



## Mémoire de Magister

Présenté à l'Université de Guelma  
Faculté des Sciences et de l'Ingénierie

Département de: Génie Electrique  
Option: Signaux et Images

Présenté par : Mr.OULED-DIAF MONCEF

---

---

### Approche Analytique Pour La Reconnaissance Hors-Ligne De l'Ecriture Arabe Manuscrite

---

---

Sous la Direction de : Pr. H.SERIDI

#### LE JURY

<b>Président</b>	Dr. A.LEMZADMI	M.C. Université de Guelma
<b>Rapporteur</b>	Dr. H.SERIDI	Pr. Université de Guelma
<b>Examineurs</b>	Dr. A.BOUALLEG	M.C. Université de Guelma
	Dr. A/K MOUSSAOUI	M.C. Université de Guelma

=====  
2008  
=====

# Remerciements

C'est avec un grand plaisir que j'apporte ce modeste travail à tous ceux qui nous a gratifié de leur soutien et de leur confiance.

Louanges à Dieu, qui m'a donné vie et santé pour le parachèvement de ce modeste travail.

J'exprime ma profonde gratitude à mon encadreur **Dr.Seridi Hamid** pour ses conseils judicieux et éclairés dont j'ai bénéficié durant le déroulement de ce mémoire , son suivi attentif et sa confiance qui m'a été très précieuse, de plus je tiens à le remercier pour m'avoir assuré l'encadrement.

J'adresse aussi mes remerciements aux membres de jury:

A monsieur **A.LEMZADMI**, chef de département de génie électrique à l'université de Guelma, pour m'avoir fait l'honneur de présider mon jury.

A monsieur **A/K Moussaoui**, docteur à l'université de Guelma et **A.Boualleg** docteur à l'université de Guelma qui ont accepté de juger mon travail.

Je remercie aussi ma famille, qui durant mes études, nous a toujours donné la possibilité de faire ce que je veux et qui ont toujours cru en moi, sans oublier les amis.

## RESUME

**D**ans ce mémoire nous avons traité le problème de la reconnaissance hors-ligne de l'écriture arabe manuscrite qui est de nature cursive et dont les caractères sont reliés sur la ligne de base. Nous avons proposé une approche analytique qui manipule un grand vocabulaire, contrairement à l'approche globale qui manipule un vocabulaire réduit. Dans un premier temps une phase de prétraitement de l'image du mot pour réduire les bruits possiblement introduits lors de l'étape d'acquisition. Une phase de segmentation des mots en parties connexes, et les parties connexes en caractères. La dernière étape à être réalisée directement sur les données présentées aux entrées du système est l'extraction des caractéristiques les plus pertinentes pour la reconnaissance. Finalement la classification est réalisée par simulation connexioniste à base de réseaux de neurones multi-couche.

**Mots clés:** Reconnaissance de l'écriture arabe manuscrite, approche analytique, segmentation des mots, réseaux de neurones.

## **ABSTRACT**

**I**n this work, we dealt with the problem of off-line recognition of the Arabic handwritten writing, in fact which is of nature cursive and whose characters are connected on the base line. We proposed an analytic approach whose process a big vocabulary in opposite of global approach whose depend with reduce vocabulary.

Initially, a phase of preprocessing of the image of word makes it possible to reduce meriting intrinsic variability as well as the noise possibly introduced at the time of acquisition. A stage of segmentation allows dividing word in connected parts; then connected parts in characters. The last step to be realized directly on the data presented at the input of the system is the extraction of the most relevant characteristic for the recognition.

Finally, the classification is carried out by the neural network.

### **Key words:**

Handwritten writing arabic words recognition, analytical method, word segmented, neural network.

## ملخص

في هذه الأطروحة قمنا بدراسة مشكلة التعرف على الكتابة العربية الخطية (المكتوبة باليد)، هذه الأخيرة توصف بأنها سريعة و تكتب حروفها متصلة وفق خط يسمى الخط الأساسي. لقد اقترحنا في هذا الشأن طريقة تحليلية للتعرف على الكلمات ذلك أنها تمكننا من التعرف على عدد كبير جدا من الكلمات مقارنة بالطريقة الأخرى المعروفة بالإجمالية و التي نتعرف من خلالها على عدد محدود من الكلمات. كما قمنا بعرض مختلف المراحل اللازمة، بدءا من إدخال الصورة عن طريق فاحص الصور (سكانير) ثم مرحلة ما بعد المعالجة و التي تهدف إلى إزالة التشوهات الناتجة عن المرحلة الأولى، فإذا أصبحت الصورة مؤهلة للمعالجة تأتي مرحلة تجزئة الكلمة إلى أجزاء متصلة والأجزاء المتصلة إلى حروف. بعد ذلك نستخلص جملة من الخصائص المشتركة بين الحروف لتأتي مرحلة التصنيف عن طريق شبكة العصبونات بحيث تكون الخصائص المستخرجة من الحروف هي مدخل الشبكة و الحروف هي مخرج الشبكة.

### الكلمات المفتاحية:

التعرف على الكتابة العربية الخطية، الطريقة التحليلية، الشبكات العصبونية.

# Sommaire

<b>Introduction Générale</b> .....	1
<b>Chapitre I : Reconnaissance des formes</b>	
I.Introduction.....	4
II.Qu'est ce que la Reconnaissance des Formes ?.....	4
III.Domaine d'application.....	6
III.1.Reconnaissance des formes d'origine humaine.....	6
III.2.Reconnaissance des formes dans les images numériques .....	6
IV.Système général de reconnaissance de formes.....	6
IV.1.Acquisition de la forme.....	7
IV.2.Prétraitement .....	8
IV.2.1.Le Filtrage .....	8
IV.2.1.1.Filtre moyenneur.....	8
IV.2.1.2.Filtre Gaussien.....	9
IV.2.1.3.Filtre médian.....	9
IV.2.1.4.Filtrage morphologique .....	10
IV.2.2.La Segmentation.....	11
IV.2.2.1.Approche frontière.....	12
IV.2.2.2.Approche région.....	12
IV.3.Paramétrisation.....	14
IV.4.La Reconnaissance.....	15
IV.4.1.Apprentissage.....	16
IV.4.2.L'étape de test.....	16
IV.4.3.La Décision .....	16
V.Les méthodes de classifications.....	17
V.1.Les méthodes statistiques .....	17
V.1.1.Méthode de BAYES.....	17
V.1.2.Classificateurs paramétriques.....	19
V.1.2.1.Classificateur euclidien .....	19
V.1.2.2.Le classificateur quadratique .....	19
V.1.3.Classification non paramétrique .....	19
V.1.3.1.La méthode du plus proche voisin.....	19
V.2.Méthode structurelle.....	20
V.2.1.Structure des arbres et des graphes .....	20

V.2.2.Structures syntaxiques (automates et grammaires).....	20
V.3.Réseaux de Neurones Artificiels.....	21
VI.Conclusion.....	21

## **Chapitre II : Reconnaissance de l'écriture**

I.Introduction.....	22
I.1.Historique de la reconnaissance des caractères arabes.....	22
II.Variabilité de la tracé écrite .....	24
II.1.Domaine d'application.....	25
III.Les différents aspects de l'OCR.....	26
IV.Les techniques de l'OCR.....	27
IV.1.Méthode de corrélation ou de masquage.....	27
IV.2.Méthodes structurelles.....	27
IV.3.Méthode des experