exemple.

Afin d'éliminer de manière progressive les parties de l'image qui ne correspondent pas à un visage, une hiérarchie de méthodes est appliquée [KP97][YK02]. L'objectif est de n'appliquer les méthodes précises (et donc coûteuses en temps) qu'aux parties candidates potentielles de l'image. Selon cette approche, une hiérarchie d'images à résolutions multiples (mêmes images avec des résolutions différentes) est utilisée. En effet, avec un très petit nombre de pixels, il est possible de sélectionner quelques zones images candidates à la représentation d'un visage. Chacune de ces zones sera ensuite étudiée plus précisément (en augmentant la résolution) à l'aide d'histogrammes et de détections de contours. Enfin, la méthode cherche à détecter les composantes caractéristiques du visage (yeux). Le taux de réussite de cette méthode reste faible avec surtout un grand nombre de fausses alertes et ne peut donc se suffire à elle-même. En revanche, elle permet une focalisation rapide sur les parties intéressantes de l'image.

## 2.2 Approches basées sur les caractéristiques invariables

L'objectif de cette méthodologie est de trouver les caractéristiques invariantes du visage.

### 2.2.1 Les traits du visage

- **Contours**

Algorithme de Sirohey [Si93] : Il s'agit d'une méthode de segmentation d'images utilisant une carte de contours et des considérations heuristiques. Il regroupe les contours qui correspondent à ceux du visage et rejette les autres. Son algorithme obtient 80 % de réussite.

- **Taches et rayures**

Chetverikov [CL93] utilise non pas des bordures mais des "taches" et des séquences linéaires de traits de bordure d'orientation similaire. L'étude se base sur un modèle de visage composé de deux taches sombres pour les yeux, trois taches claires pour les pommettes et le nez; les lignes du visage étant caractérisées par les traits d'orientation similaire. La configuration triangulaire des taches est étudiée ultérieurement.

- **Images en nuance de gris**

Un filtrage passe-bande est appliqué à l'image, permettant la mise en valeur des composants comme le nez ou la bouche (de gradient caractéristique). A partir de l'histogramme, on peut alors former des seuillages successifs pour délimiter l'image en zones.

- **Rapports de brillance**

Si les conditions de brillance changent de manière sensible entre deux visages, il n'en est pas de même des rapports de brillance entre deux parties du même visage. En ne gardant que les



"directions" (à partir de ces rapports) d'un nombre assez restreint de parties du visage, on obtient un invariant relativement robuste.

- **Distances mutuelles**

Il s'agit d'une approche probabiliste. Cette méthodologie cherche certaines composantes caractéristiques du visage telles que le nez ou la bouche. Dans un second temps, elle estime les positions relatives entre les éléments grâce à une méthode statistique. Ainsi, il est possible d'éliminer certaines zones de l'image ne correspondant pas à un visage.

- **Morphologie**

L'objectif est d'extraire des segments analogues à des yeux car il s'agit des composants les plus stables et saillants du visage humain. Ces "eye-analogue segments" sont définis comme les contours des yeux : Extraction des pixels dont l'intensité change de manière significative ? Détection des segments qui sont susceptibles d'appartenir à des yeux ? Considération de combinaisons géométriques pour détecter les autres éléments du visage candidat ? Vérification par réseau de neurones.

### 2.2.2 Les textures

Il est possible d'identifier un visage à partir de sa texture. En effet, trois types de textures sont utilisées : la peau, les cheveux, et "les autres". Les textures sont calculées sur des masques de 16\*16 pixels. Puis, un réseau de neurones réalise une corrélation en cascade dans le but de classifier de manière supervisée ces textures. Aucun test n'a encore été réalisé pour la localisation de visages.

### 2.2.3 La couleur de la peau

Il est intéressant, lorsque l'image initiale est en couleur, de pouvoir se servir de cette information afin d'améliorer la localisation du visage. Dans la plupart des cas, la peau peut être vue comme une portion de l'image d'une couleur déterminée. En utilisant les frontières de cette zone comme seuil pour une image, il est possible d'extraire les pixels dont la couleur peut s'apparenter à celle de la peau [Le03].

Des études ont montré que la variabilité de la couleur de la peau tenait plus de la différence d'intensité plutôt que de la chrominance. On peut utiliser plusieurs types d'espaces de couleurs (RGB, RGB normalisé, HSV, YcrCb, YIQ, YES, CIE XYZ, CIE LUV, etc).

L'un des espaces le plus couramment utilisé pour effectuer cette tâche est le HSV. Son principal avantage réside dans le fait qu'une de ses composantes représente la luminosité (le V de Value). Ainsi, il est possible de travailler sur le visage sans se soucier de la luminosité de l'image en général. Il suffit alors de se contenter des canaux **H** (Hue) et **S** (Saturation) pour extraire les pixels de peau de l'image.

Cette méthode n'est pas infaillible, et il faut noter que certaines détections seront incorrectes ou incomplètes. Par exemple, le front est souvent surexposé à la lumière et semble ainsi plus éclairé, la couleur en est donc modifiée. L'information de couleur est un outil efficace pour identifier des zones du visage cependant, cela n'est pas utilisable lorsque le spectre de couleurs varie de manière significative entre l'arrière plan et le visage. La Figure 5.1 illustre un exemple d'extraction ; la figure (a) représente l'image originale et la figure (b) illustre le résultat d'extraction [He02] [CR00] [Le03].



(a) Image originale  
Base d'Images BAO [Fr]



(b) Image après extraction des  
régions de peau

**Figure 5.1** : Exemple d'extraction des régions de peau

Il est intéressant de remarquer les fausses détections pour les cheveux ainsi que les régions non détectées (par exemple le front). Cela peut s'expliquer par le fait que certaines zones reflètent davantage la lumière et semblent ainsi plus éclairées, ce qui modifie la couleur de la région dans l'espace HSV. Dans ce cas, bien précis, la saturation tend vers 0 et la luminance vers 1. Autrement dit, lorsque la saturation est basse et que la luminance est élevée, la couleur tend vers le blanc, échappant donc aux seuils de détection. Pour ce qui est des cheveux, ils peuvent correspondre dans certains cas aux couleurs de la peau.

Certains auteurs ont également proposé un apprentissage automatique des couleurs représentant la peau à l'aide de réseaux neuronaux. Ceux-ci peuvent être entraînés à partir d'échantillons de pixels représentant la peau (et de contre-exemples), préalablement converti dans l'espace de couleur YCrCb. Par ailleurs, d'autres espaces de couleurs peuvent être utilisés dans ce contexte.

Les fonctions de distributions gaussiennes ont aussi été employées dans plusieurs travaux pour la modélisation de la couleur peau. D'autres auteurs ont également proposé l'utilisation d'un modèle mixte de gaussiennes (Gaussian Mixture Model), qui semble mieux représenter et modéliser la portion d'un espace de couleurs associées à la couleur peau [CR00].

D'autres études ont cependant montré que la combinaison d'analyse de formes, de segmentation de couleurs et d'informations de mouvement pour la localisation et le suivi de visages dans des séquences conduisent à des résultats intéressants.

### 2.3. Approches basées sur la mise en correspondance (Template matching)

L'appariement de gabarit, ou template matching [Le03] est certainement une des techniques de détection de visages la plus simple à mettre en œuvre. Elle consiste en effet à comparer l'intensité des pixels entre un gabarit prédéfini et plusieurs sous-régions de l'image à analyser. Ce processus correspond en pratique à effectuer plusieurs balayages couvrant toute la superficie de l'image. Les endroits les plus propices à la présence de visages seront donc facilement identifiés par des minimums de distance entre le gabarit et l'image sous-jacente. Parmi les distances envisageables, il y a notamment la norme L1 (distance de Manhattan) et la norme L2 (distance euclidienne), la cross-corrélation, etc. Il est à noter que cette technique peut également être appliquée sur une carte d'arêtes, ce qui peut simplifier le calcul.

Les plus grands désavantages de cette méthode résident, sans contredit, au niveau de la pose du visage et de son échelle. Pour résoudre ce problème, plusieurs gabarits peuvent être définis pour différentes poses, mais ceci rend la gestion des résultats assez complexe. Par ailleurs, cette technique n'étant pas intrinsèquement multi-échelle, plusieurs balayages à différents niveaux de redimensionnement du gabarit sont nécessaires.

Ceci implique donc la réalisation d'une étape de décision supplémentaire dans le but de filtrer les multiples détections appartenant aux mêmes visages. Finalement, la construction d'un gabarit efficace représente un défi en elle-même. En effet, l'utilisation d'un gabarit plus ou moins adapté au type d'objet recherché peut nuire à une détection efficace et diminuer la précision des résultats.

Pour ce qui est de la localisation des différentes caractéristiques du visage, leurs positions peuvent être déduites à partir des positions correspondantes sur le gabarit. Elles doivent être préalablement déterminées en positions relatives par rapport aux dimensions du gabarit. Il n'est cependant pas garanti que le gabarit soit parfaitement positionné en translation, en échelle et en rotation sur le visage à détecter, ce qui produira des coordonnées légèrement erronées.

Par ailleurs, certains travaux [LA00][MY99] ont utilisé une détection des caractéristiques du visage à l'aide de gabarits plus spécialisés (par exemple les yeux, la bouche, etc.). Cette méthode implique par contre une recherche intensive dans un vaste espace de solutions possibles.

### 2.4 Approches basées sur l'apparence globale: Eigenface

#### 2.4.1 La méthode Eigenface

Une des technologies de reconnaissance de visages les plus connues est développée et supportée par le MIT et se nomme "Eigenface" [FL04]. Ce système décompose l'image en une série d'images teintées de nuances de gris, chacune mettant en évidence une caractéristique particulière du visage. Les zones claires et foncées ainsi créées forment les caractéristiques uniques

du visage et ce sont elles que l'on nomme *eigenface*. On en extrait ainsi de 100 à 125 par visage.

L'élément essentiel de cette méthode repose sur la réalisation d'une analyse en composantes principales (ACP) sur une série d'image représentant les visages à détecter. Cette réduction de dimension fournit les premiers vecteurs propres (*eigen*) représentant les plus fortes différences entre les points que nous analysons : nez, yeux, etc. L'ACP permet donc de distinguer une image contenant un visage d'une image n'en contenant pas. Seulement, elle n'est pas capable de reconnaître deux visages identiques mais d'expressions différentes [Le03].

Tout comme le *template matching*, l'étape de détection de l'*Eigenface* [FL04] consiste à effectuer un balayage de la zone à traiter. Ensuite, une reconstruction avec les premiers vecteurs propres est réalisée pour chacune des imagerie à analyser. Si l'image reconstruite possède suffisamment de points communs avec l'imagerie d'origine, alors la sous-image possède les caractéristiques discriminantes de la classe générée et peut être examinée correctement avec la base générée par les vecteurs propres. La distance entre ces deux images est calculée à l'aide d'une métrique particulière telle que celle utilisée pour l'appariement de gabarit.



**Figure 5.2** : Images utilisées pour l'apprentissage et Visages propres correspondants.

#### 2.4.2 Réseau de neurones

La détection du visage à l'aide de réseaux de neurones [Le03] se résume à l'utilisation d'un classifieur à deux sorties représentant la présence ou l'absence de l'objet recherché dans une sous-région de l'image. Le principe de base, identique à certaines techniques précédentes, consiste à balayer l'image avec une fenêtre d'attention de dimensions fixes et de réaliser la détection sur les sous-images. Néanmoins, il est encore une fois nécessaire d'effectuer plusieurs balayages à différentes résolutions pour ainsi réaliser une détection suffisamment robuste.

Différents types de réseaux de neurones peuvent être utilisés, mais pour la plupart d'entre eux, les données d'entrées sont l'intensité des pixels de l'image (après un prétraitement adéquat). Par ailleurs, le réseau peut utiliser d'autres caractéristiques extraites des sous-images, comme par exemple des arêtes ou des informations fréquentielles.

Afin de réaliser l'apprentissage du réseau de neurones artificiel, une banque d'images contenant des visages est nécessaire. Avant tout, ces images devront être redimensionnées pour être compatibles avec les dimensions requises par le nombre d'entrées du réseau. Par exemple, la couche d'entrée possédera 875 neurones pour des images de dimensions 25×35 pixels.

Pour pallier aux effets produits par les rotations possibles de la tête, il est possible d'ajouter un réseau de neurones routeur qui pré-normalise la sous-image en rotation. Ce filtrage ne corrige cependant pas une pose de la tête résultant d'une rotation axiale. Pour remédier à ce problème, il est envisageable d'ajouter des images d'entraînement contenant différentes poses.

Comme dans plusieurs autres techniques multi-échelles, une étape supplémentaire de prise de décision est nécessaire afin de réduire les réponses multiples pour un même visage. Finalement, les coordonnées des caractéristiques peuvent être déduites à partir des positions moyennes dans les images d'apprentissage. Ce processus de détection, tout comme les précédents, peut cependant manquer de précision lors de cette étape de localisation.

### 3. Problèmes de la détection de visages

La détection automatique de visages fait face à beaucoup de difficultés. Malgré la multitude de solutions proposées, une solution complète et efficace pour ce problème est encore recherchée ceci est dû essentiellement aux facteurs suivants :

- **Nombre de visages à détecté**, un nombre de visages élevé augmente les fausses alarmes surtout si le nombre de visage est inconnu.
- **Décor de la scène**, plus l'arrière plan est complexe (tableaux, portraits, meubles, etc...) plus il est difficile de localiser les visages.
- **Visages atypiques**, les systèmes basés sur les méthodes à base d'exemple de visage sont moins efficaces sur les visages moins courants tels que : chauves, des barbues, les personnes portant des lunettes ou chapeaux, etc...
- **Vitesse de traitement**, c'est un facteur important surtout lorsqu'il s'agit de détection sur une séquence vidéo.
- **Variation de pose**, la vue du visage change en raison de la pose relative du capteur utilisé (frontal, 45°, profil, à l'envers, etc...)
- **Expression faciale**, les systèmes basés sur les méthodes s'appuyant sur les paramètres faciaux sont sensibles aux changements des expressions faciales.
- **Occultation**, des visages peuvent être partiellement dissimulés par d'autres objets ou d'autres visages.

- **Condition de formation de l'image**, l'allure du visage peut être affecté par l'éclairage ou les caractéristiques de l'appareil photo.

#### 4. Principe de la détection de visages

##### 4.1 Le suivi de personnes

Localiser un objet dans une image consiste à repérer une position  $(X_0, Y_0, Z_0)$  dans un ensemble de positions possibles de taille  $L \times H \times P$ ,  $L$  et  $H$  étant respectivement la largeur et la hauteur de l'image traitée et  $P$  la profondeur testée [Fe97].

Les coordonnées  $(X_0, Y_0)$  correspondent à la position dans le plan et  $Z_0$  permet d'indiquer l'échelle et donc la taille du cadre contenant le visage



**Figure 5.3** : Résultat d'une détection de mouvement réalisée par une différence entre deux images

##### 4.2 La détection de visages

En vu d'une application, les algorithmes de détections de visages s'intègrent dans un schéma plus général (figure 5.5) que l'on peut décomposer en deux modules : le système de détection et le système de suivi. Le module de détection a pour but de localiser un objet particulier comme un visage. L'algorithme de poursuite de zones permet de le suivre sans avoir à le reconnaître à chaque image [AT02][BC03][RS02].

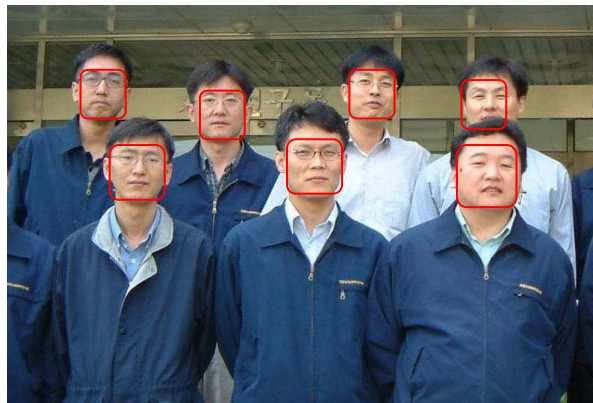
La détection de visages comprend trois étapes distinctes :

- L'utilisation d'indices, qui permet de segmenter l'image en zones qualifiées d'intérêt grâce à des critères de couleur, de mouvement, de contour, de textures,...etc.
- La normalisation des zones, qui a pour but d'obtenir la plus grande indépendance possible aux conditions d'acquisitions et d'éclairage.
- L'algorithme de détection de visages.

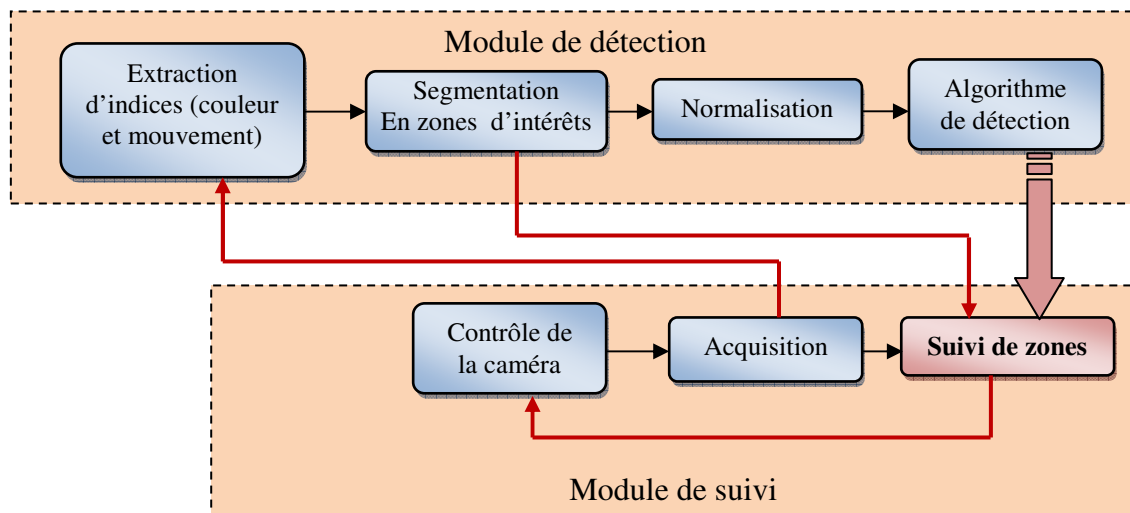
Localiser un visage dans une image soulève deux difficultés majeures :

- Le taux de fausses alarmes du module de détection doit être particulièrement bas.
- La seconde provient d'un problème technologique. Sans même tenir compte du temps de calcul nécessaire à l'algorithme, les ordinateurs ne permettent pas d'acquérir en temps réel des images de tailles plus grandes que 384 x 288. Les images traitées sont nécessairement de petites tailles et la détection doit pouvoir se faire à petite échelle.

Le système de suivi de visages est composé de deux parties : le module de détection qui permet de pointer et le module de suivi qui effectue la poursuite de la zone repérée.



**Figure 5.4 :** Visages détectés dans le plan (x,y) et en profondeur z.  
Image issue de la Base BAO [Fr]



**Figure 5.5 :** Système de détection et suivi de visages

### 4.3 Mesures des performances d'un système de détection de visages

La mesure des performances d'un système de détection de visage nécessite une connaissance des erreurs commises par ce dernier. En général, les détecteurs peuvent faire deux types d'erreur :

**Faute négative** (False negative): une région qui est un visage et qui n'est pas comptée.

**Faute positive** (False positive): une région est déterminée comme un visage par le détecteur, mais elle n'est pas un vrai visage.

Dans beaucoup de travaux récents de la détection de visages, la comparaison des résultats obtenus par plusieurs méthodes, se fait en termes de taux de détections vraies et taux de fausses alarmes tel que :

*Taux de détections vraies* = (Nombre de visages détectés / Nombre de visages réels)

*Taux de fausses détections* = (Nombre de fausses détections / Nombres de détections)

## 5. Notion de couleur

La couleur est l'un des éléments fondamentaux du système visuel. Notre œil n'est capable de percevoir qu'une dizaine de niveaux de gris alors qu'il est sensible à des milliers de couleurs. La couleur est une information utilisée par notre cerveau pour extraire les caractéristiques d'un objet observé. Dans ce contexte, l'idée d'utiliser la couleur pour faire de la reconnaissance d'objets paraît naturelle [KG93].

On peut définir la perception de la couleur d'un objet comme un phénomène qui fait intervenir la composition spectrale de la lumière incidente, les propriétés physiques de l'objet réfléchissant la lumière ainsi que le récepteur de cette lumière. Pour utiliser l'information de couleur, connaissant la lumière incidente et le type de capteurs permettant de récupérer la lumière réfléchie sur l'objet, on peut déterminer certaines propriétés physiques de l'objet. Deux types d'approches permettent de représenter la couleur [KG93]:

- La représentation fréquentielle.
- La représentation par espace de couleur.

### 5.1 Représentation fréquentielle

La représentation fréquentielle de la couleur est basée sur la composition spectrale de la lumière réfléchie. Etant donné une série de capteurs filtres  $N_i$ , dont les sensibilités spectrales sont aussi fines que possible, on peut caractériser la composition spectrale  $f(\lambda)$  de la lumière réfléchie par la réponse  $R_{N_i}$  du capteur  $N_i$  à cette lumière [KG93]:

$$R_{N_i} = \int_{\lambda^{(i)-c}}^{\lambda^{(i)+c}} f(\lambda) d\lambda \quad \text{Avec : } N : \text{ nombre de capteurs}$$

$C$  : la  $\frac{1}{2}$  largeur de la bande de chaque filtre.

Le problème majeur de cette approche est lié au très grand nombre de capteurs  $N_i$  nécessaires pour caractériser la lumière.



## 5.2 Espace de couleurs

Afin de pouvoir manipuler correctement des couleurs et échanger des informations colorimétriques, il est nécessaire de disposer de moyens permettant de les catégoriser et de les choisir. Ainsi, il n'est pas rare d'avoir à choisir la couleur d'un produit avant même que celui-ci ne soit fabriqué. Dans ce cas, une palette de couleurs est présentée, dans laquelle la couleur convenant le mieux au besoin est choisie. La plupart du temps le produit (véhicule, bâtiment, etc.) possède une couleur qui correspond à celle choisie.

En informatique, de la même façon, il est essentiel de disposer d'un moyen de choisir une couleur parmi toutes celles utilisables. Or la gamme de couleurs possibles est très vaste et la chaîne de traitement d'images passe par différents périphériques : par exemple un numériseur (scanner), puis un logiciel de retouche d'images et enfin une imprimante. Il est donc nécessaire de pouvoir représenter avec fiabilité la couleur afin de s'assurer de la cohérence entre ces différents périphériques [JR][DB00].

On appelle ainsi «espace de couleurs» la représentation mathématique d'un ensemble de couleurs. Il en existe plusieurs, parmi lesquels les plus connus sont :

- L'espace RGB (Rouge, Vert, Bleu, en anglais RGB, Red, Green, Blue).
- L'espace RGB normalisé.
- L'espace HSV (Hue, Saturation, Value, ou en français Teinte, Saturation, Luminance).
- L'espace YCbCr.

Le spectre de couleurs qu'un périphérique d'affichage permet d'afficher est appelé gamut ou espace colorimétrique. Les couleurs n'appartenant pas au gamut sont appelées couleurs hors-gamme.

### 5.2.1 Espace RGB

Le codage RGB, mis au point en 1931 par la Commission Internationale de l'Eclairage (CIE) consiste à représenter l'espace des couleurs à partir de trois rayonnements monochromatiques de couleurs [KG93]:

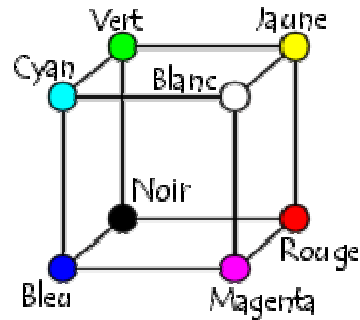
- Rouge (de longueur d'onde égale à 700,0 nm),
- Vert (de longueur d'onde égale à 546,1 nm),
- Bleu (de longueur d'onde égale à 435,8 nm).

Cet espace de couleur correspond à la façon dont les couleurs sont généralement codées informatiquement, ou plus exactement à la manière dont les tubes cathodiques des écrans d'ordinateurs représentent les couleurs.

Ainsi, le modèle RGB propose de coder sur un octet chaque composante de couleur, ce qui correspond à 256 intensités de rouge, 256 intensités de vert et 256 intensités de bleu, soient

16777216 possibilités théoriques de couleurs différentes, c'est-à-dire plus que ne peut en discerner l'oeil humain (environ 2 millions). Toutefois, cette valeur n'est que théorique car elle dépend fortement du matériel d'affichage utilisé [JR].

Etant donné que le codage RGB repose sur trois composantes proposant la même gamme de valeurs, on le représente généralement graphiquement par un cube dont chacun des axes correspond à une couleur primaire :



**Figure 5.6 :** Présentation des couleurs dans l'espace RGB

### 5.2.2 Espace RGB Normalisé

Les couleurs normalisées ont été efficacement employées pour segmenter des images de couleur dans beaucoup d'applications. La représentation des couleurs dans l'espace RGB n'est pas appropriée à caractériser certaines couleurs et textures tel que la couleur de peau.

Dans l'espace RGB, le composant triple (R, G, B) représente non seulement la couleur mais également la *luminance* qui peut être éliminée par un procédé de normalisation permettant une représentation fiable des différentes couleurs de peau, et aux différents états d'éclairage [CR00].

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad (5.1)$$

Comme, la somme des trois composantes (r g b) est connue ( $r + b + g = 1$ ), la valeur de la troisième composante devient inutile et peut être omise, ce qui réduit la dimension de l'espace.

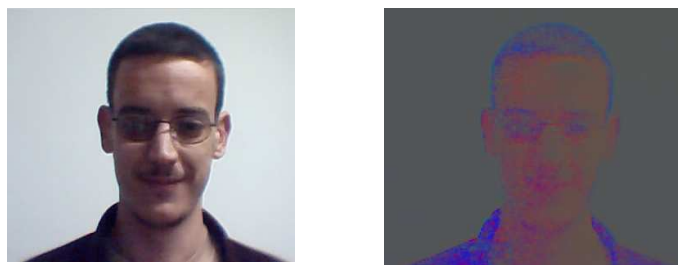


Image avant normalisation    Image après normalisation

**Figure 5.7 :** Exemple de normalisation d'une image RGB

### 5.2.3 Espace HSV (HSL)

Le modèle HSV, s'appuyant sur les travaux du peintre Albert H. Munsell, est un modèle de représentation dit "naturel", c'est-à-dire proche de la perception physiologique de la couleur par l'oeil humain. En effet, le modèle RGB aussi adapté soit-il pour la représentation informatique de la

couleur ou bien l'affichage sur les périphériques de sortie, ne permet pas de sélectionner facilement une couleur. En effet, le réglage de la couleur en RGB dans les outils informatiques se fait généralement à l'aide de trois glisseurs ou bien de trois cases avec les valeurs relatives de chacune des composantes primaires, or l'éclaircissement d'une couleur demande d'augmenter proportionnellement les valeurs respectives de chacune des composantes. Ainsi le modèle HSV a-t-il été mis au point afin de pallier à cette lacune du modèle RGB [JR].

Le modèle HSV consiste à décomposer la couleur selon des critères physiologiques :

- **La teinte** (en anglais *Hue*), correspondant à la perception de la couleur,
- **La saturation**, décrivant la pureté de la couleur, c'est-à-dire son caractère vif ou terne,
- **La luminance** (en anglais *value*), indiquant la quantité de lumière de la couleur, c'est-à-dire son aspect clair ou sombre.

Le modèle HSV a été mis au point dans le but de permettre un choix interactif rapide d'une couleur, pour autant il n'est pas adapté à une description quantitative d'une couleur.

Voici une représentation graphique du modèle HSV, dans lequel la teinte est représentée par un cercle chromatique et la luminance et la saturation par deux axes (figure 5.8) :

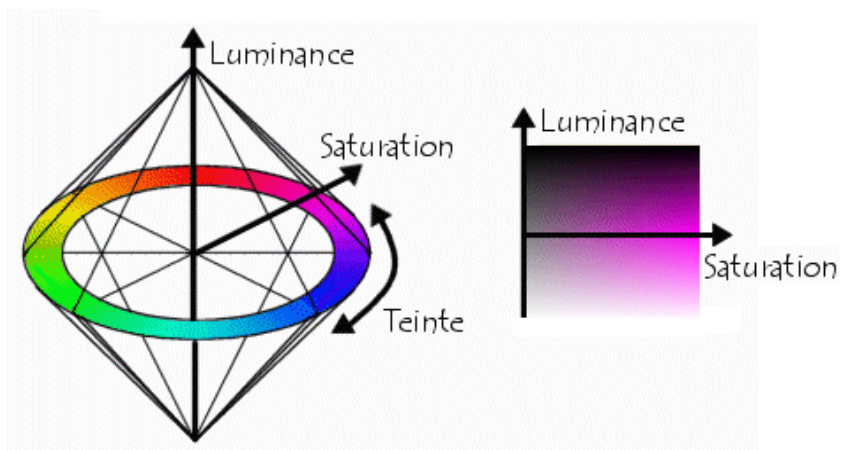


Figure 5.8 : Présentation des couleurs dans l'espace HSV

#### 5.2.4 Espace YCbCr

Utilisée souvent dans la compression des images, la luminance (appelée aussi Luma) est séparée de la couleur, présentée par la valeur Y et calculée par une somme pondérée des composantes de R, G et B [JR] :

$$Y = 0,299 * R + 0,587 * G + 0,114 * B \quad (5.2)$$

Les deux autres composantes de cette espace représentent l'information chromatique et elles sont calculées à partir de *Luma* :

$$Cr = R - Y \quad (5.3)$$

$$Cb = B - Y \quad (5.4)$$

## 6. Traitement des images binaires

### 6.1 Images binaires

Une image binaire est une image pour laquelle chaque pixel ne peut avoir pour valeur que 0 ou 1. La manipulation de telles images regorge d'outils spécialisés ainsi que de théories mathématiques pour plusieurs raisons :

- Les débuts du traitement des images numériques ne permettaient pas le traitement d'images complexes (problème de temps de calcul, d'espace mémoire disponible et qualité des périphériques de sortie). De plus, les premières applications (reconnaissance de caractères, analyse de traces laissées dans les chambres à bulles par des particules) vers 1950 s'adaptaient bien à ce type d'images.
- Les images binaires sont un contexte simple permettant une formalisation mathématique des problèmes par des outils tels que la topologie.

Dans le domaine de la vision industrielle (détection de défauts, contrôle de qualité, mesure, ...) on considère souvent l'image binaire comme un passage obligé, suivant en général la phase de segmentation.

Deux catégories d'outils sont alors nécessaires pour d'une part le codage efficace (et éventuellement la compression) et d'autre part pour le traitement (analyse et description des formes) [BM01].

### 6.2 Méthodes de traitement des images binaires

Les traitements les plus connus sur les images binaires sont bien sûr ceux issus de la morphologie mathématique.

Un des traitements les plus importants réalisés sur une image binaire consiste à filtrer et/ou partitionner les objets dans un but de représentation ou de modélisation. Les méthodes de partitionnement, généralement contraintes par les propriétés géométriques des éléments résultants (triangles, polygones, spirales (une seule suite d'arêtes à sommets concaves), carrés, rectangles), fournissent une partition qui n'est pas nécessairement en accord avec la morphologie de la forme. C'est pourquoi on adopte plus volontiers les décompositions par recouvrement qui sont plus souples et aussi plus anthropomorphiques, c'est à dire qui coïncident davantage avec les essais humains de compréhension des formes par décomposition.

La morphologie mathématique est une manière d'aborder ce problème. La théorie des ensembles est l'un des fondements de cette discipline; son espace de travail est l'espace discret.

L'aspect de description de formes est fourni par définition de formes géométriques élémentaires appelées éléments structurants. Ceux-ci définissent des primitives de base dont on pourra confirmer ou infirmer la présence dans la forme étudiée, au moyen d'opérations

ensemblistes. Ceci amène à la construction de procédés de filtrage et de décomposition [BC97] [Se89].

Les principaux opérateurs morphologiques sont :

- L'érosion,
- La dilatation,
- L'ouverture,
- La fermeture,
- La squelettisation.

### 6.2.1 Erosion

Soit  $B$  un élément structurant et  $B_x$  cet élément centré en un pixel  $x$ . L'érosion consiste à poser en chaque pixel  $x$  d'un objet  $X$ , la question : " $B_x$  est-il contenu entièrement dans  $X$  ?".

L'ensemble des positions  $x$  correspondant à une réponse positive forme le nouvel ensemble  $Y$ , appelé **érodé de  $X$  par  $B$** . Cet ensemble satisfait l'équation [BC97]:

$$Y = \{ x \text{ dans } X : B_x \text{ inclus dans } X \}$$

On note l'érosion (figure 5.9.b) par  $E^B$  :

$$Y = E^B(X) \tag{5.5}$$

### 6.2.2 Dilatation

L'opération de dilatation se définit de manière analogue à l'érosion. En prenant le même élément structurant  $B_x$ , on pose pour chaque point  $x$  la question " $B_x$  touche-t-il l'ensemble  $X$ ?" C'est à dire, y a-t-il une intersection non vide entre  $B_x$  et  $X$  ?

L'ensemble des points de l'image correspondant aux réponses positives forme le nouvel ensemble  $X$  des dilatés de  $X$  [BC97].

$$Y = \{ x \text{ dans } I : B_x \div X \neq \emptyset \} \text{ (où } \div \text{ désigne l'opérateur intersection)}$$

On note l'opération de dilatation (figure 5.9.a) par  $DB$  :

$$Y = DB(X) \tag{5.6}$$

### 6.2.3 Ouverture

L'ouverture est l'application de l'opérateur érosion puis de l'opérateur dilatation avec le même élément structurant :

$$OB(X) = DB(EB(X)) \tag{5.7}$$

En général, on ne retrouve pas l'ensemble de départ car une partie de la forme éliminée par l'érosion ne peut être recrée par une dilatation. L'ensemble  $OB(X)$  est plus régulier (moins de détails au niveau du contour) que l'ensemble initial  $X$ . En termes "géographiques" ou morphologiques, on

dit que l'ouverture adoucit les contours, coupe les isthmes étroits, supprime les petites îles et les caps étroits [BC97] (Figure 5.9.c).

#### 6.2.4 Fermeture

La fermeture est l'opération "inverse" de l'ouverture, c'est à dire que l'on applique tout d'abord une dilatation puis une érosion (toujours en gardant le même élément structurant):

$$FB(X) = EB(DB(X)) \quad (5.8)$$

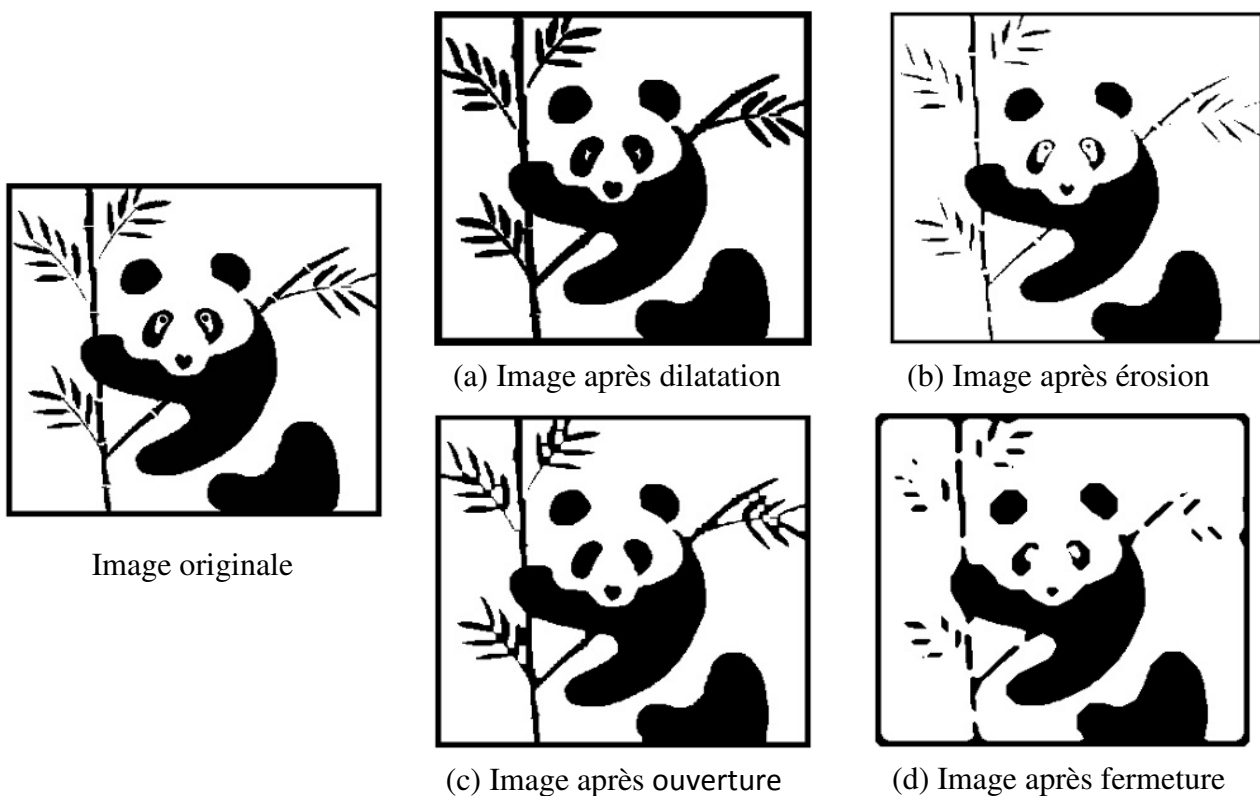
Un ensemble fermé est également moins riche en détails que l'ensemble initial. La transformation par fermeture bouche les canaux étroits, supprime les petits lacs et les golfes étroits.

Les deux opérations d'ouverture et de fermeture sont idempotentes, c'est à dire que le résultat est invariant après itérations :

$$OB(X) = OB(OB(X)) \quad (5.9)$$

$$FB(X) = FB(FB(X)) \quad (5.10)$$

Ces transformations ne sont pas homotopiques, i.e. elles ne préservent pas la connexité car elles peuvent scinder une forme en deux (érosion) ou fusionner deux formes (dilatation). Elles ne présentent donc pas de propriétés topologiques intéressantes. Ce ne sont que des transformations d'aspect, de simplification ou de filtrage [BC97]. La figure 5.9.d illustre cette opération.



**Figure 5.9 :** Application des différents opérateurs morphologiques

### 6.2.5 Squelette

Par contre, le squelette est une représentation d'une forme très utilisée car il conserve les propriétés topologiques de la forme qu'il représente. La notion de squelette est apparue pour l'étude des objets minces. La reconnaissance des caractères en constitue un exemple typique.

L'idée de squelettisation consiste à centrer dans la forme un squelette qui soit significatif de l'élongation et des déformations de celle-ci. Typiquement, le squelette d'un cercle va alors être son centre, celui d'une ellipse son grand axe, ... Malheureusement, le squelette, pour être fidèle à la forme initiale, est très sensible au bruit (petite déformation du contour, présence d'un trou, ...). Le calcul du squelette en morphologie mathématique est obtenu par une transformation de voisinage décrite dans la suite [BC97].

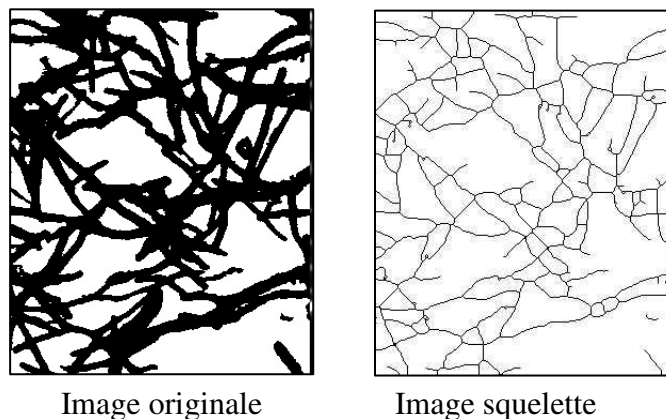


Figure 5.10 : Exemples de squelettisation

## 7. Méthode adoptée

Après avoir comparé les différentes méthodes de détection de visages, une méthode hybride a été retenue, alliant l'approche de détection de visages basée sur la couleur de peau et celle du Template matching [CR00]. Cette solution est composée des étapes suivantes :

1. Détection des pixels représentant la peau à partir d'une image d'entrée.
2. Filtrage des régions de peau.
3. Détection du visage par appariement du gabarit "Template Matching".

Les sous-sections suivantes aborderont donc en détails les différentes étapes de la solution retenue.

### 7.1 Détection de la peau

Le processus de détection de peau peut se réaliser de deux façons différentes soit par simple seuillage selon l'espace de couleur utilisé ou bien par un modèle gaussien de peau préalablement créé. Les choix d'un espace donnée pour la représentation et la détection de la couleur de peau est un problème qui n'a toujours pas été résolu pour cette raison et à titre comparatif nous avons opté

pour basée sur trois des espaces les plus utilisés : HSV, RGB et l'espace YCbCr. La figure 5.11 illustre ce processus.

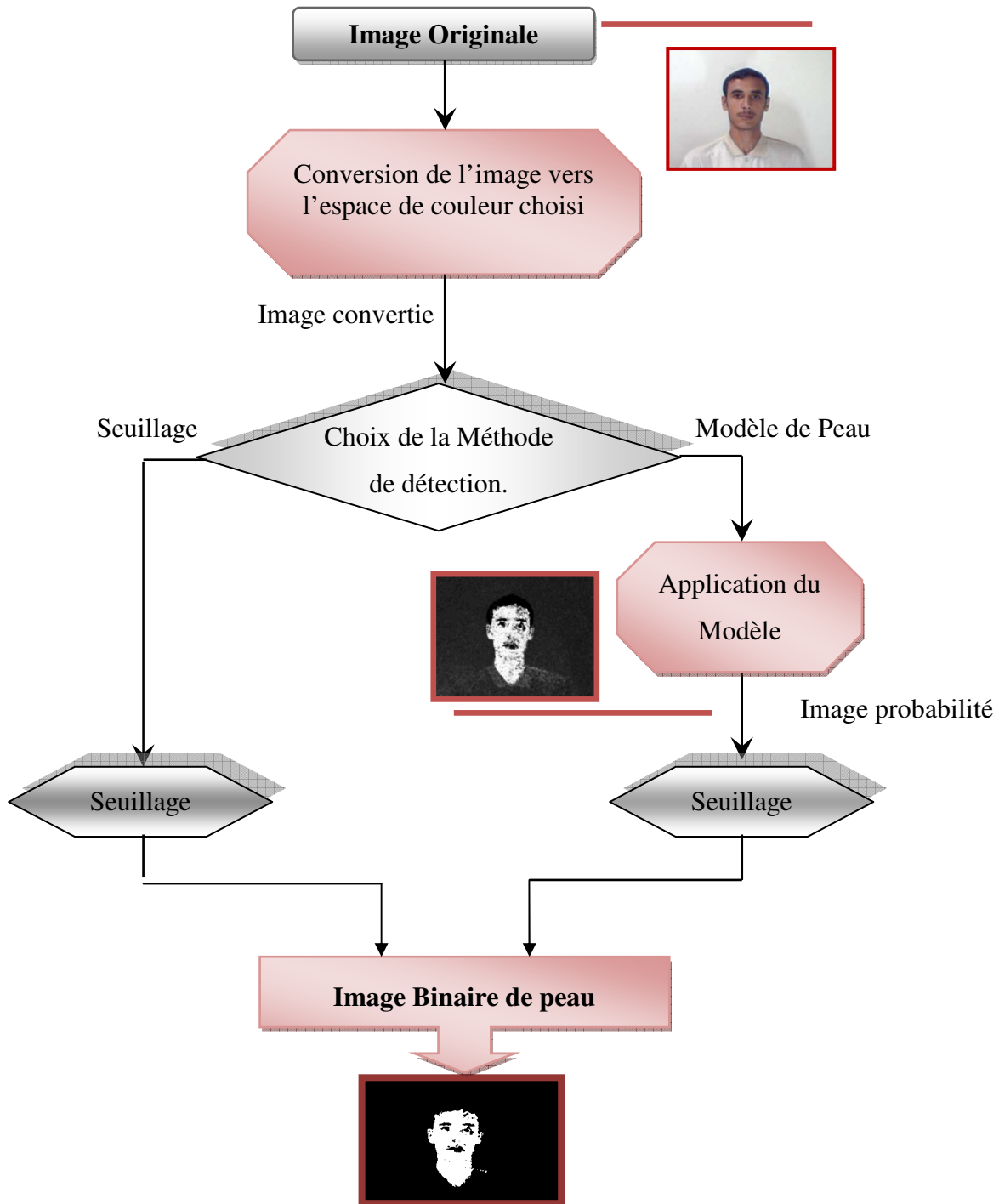


Figure 5.11 : Illustration du processus de détection de peau



### 7.1.1 Détection de peau par seuillage

Les seuils de segmentation par couleur de peau change avec l'espace de couleur choisie. Ces seuils, trouvés dans la littérature scientifique [TP98] [KP03][KN04] [CN99], ont été fixé après une étude approfondie de la distribution de la couleur de peau représenté dans chaque espace de couleur.

#### a. Cas de l'espace HSV

Dans l'espace HSV, l'information d'intensité est représentée par le canal V (value en anglais), pour cette raison, ce canal doit être négligé dans le processus de détection de la peau, on ne considère que les canaux H et S qui représentent l'information chromatique. Donc, on procède d'abord à effectuer une conversion de l'espace RGB vers l'espace HSV. Ensuite, le résultat subi une segmentation par seuillage; les seuils sont les limites supérieures et inférieures de chaque canal (H, S). Les intervalles optimaux pour le processus de détection de la peau sont décrits ci-dessous [TP98]

$$\begin{aligned} 0 < H < 27^\circ \\ 0.2 < S < 0.8 \end{aligned} \quad (5.11)$$

#### b. Cas de l'espace RGB

Une autre méthode a été implémentée afin de segmenter les régions de peau des régions non peau est celle de l'espace RGB, cette méthode ne nécessite aucun modèle de peau et aucune transformation des couleurs, elle vérifie simplement une suite des contraintes pour décider si un triplet de couleur (R, G, B) représente une couleur de peau ou non, ces contraintes sont divisées en deux selon l'état de l'éclairage, fort (jour) ou faible (nuit) [KP03].

##### *Eclairage fort :*

$$\begin{aligned} (R > 95) \text{ ET } (G > 40) \text{ ET } (B > 20) \text{ ET} \\ (\text{MAX}(R, G, B) - \text{MIN}(R, G, B) > 15) \text{ ET} \\ (\text{ABS}(R-G) > 15) \text{ ET } (R > G) \text{ ET } (R > B) \end{aligned} \quad (5.12)$$

##### *Eclairage faible :*

$$\begin{aligned} ((R > B) \text{ ET } (G > B)) \text{ OU} \\ ((R > 220) \text{ ET } (G > 210) \text{ ET } (B > 170) \text{ ET} \\ (\text{ABS}(R-G) \leq 15)) \end{aligned} \quad (5.13)$$

Cette méthode est très pratique en termes de temps d'exécution, et donne des résultats acceptables.

### c. Cas de l'espace YCbCr

Dans l'espace YCbCr, l'information d'intensité est représentée par le canal Y, pour cette raison, ce canal doit être négligé dans le processus de détection de la peau, on ne considère que les canaux Cb et Cr qui représentent l'information chromatique.

L'image est convertie en YCbCr, puis elle subit une segmentation par seuil, une limite supérieure et inférieure pour les canaux Cb et Cr est fixé [KN04] [CN99]

$$77 \leq C_b \leq 135 \quad \text{et} \quad 133 \leq C_r \leq 180 \quad (5.14)$$

#### 7.1.2 Détection de peau par modèle Gaussien

L'inspiration pour utiliser l'analyse de couleur de peau pour la classification initiale d'une image dans des régions probables de visages et de non visages provient d'un certain nombre de caractéristiques simples mais puissantes de couleur de peau tel que son insensibilité à la rotation et à la translation.

La Modélisation de la distribution de la couleur de peau peut faire l'objet de modèles paramétriques ou non paramétriques. En cas de modélisation paramétrique, un modèle statistique prédéfini est choisi pour modéliser la distribution de peau.

Dans le cas de la modélisation non paramétrique, une analyse d'histogramme des données d'apprentissage suit une fonction de probabilité décrivant la distribution de la couleur de peau.

##### a. Création du modèle de couleur de peau

La distribution des couleurs de peau des différentes personnes est groupée dans un petit secteur de l'espace de couleur normalisé. Les couleurs de peau sont très étroites, mais elles diffèrent principalement dans les intensités. A partir de cette conclusion, nous avons mis en œuvre un modèle de couleur de peau dans chaque espace de couleur chromatique.

La création du modèle débute par la préparation des échantillons de couleur de peau provenant de plusieurs images en couleur. Ces échantillons de peau ont été filtrés à l'aide d'un filtre passe-bas (Filtre Médian) pour réduire l'effet du bruit.

La figure 5.12 illustre un ensemble d'échantillons de peau (issue de, mains, fronts, cou, joue...etc) tirés de la base utilisée.

La base de segments de peau utilisée contient 168 segments de peau, extraits à partir de photos de personnes de différentes couleurs de peau, la totalité de ces segments contient 136745 pixels. Cette base sera utilisée pour créer le modèle de couleur de peau.

**Début :**

1. Convertir l'image originale vers l'espace de couleur choisi pour détecter les régions peau.
2. Créer un modèle de peau pour chaque espace de couleur, à partir de la base des imagerie peau (figure 5.12)
3. Chacun de ces échantillons était alors filtré par un filtre passe bas afin de réduire le bruit dans chaque image.
4. Après avoir choisi les canaux C1 et C2 de l'espace de couleur, ensuite le calcul de la moyenne de l'espace de couleur utilisé :

$$\text{moyenne: } m = E\{x\} \quad \text{ou} \quad x = (C1, C2)^T \quad (5.15)$$

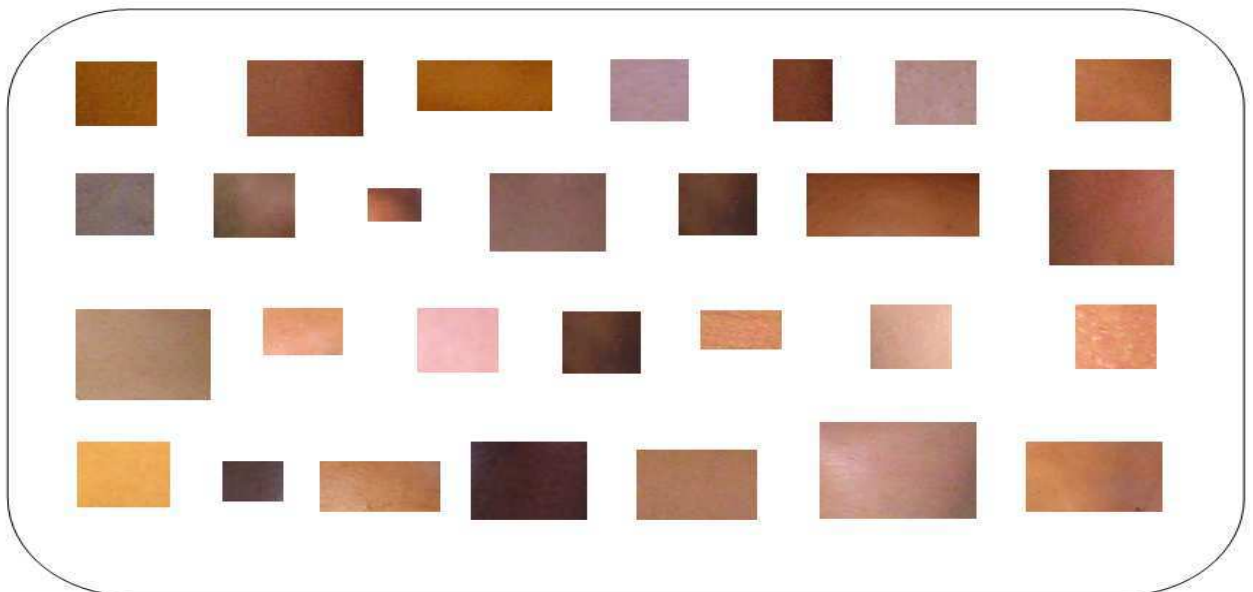
Avec **C1** et **C2** les canaux porteur d'information de chrominance pour chaque espace de couleur choisi tel que :

- YCbCr : C1=Cb et C2=Cr
- RGBN : C1=R et C2=G
- HSV : C1=H et C2=S

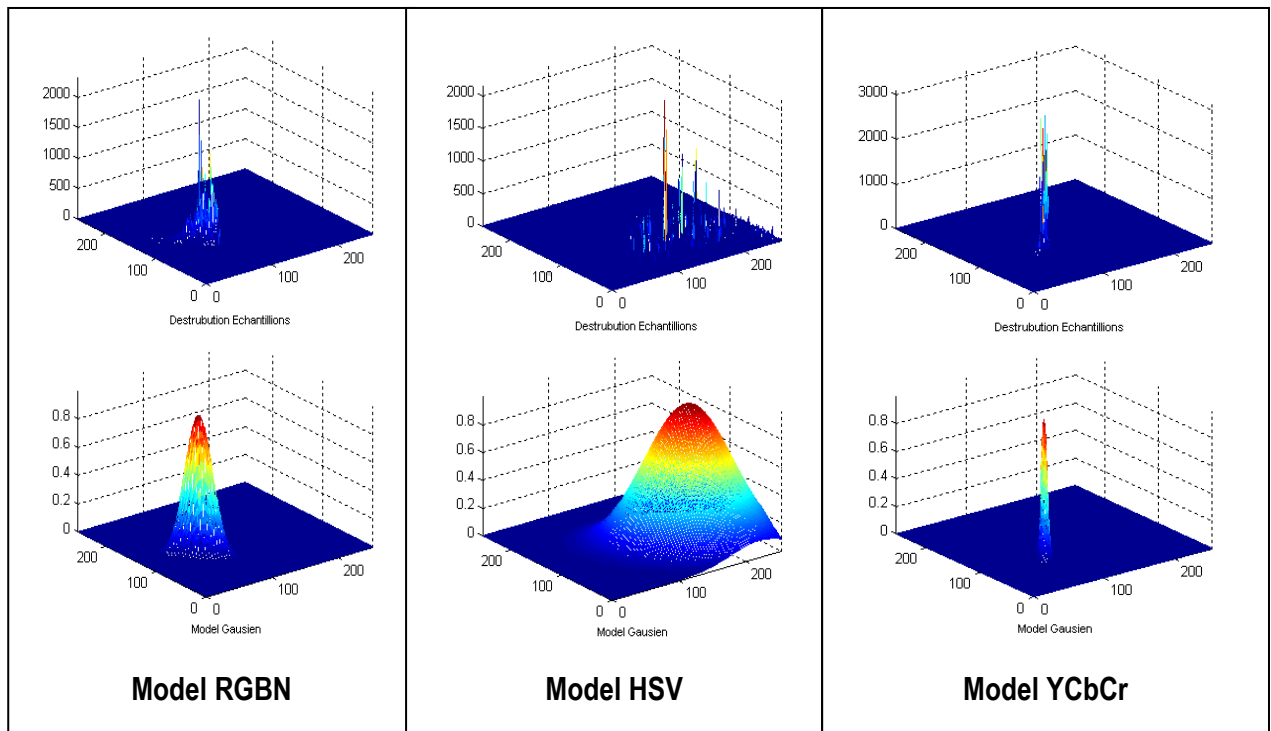
5. covariance :  $C = E\{(x - m)(x - m)^T\}$  (5.16)

6. En utilisant les valeurs de la moyenne et de covariance, le modèle de couleur de peau peut être adapté à un modèle gaussien par :

$$P(C1, C2) = \exp[0.5 * (x - m)^T C^{-1} (x - m)] \quad \text{ou} \quad x = (C1, C2)^T \quad (5.17)$$

**FIN.**

**Figure 5.12 :** Quelques exemples extraient de la base d'échantillons de peau



**Figure 5.13 :** Distribution (en haut) et model gaussien (en bas) de peau selon les espaces RGBN, HSV et YCbCr.

**b. Application du Modèle :**

Avec ce modèle gaussien, nous pouvons maintenant obtenir la probabilité de la peau pour n'importe quel pixel d'une image. Soit un pixel ayant une valeur chromatique (C1, C2), la probabilité pour que ce pixel soit une peau, dans l'espace choisie, peut alors être calculée par l'équation (5.17)

Par conséquent, ce modèle peut transformer une image couleur en une image en niveau de gris tel que la valeur de chaque pixel correspond à sa probabilité d'appartenance à la peau.

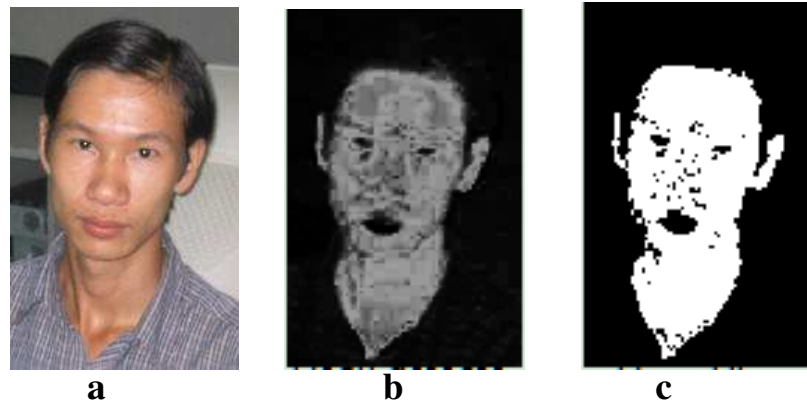
Avec un seuillage approprié, les images en niveau de gris peuvent alors être encore transformées en des images binaires montrant les régions de peau et les régions non peau, puisque les régions de peau sont plus lumineuses que les autres parties de l'image, ces régions peuvent être segmentées du reste par un seuillage. On produit alors une image binaire sur laquelle les « 1 » représentent les pixels couleur de peau et les « 0 » pour les autres pixels. Un seuil optimal est pris dans l'intervalle 0.35 à 0.45 [CR00] [FL04].

$$\begin{aligned}
 x < \text{Seuil} & \Rightarrow x=0 \\
 x \geq \text{Seuil} & \Rightarrow x=1
 \end{aligned}
 \tag{5.18}$$

Tel que x représente les valeurs des Pixels de l'image de probabilité.

**Remarque**

Il est important de noter que les régions détectées ne correspondent pas toujours à la peau, mais il se peu que sa couleur soit semblable à celle de la peau.

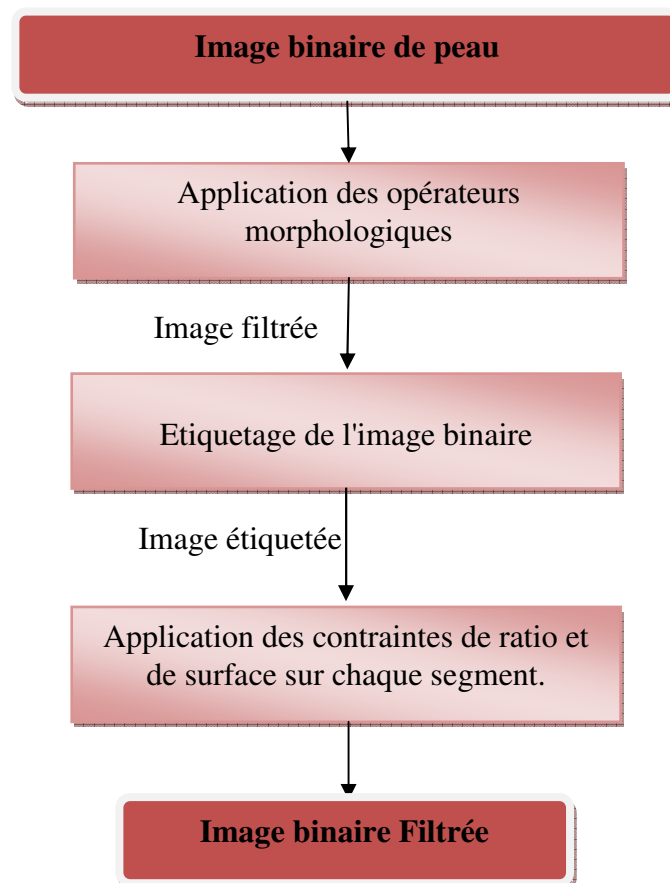


**Figure 5.14:** Application du model Gaussien a) Image Originale [Fr].  
b) Image Probabilité "**likelihood**". c) Image Binaire de peau.

## 7.2 Filtrage et séparation des régions

Il est clair que, quelque soit l'espace de couleur utilisé, les régions de peau détectées ne contiennent pas toujours des visages. Ils peuvent correspondre à toute autre partie exposée du corps, ou aux objets ayant des couleurs semblables à ceux de la peau. Par conséquent, on fait appel à une étape de filtrage des régions pour éliminer les segments qui ne vérifie pas certaines contraintes.

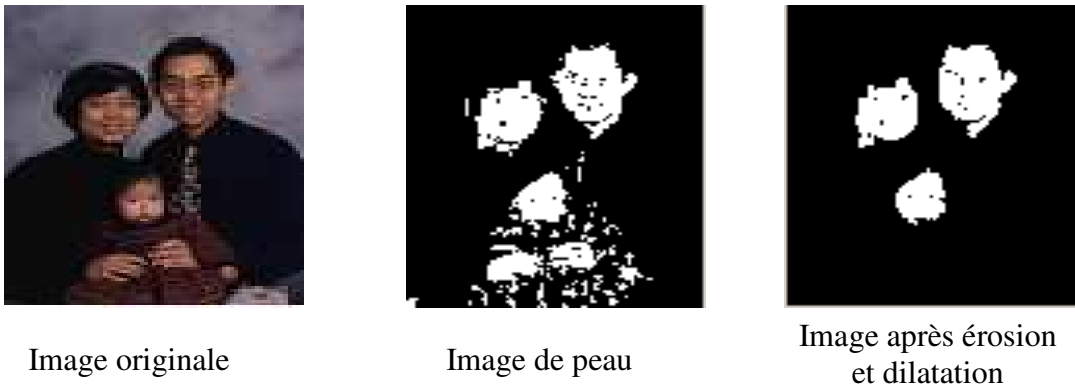
Pour arriver à une vitesse d'exécution acceptable nous avons choisie de traiter chaque segment de peau seul, pour cela nous avons appliqué un filtrage des régions peau puis une séparation des différents segments détectés. Le processus de cette étape est illustré par la figure 5.15



**Figure 5.15 :** Illustration du processus de filtrage des régions de peau

### 7.2.1 Application des opérateurs morphologiques

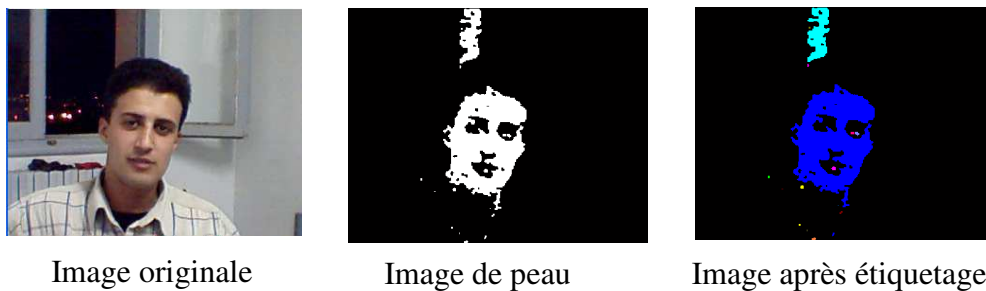
L'application des opérateurs morphologique, à savoir l'érosion et la dilatation, sur les zones peau précédemment détectés nous permettra de réduire les défauts de ces zones tel que : l'égalisation des bords, le remplissage des trous et l'élimination des petits segments.



**Figure 5.16:** Exemples d'application des opérateurs morphologiques (Érosion et dilatation)

### 7.2.2 Etiquetage des régions

Nous devons d'abord déterminer le nombre de régions de peau dans l'image, en associant à chaque région une valeur entière appelée *étiquette*. Afin de déterminer l'étiquette d'un Pixel, on utilise pour cela un voisinage d'ordre 8 [CR00]. Le nombre de régions dans l'image segmentée sera donc la valeur de l'étiquette maximale (figure 5.17).



**Figure 5.17 :** Exemple d'étiquetage des régions peau  
Base visage LFDB [La]

### 7.2.3 Extraction des caractéristiques des régions [CR00]

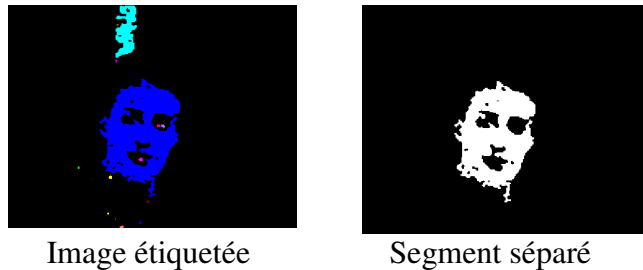
Cette étape consiste à séparer les régions et en extraire les différentes informations tels que :

- La position,
- Les dimensions,
- La surface,
- Le rapport entre la hauteur et la largeur "Ratio",
- Le centre de gravité,

- La rotation.

### a) Séparation des régions

Afin de séparer une région particulière, nous balayons l'image segmentée à la recherche des pixels qui correspondent à l'étiquette de la région concernée. Le résultat sera une image binaire qui ne contient que la région en question. La figure suivante illustre cette opération :



**Figure 5.18 :** Exemple de séparation d'une région

### b) Estimation de la position du segment

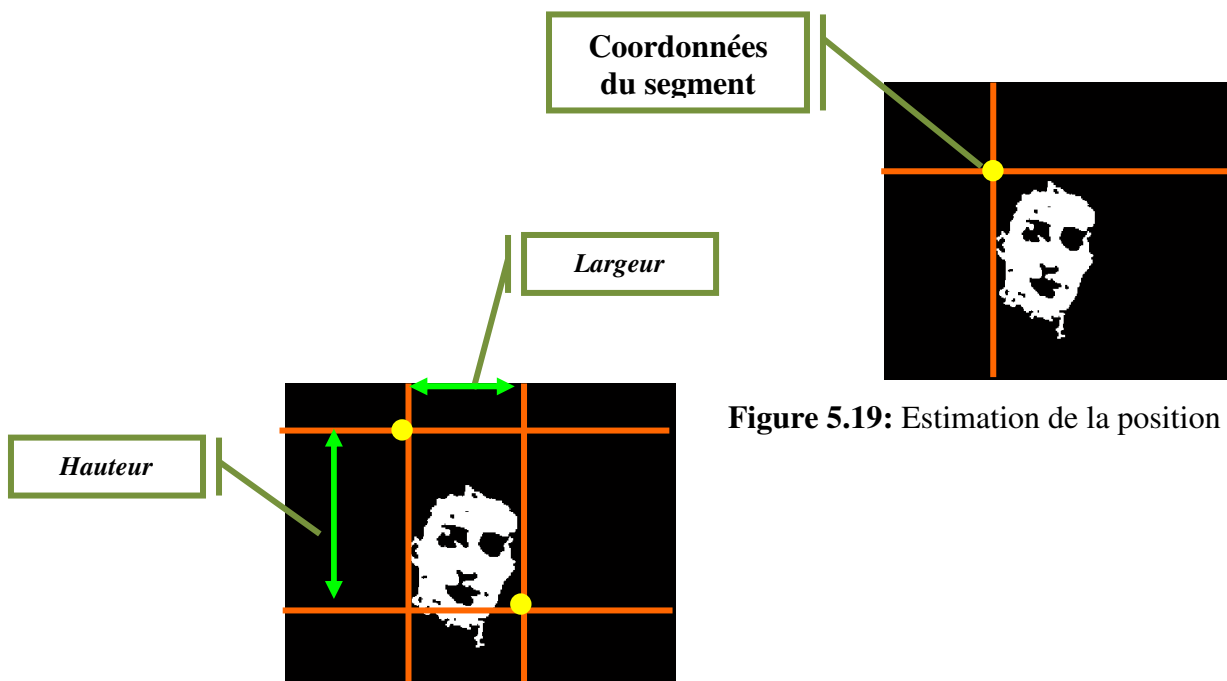
Il consiste à déterminer les coordonnées  $(X, Y)$  du coin supérieur gauche du segment par rapport à l'image segmentée (figure 5.19).

### c) Calcul des dimensions du segment

En désigne par dimension, la largeur et la hauteur du segment, cela s'effectue en calculant la différence entre les coordonnées  $(X1, Y1)$  du coin inférieur droite du segment et ceux  $(X2, Y2)$  de son coin supérieur gauche (figure 5.20), comme le montre la formule suivante:

$$\text{Largeur} = X1 - X2$$

$$\text{Hauteur} = Y1 - Y2$$



**Figure 5.19:** Estimation de la position d'un segment

**Figure 5.20 :** Calcul des dimensions du segment

**d) Calcul de la surface du segment**

Il consiste à compter le nombre de pixels appartenant à cette région.

$$Surface = \sum_i^M \sum_j^N I(i, j) \quad (5.19)$$

**e) Calcul du ratio segment**

Il consiste à calculer le rapport entre la largeur et la hauteur de la région. Il est défini par la formule suivante :

$$Ratio = \text{Hauteur} / \text{Largeur}$$

**f) Calcul du centre de gravité du segment**

Pour étudier une région, nous devons déterminer son centre. Une manière efficace est de calculer le centre de gravité avec la formule suivante :

$$\bar{x} = \frac{1}{A} \sum_{i=1}^n \sum_{j=1}^m jB[i, j] \quad (5.20)$$

$$\bar{y} = \frac{1}{A} \sum_{i=1}^n \sum_{j=1}^m iB[i, j] \quad (5.21)$$

Où B est la matrice [n x m] qui représente la région et A est la surface de cette région.

**g) Calcul de la rotation du segment**

Habituellement les visages sont orientés verticalement, cependant on peut avoir une petite inclination  $\theta$ , qui peut affecter les résultats de l'étape suivante, ce qui nous amène à calculer cette inclinaison pour rendre droite cette région, par la formule suivante [CR00]:

$$\theta = \frac{1}{2} \arctan \frac{b}{a-c} \quad (5.22)$$

$$\text{Où} \quad a = \sum_{i=1}^n \sum_{j=1}^m (x'_{ij})^2 B[i, j]$$

$$b = 2 \sum_{i=1}^n \sum_{j=1}^m x'_{ij} y'_{ij} B[i, j] \quad (5.23)$$

$$c = \sum_{i=1}^n \sum_{j=1}^m (y'_{ij})^2 B[i, j]$$

$$\text{Et} \quad \begin{aligned} x' &= x - \bar{x} \\ y' &= y - \bar{y} \end{aligned} \quad (5.24)$$

**7.2.4 Sélection des régions**

La sélection des régions consiste à balayer toutes les régions détectées, en vérifiant un ensemble de contraintes. Ces contraintes sont aux nombres de deux, l'une concerne le ratio de la région et l'autre concerne sa surface. Leur vérification est précédée d'une étape préliminaire de

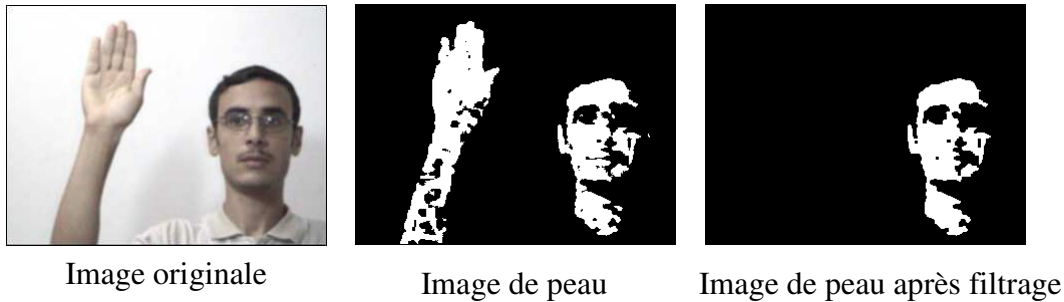


traitement de l'image binaire par les opérateurs morphologiques, et qui vise essentiellement à éliminer le bruit de l'image binaire.

#### a) Vérification de la contrainte du ratio

Le ratio est le rapport entre l'hauteur et la largeur d'un segment, habituellement, dans le cas des visages, se rapport ne dépasse pas un certain intervalle, donc tous les segments qui ont un ratio qui ne fait pas partie à cet intervalle seront éliminés. L'intervalle convenable utilisé est le suivant :

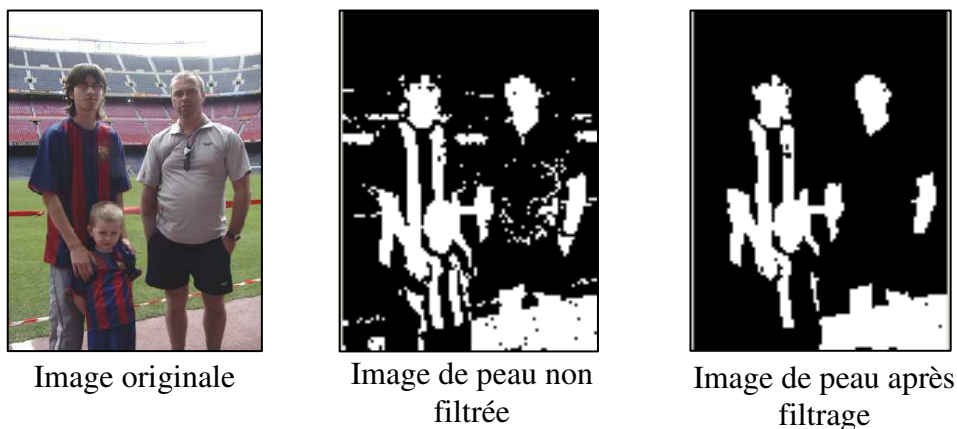
$$0.8 < \text{ratio} < 1.8 \quad (5.25)$$



**Figure 5.21 :** Exemple de filtrage avec la contrainte de ratio Base visage LFDB [La]

#### b) Vérification de la contrainte de surface

Pour augmenter les performances de la méthode choisie, nous avons introduit une autre contrainte pour éliminer les segments de tailles négligeables, dont il est inutile de traiter. Pour cette raison, nous fixons un seuil minimal (dans notre cas 50) et tous les segments qui ont une surface inférieure à cette valeur seront supprimés. Une autre contrainte qui concerne la surface est le rapport entre la surface en question et la surface dont le segment est le plus grand, qui doit être inférieur à un certain seuil (dans notre cas 20).



**Figure 5.22 :** Exemple de filtrage avec les contraintes de la surface

### 7.3 Template Matching

L'utilisation de la méthode de "Template matching" se justifie principalement par la contrainte de temps réel qu'il faut appliquer au programme : il est nécessaire d'implémenter une

fonction rapide, afin d'éviter que le système ne mette trop de temps pour réagir. Un autre avantage de cette méthode est la simplicité de l'implémentation.

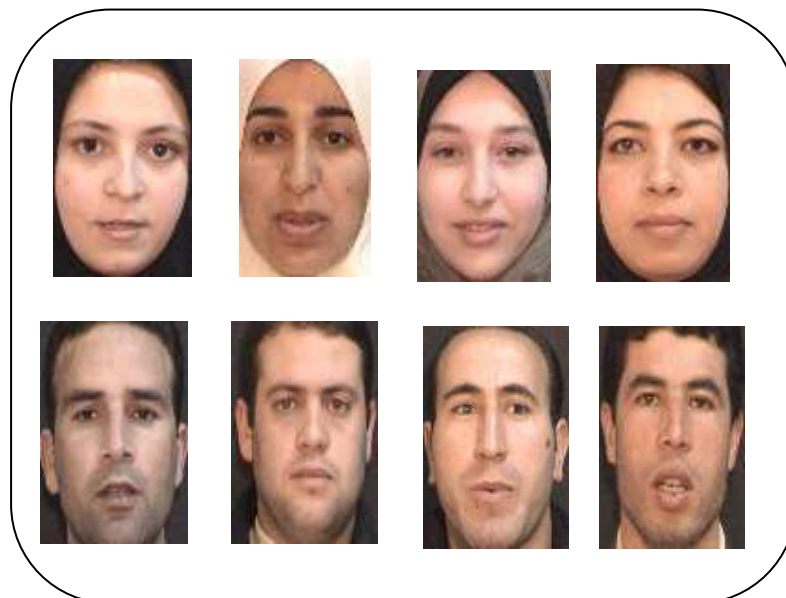
Le principe de cette méthode est le suivant (figure 5.25) : il s'agit, étant donné une image quelconque et un modèle prédéfini, de déterminer le degré d'appartenance (similarité) de ce modèle à cette image [CR00].

### 7.3.1 Création d'un modèle de visages

La caractéristique la plus importante de cette méthode est qu'elle utilise les visages pour décider si une région de peau correspond à un visage ou non. Notre modèle a été créé en effectuant un calcul de moyenne à partir de plusieurs images en niveaux de gris, qui ne contiennent que des visages (figure 5.23). La base (B-LAIG) [La] que nous avons utilisé pour créer le modèle de visages a été réalisée au LAIG (*Laboratoire d'Automatique et Informatique de Guelma*) [Bo05], elle contient 1823 visages pris dans différents profils et dans différentes conditions d'éclairage (figure 5.24).



**Figure 5.23** : Exemples de modèles, (c) est le modèle utilisé



**Figure 5.24** : Quelques exemples extraits de la base de modèles de visages B-LAIG [La]

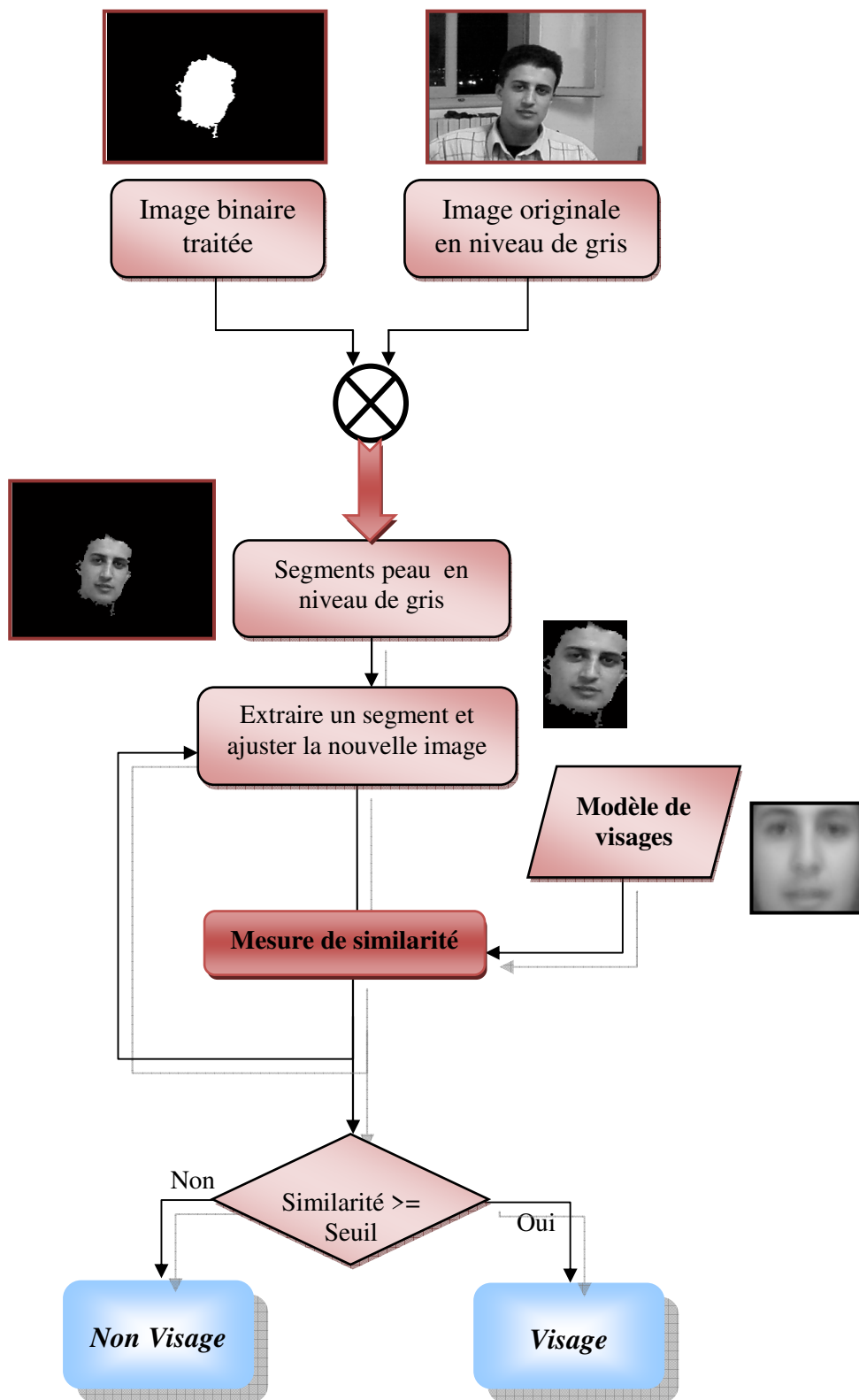


Figure 5.25 : Schéma illustratif de l'application du "Template matching"

### 7.3.2 Mesures de similarité

Il existe plusieurs manières de mesurer la similarité entre deux images.

- La première mesure appliquée est une simple différence de normes, données par les deux formules suivantes (pour les normes L1 et L2).

$$\text{Norme L1 : } \sum_{(i,j)} |a_{(i,j)} - b_{(i,j)}| \quad (5.26)$$

$$\text{Norme L2 : } \sqrt{\sum_{(i,j)} a_{(i,j)}^2 - b_{(i,j)}^2} \quad (5.27)$$

Où les  $a_{(i,j)}$  et  $b_{(i,j)}$  sont la valeur du pixel (i,j) dans les deux images à comparer. Il ressort donc que plus ces valeurs sont petites, plus les images sont proches [He02].

- La deuxième mesure calcule la cross-correlation, donnée par la formule suivante :

$$R_x(r,c) = \sum_{j=1}^{tplRows-1} \sum_{i=1}^{tplCols-1} t(i,j) x\left(r+j-\frac{tplRows}{2}, c+i-\frac{tplCols}{2}\right) \quad (5.28)$$

Cette mesure est un peu plus complexe que la précédente, puisqu'elle n'apporte pas des résultats directs. Effectivement, le résultat est en fait une image, et la similarité peut se définir de la manière suivante : plus la probabilité d'apparition du Template est grande, plus grande est l'intensité de l'image résultante [He02].

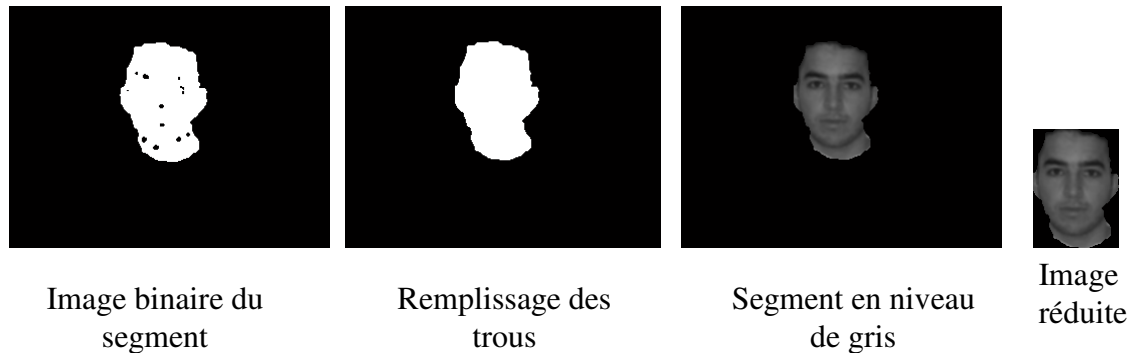
- Finalement, une troisième technique a été appliquée aux images binaires : Il s'agit d'un OU-Exclusif suivi d'une négation ( $\overline{XOR}$ ) sur chaque Pixel correspondant aux deux images. De cette manière les Pixels qui ont la même valeur seront blancs dans l'image résultante et noire aux endroits où ils sont différents [He02].

### 7.4 Mise en œuvre du modèle

Pour effectuer l'appariement avec un segment détecté en utilisant le modèle du visage créé, nous passons par les étapes suivantes :

- Extraction de l'image niveau du gris : d'abord, on doit extraire l'image binaire du segment, et nous remplissons les trous existants. Nous multiplions l'image résultante par l'image originale en niveau de gris, le résultat sera une image en niveau de gris qui représente le segment en question. En fin, nous réduisons la taille de l'image pour qu'elle ne contienne que les Pixels du segment. La figure 5.26 illustre ce que nous venons d'expliquer.
- Normalisation du segment : cette procédure de normalisation porte sur l'image en niveau de gris du segment, afin de rendre possible et d'améliorer l'opération d'appariement. La normalisation consiste à : ajuster l'inclinaison du segment selon son orientation, rendre les dimensions du segment égales à celles du modèle.

- Maintenant, le segment est prêt à être comparé au modèle, cette comparaison s'effectue en balayant la surface du segment et en calculant la cross-corrélation entre le modèle et la région sélectionnée du segment normalisé. Le résultat de cette procédure est la valeur maximale calculée ainsi que la position de la région correspondante. La valeur est comprise dans l'intervalle des valeurs réelles :  $[0.0, 1.0]$ .
- Décision : en fixant un seuil de similarité, nous pouvons juger un segment de peau comme étant "visage" ou "non".

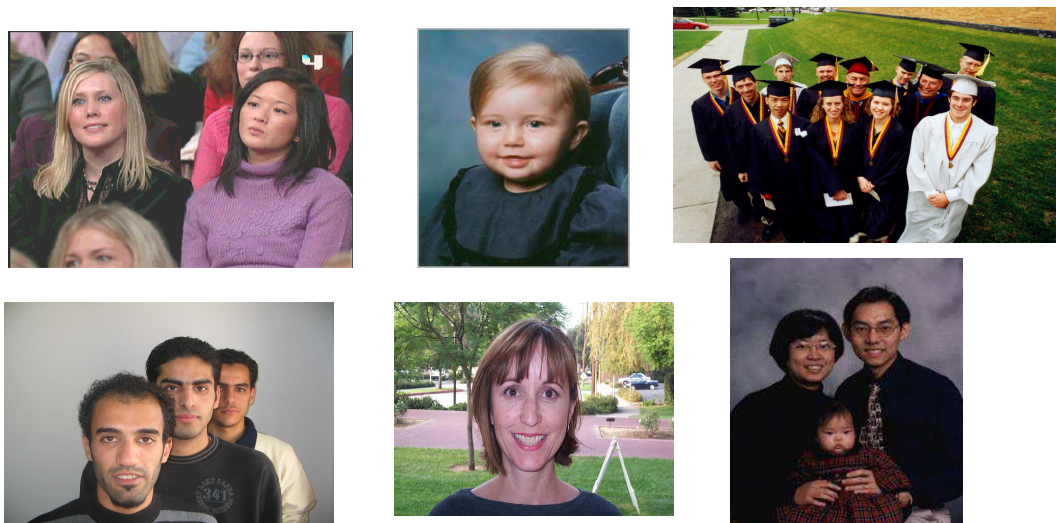


**Figure 5.26:** Extraction du segment visage

## 8. Résultats et interprétations

Afin d'évaluer les performances du système de détection de visages nous avons effectué deux types de tests : le premier pour la détection de peau et le second pour la détection de visages.

Les tests ont été effectués sur un ensemble d'images que nous avons nommé "base test" constituée essentiellement des bases : BAO [Fr], CALTECH [Fr] et LFDB [La]. Rappelons qu'il n'y a aucune condition d'éclairage, orientation, d'échelle ou autre sur les images utilisées.



**Figure 5.27 :** Extrait de la base test utilisée

### 8.1 Test de Détection de couleur de peau

Sur notre système, la détection de couleur de peau peut être faite sur plusieurs espaces de couleur, pour cela et afin d'en choisir le plus représentatif de la couleur de peau nous avons choisi de faire une comparaison entre la détection avec ces espaces en utilisons la détection par seuillage et par model de couleur de peau.

Les performances sont exprimées par deux paramètres :

$$\text{Taux de Détections de Peau (TDP)} = \frac{\text{Nbr de segment peau détectés}}{\text{Nbr de segment peau réels}}$$

$$\text{Taux de Fausses Détections de Peau (TFDP)} = \frac{\text{Nbr de fausses détections}}{\text{Nbr de détections}}$$

Il est clair que le choix du meilleur espace de représentation repose sur les deux paramètres précédant à savoir un taux de détection maximal et un taux de fausse alarme minimal.

Les résultats obtenus sont reportés sur le tableau 5.1 sur lequel nous remarquons que dans le cas de la détection de peau par seuillage le meilleur taux de détection a été obtenu dans la représentation de l'espace RGB (94 %) en contre partie un taux de fausses détection le plus élevé (31 %) d'un coté le meilleur taux de fausses détection a été obtenu dans la représentation de l'espace YCbCr (15 %) face à un taux de détection acceptable (90 %).

Dans le cas de la détection par modèle de peau gaussien avec l'espace YCbCr nous avons obtenus les meilleurs taux de détections (96 %) et de fausses alarmes (14 %).

D'après ces derniers résultats nous pouvons conclure que le meilleur espace pour la représentation de la couleur chair est l'espace YCbCr et la meilleure méthode (parmi les deux testées) de la détection de peau est celle basée sur le modèle gaussien de peau.

Méthode de détection	Seuillage			Model de couleur de peau Gaussien			
	Espace de couleur	YCbCr	RGB	HSV	RGBN	HSV	YCbCr
<b>TDP</b>		90 %	94 %	79 %	96 %	86 %	96 %
<b>TFDP</b>		15 %	31 %	24 %	28 %	19 %	14 %

**Tableau 5.1** : Résultats de détection de peau (Tests réalisés sur 100 images)

## 8.2 Test de détection de visages

Les résultats de la détection de peau nous ont permis de fixer l'espace de représentation et la méthode de détection de peau adoptée pour les tests de la détection de visages.

Les performances sont exprimées par deux valeurs : le taux de détection et le taux de fausses détections.

$$\text{Taux de Détections de visages (TDV)} = \frac{\text{Nbr de visages détectés}}{\text{Nbr de visages réels}}$$

$$\text{Taux de Fausses Détections de Visages (TFDV)} = \frac{\text{Nbr de fausses détections}}{\text{Nbr de détections}}$$

Les performances du système de détection de visages ont été testées sur 100 images respectivement de chacune des bases : BAO (un seul visage), CALTECH, ainsi qu'une troisième base (conçu par nos soins) constituée d'images à plusieurs visages dans différentes positions et éclairages que nous avons nommé LFDB (LAIG Face DataBase) [La] (figure 5.29), les taux obtenus sont illustré sur le tableau 5.2 et le graphe de la figure 5.28. Où nous pouvons constater que les performances du système sont considérables sur les bases BAO et LFDB soit un TDV max de 86 % et un TFDV min de 10 % surtout vu la complexité de l'environnement des d'images utilisées pour le test (figure 5.27), ceci est dû d'une part au choix de l'espace de représentation et la méthode de détection de peau adoptée et d'autre part au étapes de filtrages qui nous ont permis d'éliminer les segments superflus et de minimiser l'espace de recherche de visages par le « Template Matching ». Concernant la base CALTECH les échecs du système ont été signalés surtout sur les images ayant un mauvais contraste.

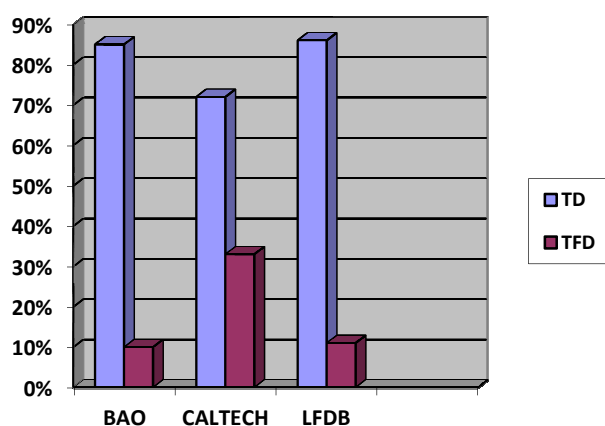


Figure 5.28 : Performance du système de détection de visages

Taux	TDV	TFDV
Base		
BAO (One Face)	85%	10%
CALTECH	72%	33%
LFDB	86%	11%

Tableau 5.2 : Résultats de détection de visages



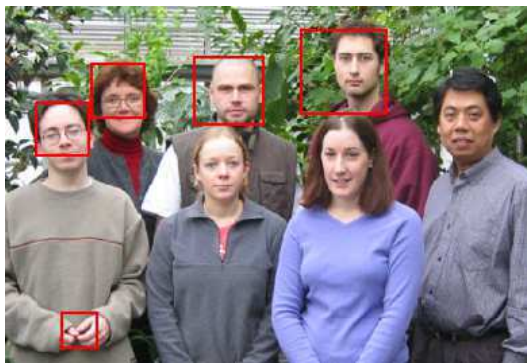
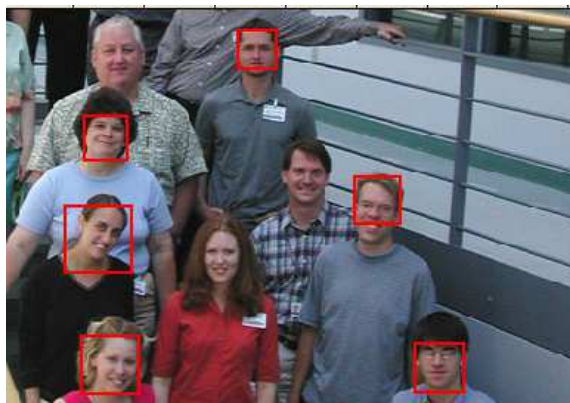
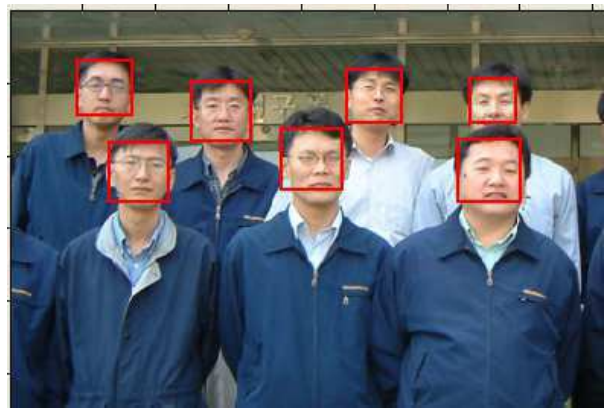
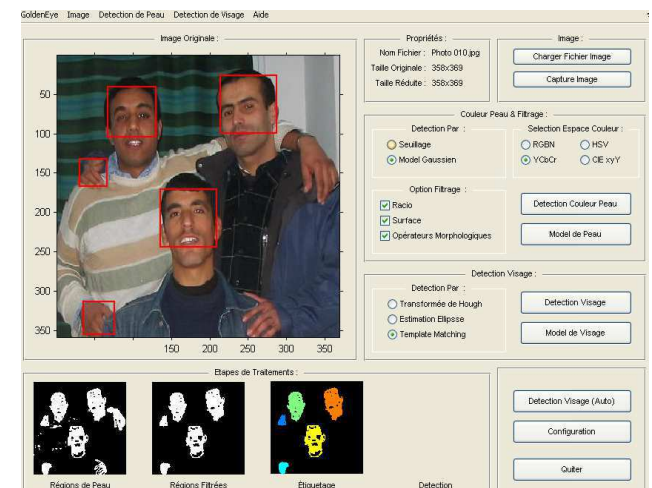
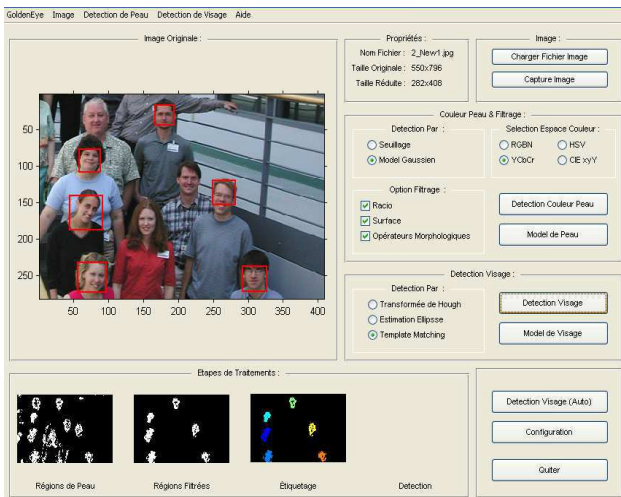
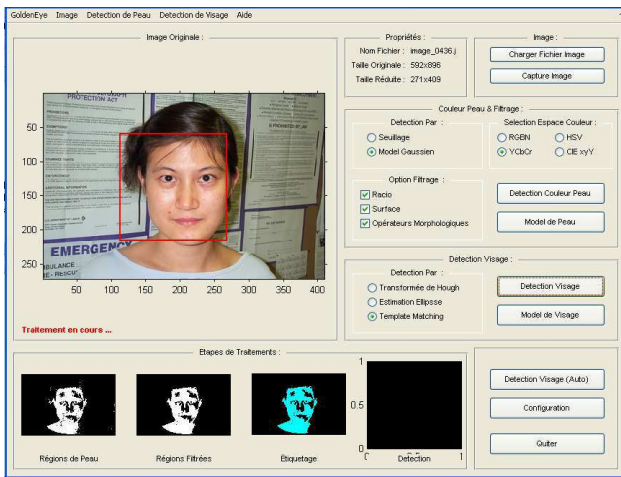
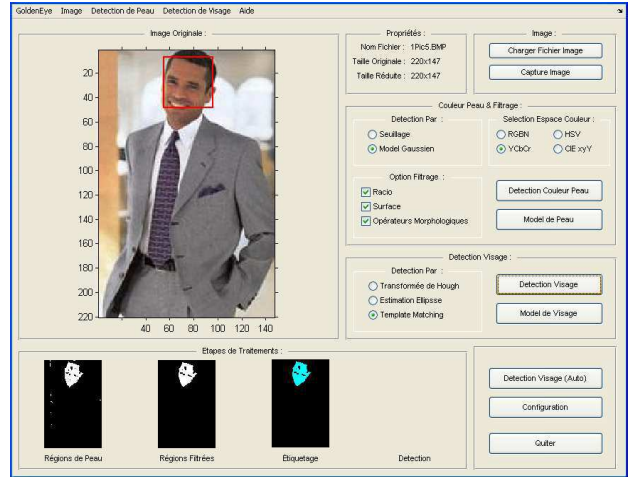
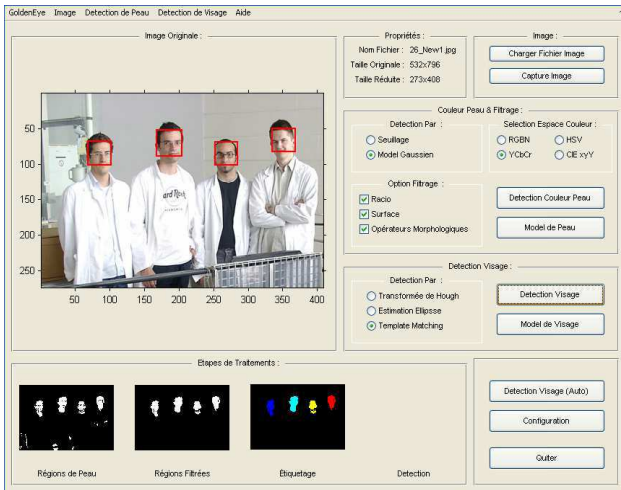


Figure 5.29 : Illustration de quelques résultats de détection de visages





**Figure 5.30 :** Démonstration de quelques résultats sur notre interface utilisateur avec (en bas de l'interface) les trois fenêtres successives de la détection, le filtrage et l'étiquetage des régions peau.

### 8.3 Discussion des résultats

D'après les tests que nous avons effectués, nous pouvons constater ce qui suit :

1. Les détections qui ne correspondent pas à un visage (fausses détections) proviennent de deux sources essentielles :
  - a. Erreur de détection de peau : dans ce cas, les objets qui ont une couleur similaire à celle de la peau sont considérés comme étant une peau, si ces segments s'échappent à l'étape de filtrage et le Template matching ils seront considérés comme étant des visages.
  - b. Erreur de filtrage et de Template matching : dans ce cas, le segment peut correspondre à une peau, mais n'est pas un visage, cependant il peut être signalé comme visage.
2. Les cas de "non détection" de visages existants peuvent correspondre aux situations :
  - a. La détection de peau a échoué, à cause des conditions d'éclairage, ou de la qualité de l'image acquise, ou bien à cause d'un mauvais choix de l'espace de couleur.
  - b. Les visages sont détectés comme peau, mais ils ont été supprimés par la suite par le filtrage, car ils sont trop petits.
  - c. Une collision de peau : dans ce cas, les régions sont assemblées dans un seul segment de peau, et il se peut qu'il soit supprimé car il ne vérifie pas les conditions exigées tel que le ratio.
  - d. Un visage n'a pas été détecté car il se trouve dans une région de peau très large, il se peut que cette région ne vérifie pas les conditions, et sera, de ce fait, supprimé au filtrage, ou que le Template matching ne donne pas un bon résultat, car il portera sa recherche du visage sur toute la région.

### 9. Conclusion

La détection de visages demeure l'un des problèmes les plus complexes, malgré les recherches actives actuelles. Il y a de nombreuses conditions réelles, difficiles à modéliser et prévoir, qui limitent les meilleurs systèmes actuels.

Dans le présent chapitre, nous avons présenté une méthode de détection de visages hybride alliant la méthode de détection de peau et celle du Template matching.

Les résultats de détection de peau sont fiables dans la plus part des cas. Cependant, dès que l'image d'entrée devient plus complexe le taux de mauvaises localisations augmente. Cette phase est suivie d'une phase de filtrage des régions détectées pour supprimer tout segment inutile. Après la création d'un modèle de visage, tous les segments issus de la phase de détection de peau sont comparés avec ce modèle pour déterminer leurs taux d'appartenance à une classe de visages.

CONCLUSION  
GÉNÉRALE

# CONCLUSION GÉNÉRALE

---

La détection de visages et la reconnaissance d'individus demeurent des problèmes complexes, malgré les recherches actives actuelles. Il y a de nombreuses conditions réelles, difficiles à modéliser et prévoir, qui limitent les meilleurs systèmes actuels.

Le travail que nous présentons dans cet ouvrage est réparti en deux grandes parties :

Dans la première partie, nous proposons un système de classification et de reconnaissance de visages dont les tests sont réalisés sur la base de visages créée au LAIG (Laboratoire d'Automatique et d'Informatique de Guelma) B-LAIG.

La richesse en information et la complexité des objets manipulés (Visages) nous a mené à exploiter les informations issues des visages, pour leur classification, de deux manières différentes : la première consiste à extraire les paramètres géométriques définissant les traits faciaux, alors que la seconde se contente de prendre l'image du visage entière (traitement de tous les pixels du visage). Comme outil de classification, nous avons utilisé en premier lieu le PMC ; connu par ses performances dans ce domaine. La seconde méthode utilisée est l'Analyse en Composantes Principales (ACP), certes c'est une méthode des plus anciennes en classification de données mais elle nous a permis d'augmenter le taux de reconnaissance de 0.5 %. Une troisième idée nous a mené à de bien meilleurs résultats, c'est l'hybridation de la première méthode avec la seconde que nous avons appelé méthode Neuro-ACP. L'association des performances du PMC avec ceux de l'ACP nous a permis d'augmenter le taux de reconnaissance de 3% ; soit un taux de reconnaissance de 94.5 %.

La seconde partie de ce travail, que nous considérons comme le complément de la première partie, consiste à détecter les visages dans une scène quelque soit la position, l'éclairage, l'orientation,...etc. Pour cela, nous proposons une méthode hybride alliant la méthode de détection de couleurs de peau et la méthode du « Template Matching ». Cette hybridation nous a permis de contourner les inconvénients de chaque méthode et de minimiser les risques de fausses alertes dans la décision "visage" ou "non-visage".

Les résultats de détection de peau sont fiables dans la plus part des cas, cependant dès que l'image d'entrée devient plus complexe, le taux de mauvaise localisation augmente. Il est donc nécessaire de pallier à cette limite en affinant les traitements préalables afin d'augmenter leurs précisions (élimination des bruits, ... etc).

La phase de détection de peau est suivie par une phase de filtrage des régions détectées pour supprimer tout segment inutile. Après la création d'un modèle de visage, tous les segments issus de

la phase de détection de peau sont comparés avec ce modèle pour déterminer leurs taux d'appartenance à la classe des visages.

Le but lointain de cette seconde partie c'est d'arriver à créer un système qui procède à la détection et la reconnaissance des individus en temps réel quel que soit leurs positions dans la scène et dans n'importe quelle condition d'éclairage.

Comme perspectives de ce travail nous proposons :

Dans le cadre de la classification de visages :

- ✓ Développer de nouvelles applications par notre approche (Neuro-ACP sur d'autres types d'images : images texturées par ex.).
- ✓ Enrichir notre base géométrique par l'ajout d'autres caractéristiques géométriques tel que les contours et les distances du profil gauche et droite.
- ✓ ...etc.

Dans le cadre de la détection de visages :

- ✓ Généralisé le système en un système de détection et reconnaissance de visages en temps réel.
- ✓ Intégration de plusieurs modèles de visages en plusieurs profils (gauche et droite).
- ✓ Enrichir la base de modèles de peau.
- ✓ Combiner la détection de peau avec d'autres méthodes tel que : Transformée de Hough ellipse, réseau de neurones, Adaboost,....etc.
- ✓ Généralisé le module de détection de visages sur les images couleur et niveau de gris.
- ✓ ...etc.

# BIBLIOGRAPHIE

# BIBLIOGRAPHIE

---

- [AR88] J. A. Anderson, E. Rosenfeld. « *Neuro Computing Foundations of search* ». MIT PRESS, Cambridge, 1988.
- [AT02] E. Acosta, L. Torres, « *An Automatic Face Detection And Recognition System For Video Indexing Applications* » Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP'02). IEEE International Conference on. Volume: 4, On page(s): IV-3644- IV-3647 Publication Date: 2002.
- [BB92] A. Belaïd, Y. Belaïd, « *reconnaissance des formes, méthodes et application* », Interedition, paris, 1992.
- [BB03] A/H. Boualleg, Ch. Bencheriet, H. Tebbikh, « *Classification et Reconnaissance de Visages par Réseaux de Neurones* », 1<sup>ères</sup> Journées sur les Scientifiques et Technologies Avancées (JSTA'2003), 24-25 Mai 2003, Guelma.
- [BB04] A. Baccini, P. Besse, « *Data mining, Exploration statistique* », Publication du laboratoire de statistique et probabilité, Université Paul Sabatier Toulouse III, Version septembre 2004.
- [BB05] A/H. Boualleg, Ch. Bencheriet et H. Tebbikh, « *Reconnaissance des visages par la méthode d'analyse en composante principale* », Journée sur l'Informatique Industrielle et l'Imagerie, Guelma, le 18 Avril 2005.
- [BB06] Ch. Bencheriet, A.H. Boualleg and H. Tebbikh, « *Hybrid PMC-PCA Approach For Automatic Faces Recognition* », WSEAS TRANSACTIONS on SIGNAL PROCESSING Journal, Issue N°1, Volume N°2, pp 37-44, January 2006.
- [BC97] S. Bres, P. Champin, « *Traitement d'image et vision artificielle* », Notes de cours INSA Lyon 1997.
- [BC03] O. Bernier, M. Collobert, D. Collobert, « *Détection et suivi robuste de visages en temps réel* » Coresa 2003, Lyon, Janvier 2003.
- [Be00] Ch. Bencheriet, « *Approche géométrique pour la reconnaissance de visages* », Mémoire de magister, Université de Guelma, septembre 2000.
- [Bi03] SM. Bileschi, « *Advances in Component-based Face Detection* », Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on Publication Date: 17 Oct. 2003 On page(s): 149- 156 Massachusetts institute of technology, January 2003.
- [BL90] S. Brunak and B. Lautrup. « *Neural Networks, Computers with Intuition* ». World Scientific edition, Singapore, 1990.
- [BM01] : A. Ben Driss et F. Mekki, « *MEDSeen : système de traitement des images médicales pour l'aide au diagnostic* », Mémoire de magister Université de Annaba 2001
- [Bo92] H. Bourlard. « *Continuous Speech Recognition, From Hidden Markov Models to Neural Networks* » Invited paper, EUSIPCO 92, Brussels (Belgium). In Signal Processing VI, Vol. I, P. 63-70, Elsevier, 1992.

- [Bo05] A/H. Boualleg « La reconnaissance Automatique des Visages », Mémoire de magister en électronique, Université de Guelma, 2005.
- [BP93] R. Brunelli. T. Poggio : « *Faces récoognition : features versus templates* ». IEEE transaction on pattern analysis and machine intelligence, vol 15, p 1024-1052, october 1993.
- [BS01] D. Bryliuk, V.Strovoitov « *Application of recirculation Neural Network and Principal Component Analysis for Face Recognition* ». The 2nd International Conference on Neural Networks and Artificial Intelligence` ICNNAI'2001`, Minsk, Belarus, pp. 136-142, October 2-5, 2001.
- [BT00] H. Bessho, Twaix, Yachida, « *Detecting human face and recognizing facial expressions using potential net* », Proceeding of the 15<sup>th</sup> international conference on pattern recognition ICPR 2000, N 1, P. 387-397. Barcelona, September 3-8, 2000.
- [Bu91] G. Burel, « *Réseaux de neurones en traitement d'images : des modèles théoriques aux applications industrielles* », Thèse de Doctorat de l'université de Bretagne occidentale, 1991.
- [BW89] H. Bourlard et C. Wellkens. « *links between Markov models and multiplayer perceptions* », D.S. Tourezki, Editors, Advances in neural information processing systems, Volume 1 P 502-510, San Muteo, CA, IEE. Morgan Kaufmann, 1989.
- [CE02] P. Cayouette., B. Eng, « *Réseau de Neurones probabilistes pour la Discrimination de Cibles par Image Infrarouge* », Maitrise des Science (Informatique) p 05-18 Mai 2002.
- [CF97] P. Charbonnier, LB Ferrou, G. Aubert, M. Barlau.: « *Deterministic Edge Presenting Regularization in computer imaging* ». IEEE transaction on image processing vol. 6, number 2, p 298, February 1997.
- [CG93] M. Cottrell, B. Girard, Y. Girard, M. Mangears, « *Times Series and Neural network: a statistical method for weight elimination* », 1<sup>st</sup> European Symposium on Artificial Neural Networks, Brussels (Belgique), D factio Editeur, p. 157-164, 1993.
- [CL93] D. Chetverikov and A. Lerch, "Multiresolution Face Detection" Theoretical Foundations of Computer Vision, vol. 69, pp. 131-140, 1993.
- [CN99] D. Chai, K. N. Ngan, « *Face segmentation using skin color map in videophone applications* », IEEE Transactions on Circuits and Systems for video Technology, pages 551–564, 1999.
- [CO03] A. Chau, E. Oji, J. Walters « *The detection of face in Color Images* » Rapport de project EE368, 2003.
- [Co91] P. Comon « *Classification supervisée par des réseaux multicouches* », Traitement du signal, Vol. 6 n°6, P 387-407, 1991.
- [CR00] H. Chang, U. Robles, « *Face Detection* », project report <http://www-cs-students.Stanford.edu/~robles/ee368/skincolor.html>. May 25, 2000.
- [CT03] L. Chouali, S. Tifouti, « *Application de la méthode des hypercubes à la reconnaissance de visages* », Mémoire d'ingénieur, option Informatique université de Guelma, juin 2003.
- [DB00] L. Delahoche et E. Brassart, « *Introduction aux traitements des images* » Notes de cours, DUT Informatique 2ème année, avril 2000.



- [De93] E. Deday. « *From data to knowledge, Boolean, probabilistic, possibilist and belief objects for symbolic data analysis* ». In tutorial at the fourth conference of international federation of classification societies 1993.
- [DK91] E. Deday and Y. Kodratoff. « *Introduction symbolique et numérique à partir des données* ». Capadués édition, 1991
- [DN93] E. Davalo, P. Naim, F.M Clément, « *Des Réseaux De Neurones* », EYROLLES, 1993.
- [Do03] B. Dousset « *Analyse De Données* », notes de cours, Institut de Recherche en Informatique de Toulouse 2003.
- [Fe97] R. Feraud, « *Un modèle connexionniste utilisant un principe longueur de description minimale : application à la détection de visage* », Neural Processing letters volume 5, numero 2, pages 227-233, 1997.
- [FF88] J. Feldman, M. A. Fanty, and N. H. Goddard. « *Computing with structured neural networks* ». IEEE Computer Magazine, 21(3):-103, March 1988.
- [FL04] D. Ferguson, F. Levasseur, « *Détection de visages* » Publication du Laboratoire LAPI - ISICO 2004.
- [Fr] R.Frischholz, BAO and CALTECH face database at the face detection homepage. Available: <http://www.facedetection.com>.
- [FW90] M. Franzini, A. Waibel, K.F. Lee, « *Connectionist Viterbi Training: A New Hybrid Method for Continuous Speech Recognition* » IEEE Transactions on Neural Networks, PP. 425-428, 1990.
- [GB05] B. Guerzize W. Belguidoum « *Système de détection de visages par méthode hybride, détection de peau et Template matching* ». Mémoire d'ingénieur en Informatique. Université de Guelma juin 2005.
- [Gb05] V. Gouet-Brunet, Notes de Cours « *Rappels sur les bases de l'image* » Master STIC – Informatique Bases de la Vision par Ordinateur, 2005/2006.
- [Go00] B. Gosselin « *traitement de l'information, classification et Reconnaissance de formes* », Notes de cours, Faculté polytechniques de Mons, 2000.
- [GP04] P. Gejgus, J. Placek, « *Skin color segmentation method based on mixture of Gaussians and its application in Learning System for Finger Alphabet* », International Conference on Computer Systems and Technologies - CompSysTech'2004.
- [GY00] R. Gross, J. Yang, A. Walbel « *Gaussian Mixture Models for Pose invariant Face recognition Proceedings* » From the international, conference on pattern Recognition ICPR'2000, Vol. 1, p1088-1091, Barcelona, September 3-8, 2000.
- [Ha94] S. Haykin. « *Neural Network, A Comprehensive Foundation* », Macmillan College Publishing Company, New York, 1994.
- [He00] J-Y Hervé « *Vision Artificielle* » Notes de cours, université Rhode Islande 2000.
- [He02] G. Heusch: « *Détection automatique de visages dans une séquence vidéo* », Rapport SSC Semestre 5, VRLAB, Ecole Polytechnique Fédérale de. Lausanne, 18 février 2002.

- [Ho33] H. Hotelling, « Analysis of a complex of statistical variables into principal components », *Journal of Educational Psychology*, 24:417-441,498-520, 1933
- [Ho82] J. J. Hopfield. « *Neural networks and physical systems with emergent collective computational abilities* ». In Proc. Nat'l Academy of Sciences, pages 2554-2558, 1982.
- [La] Laboratoire d'Automatique et d'Informatique de Guelma, LFDB et B-LAIG, Bases de données visages disponible au laboratoire LAIG, <http://www.univ-guelma.dz/recherche/labs/laig/LAIG2009>
- [La98] : M-T. Laurence INRA Notes de cours « *ACP : Analyse en Composante principale* » 1998.
- [Ja01] W. Jacquet, « *Paradigme Connexionniste pour l'apprentissage de fonctions robotiques : Application au problème d'asservissement visuel de bras manipulateur* » Thèse de Doctorat, Juillet 2001.
- [JL02] Y. Jae, S. Lee, I. Yoo « *An Elliptical Boundary Model for Skin Color Detection* », School of Computer Science and Engineering, Seoul National University Shilim-Dong, Gwanak-Gu, Seoul 151-742, Korea 2002.
- [Jo94] J-F. Jodouin, « *Les réseaux neuromimétiques modèles et applications* » Edition Hermès 1994.
- [JR] Dr. John, C. Russ, « *The Image Processing and Measurement Cookbook*» <http://www.reindeergraphics.com/tutorial/index.shtml>.
- [JS99] J-C. Terrillon, A. Shigeru « *Comparative Performance of Different Chrominance Spaces for Color Segmentation and Detection of Human Faces in Complex Scene* », Images Vision Interface '99, Trois-Rivières, Canada, 19-21 May 1999.
- [Ke89] G. kengo « *Approximation by superposition of sinusoidal functions*» Mathematics of control, signals and systems, Springer London 2, PP 303-314, 1989.
- [KG93] M. Kunt, G. Granlund, M. Kocher, « *Traitement numérique des Images* », Presses Presses polytechniques et universitaires romandes, 1993.
- [KN04] G. Kukharev, A. Novosielski, « *Visitor identification - elaborating real time face recognition system* », in Proc. 12th Winter School on Computer Graphics (WSCG), Plzen, Czech Republic, Feb. 2004.
- [KP90] J.F Kolen et J.B Pollack, « *Backpropagation is sensitive to initial condition* », Ohio State University, Computer information Science Research Center, TPR 90-JK-BPSIC. 1990.
- [KP97] C. Kotropoulos et I. Pitas, « *Rule-based Face Detection in Frontal View*», Proc. Intl Conf. Acoustics, Speech, and Signal Processing, vol. 4, pp. 2537-2540, 1997.
- [KP03] J. Kovac, P. Peer, F. Solin, « *2d versus 3d color space face detection* » 4th EURASIP Conference on Video/Image Processing and Multimedia Communication, Croatia, pages 449-454, 2003.
- [LA00] L. Liang, H. Ai, G. Xu, « *Face detection based on template matching and neural network verification*». In: Y. J. Zhang eds. First International Conference on Image and graphics (ICIG'2000). Journal of Image and Graphics (JIG), Vol.5, Supp. 553~556 (2000).

- [LB89] Y. Le Cun, B. Boser, J.S Denker « *Backpropagation applied to handwritten zip code recognition* » Neural computation MIT Press, Cambridge, P.541-551, 1989
- [LC96] B. Leroy, A. Chouakria, I. Herlin, E. Diday, « *Approche géométrique et classification pour la reconnaissance de visage.*» Publication INRIA, Domaine de Voluceau, Recquencourt, France 1996.
- [LC02] M.T Laskri, D. Chefrour, « *who-is : Système d'identification des visages humains* », département d'informatique université de ANNABA, ARIMA volume 1,.P.39-61. 2002
- [Le03] A. lemieux, « *système d'identification de personnes par vision numérique* » pour l'obtention du grade de maître des sciences (M.Sc.), Université Laval QUEBEC, décembre 2003.
- [LM99] G. Langis, L. Marc, Rapport de veille stratégique sur « *la vision artificielle et le traitement d'images* » publication de l'équipe Parole et Vision, au Centre de recherche Informatique de Montréal (CRIM) 1999.
- [Lo92] V. Lorquet, « *Etude d'un codage semi-distribuer adaptatif pour les réseaux de neurones multicouches. Application au diagnostic, à la commande* », Thèse de Doctorat de L'ENST Paris, 1992.
- [LW90] K. Lank , A. Waibel, G. Hinton, « *A Time Delay Neural Architecture for Isolated Word Recognition* » , neural networks, Vol. 3, P. 23-43, 1990.
- [MH91] J.M. Martinez, M. Houkari, CH. Pary, C. Barret, P. Grrizo, «*La retropropagation sous l'angle de la théorie du contrôle* », 4th Int. Conf. On Neural Networks and their Application, Neuro-Nimes 91(France), EC2, P. 279-290, 1991.
- [Mo93] M.C Mozer « *Neural Net Architectures for Temporel Sequence Processing* » Prediction the future an understanding the past, A. weigend & N. greshenfeld (ed.), Addison-wesley Publishing. Redwood city, CA, 1991.
- [MP43] W. S. Mc Culloch and W. Pitts. « *A logical calculus of ideas immanent in nervous activity* ». Bull. Mathematical Biophysics, 5:-133, 1943.
- [MP69] M. Minsky, S. Papert. « *Perceptrons: An Introduction to Computational Geometry*». MIT Press, 1969.
- [MP94] B. Moghaddam, A. Pentland, « *A Face recognition using view-based and modular eigenspaces* ». Technical report 301, MIT Media Laboratory P C Section, 1994
- [MP03] M. Milgram, L. Prevost, R. Belaroussi « *Détection de visages sur des images fixes par combinaison de classifieurs discriminants et de modèles* » 19<sup>e</sup> colloque GRETSI sur le traitement du signal et des images Paris 8 - 11 septembre 2003.
- [MR] O. Millon, V. Rouilly, « *Détection d'ellipses par transformation de Hough*» Projet MBAI.
- [MY99] J.Miao, B.Yin, K.Wang, L. Shen et X. Chen «*A Hierarchical Multiscale and Mul-tiangle System for Human Face Detection in a Complex Background Using Gravity-Center Template* », Pattern Recognition, vol 32, no.7, pp1237 -1248, 1999.
- [NP91] K. Narendra, K. Pathasarathy. « *Gradient methods of optimisation of dynamical systems containing neural networks*»,IEEE trans. On NN, Vol 2, 2, P.252-262, 1991.

- [Ou98] S. Ouchtati, « *Sur les choix d'une représentation de caractères manuscrits Arabes* », Mémoire de Magistère, Université Badji Mokhtar, Annaba, 1998.
- [Pe93] B.A Pearlmutter et al. « *Gradient Descent: Second-Order Momentum and Saturating Error*» Neural Information Processing Systems, Vol. 4 Morgan Kauffmann, San Mateo, PP. 887-894, 1993.
- [Pf00] G. Pfozter, « *Méthodes d'authentification vocale d'utilisateurs dans les systèmes informatiques* », Mémoire d'Ingénieur C.N.A.M. en Informatique, conservatoire national des arts et métiers centre régional associé de Strasbourg, mai 2000.
- [Po87] J. Postaire, « *De l'image à la décision*» Dunod, Paris. mars 1987.
- [Ro62] R. Rosenblatt. « *Principles of Neurodynamics* ». Spartan Books, 1962.
- [RM86] D. E. Rumelhart, J. L. Mc Clelland. « *Parallel Distributed Processing: Exploration in the MicroStructure of Cognition* ». MIT Press, Cambridge, Mass., 1986.
- [RS03] A.Rojbi, J-C. Shmitt, A. Georges, P.Boissonade, « *localisation des personnes dans une séquence vidéo* » CORESA, Lyon 2003.
- [Se89] J. Serra, « *Image analysis and mathematical morphology* », Volume 1, edition: Academic Press, 1989.
- [SH94] F.S Samaria, AC. Harter « *Parameterization of stochastique model for human face identification* », IEEE Computer. Sco. Press, P 138-180, 1994.
- [Si93] S. A. Sirohey, « *Human Face Segmentation and Identification* », Technical Report, CS-TR-3176, Univ. of Maryland, pp.25-29, 1993.
- [TB00] D. Tidjet, N. Bouzida, K. Mokrani, A. Alliche. : « *Reconnaissance automatique des visages. Application de l'analyse en composantes principales* », Conférence sur la vision artificielle CVA'2000 Tizi Ouzou, Novembre 2000.
- [To92] C. Touzet, « *Les réseaux de Neurones Artificiels, Introduction au Connexionnisme, Cours, Exercices et Travaux Pratiques* », 1992.
- [TP91] M. Turk, A. Pentland, « *Eigenfaces face recognition* ». Journal of cognitive neuroscience, 3(1) : 71-86, 1991.
- [TP98] S. Tsekeridouand, I. Pitas, « *Facial feature extraction in frontal views using biometric analogies* », Proc. of the IX European Signal Processing Conference, I:315–318, 1998.
- [TW88] V. Tolat, B. Widrow, « *An Adaptative Broom Balancer with Visual Inputs*» IEEE Int. Conf. on Neural Networks, San Diego (CA), USA, TOME II, P. 641-647, 1988.
- [We74] P. Werbos. « *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences* ». PhD thesis, Dept of Applied Math., Harvard University, Cambridge, Mass, 1974.
- [YK02] M-H Yang, D. J. Kriegman et N. Ahuja, « *Detecting Faces in Images : A Survey*», IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp.34-58, Jan. 2002.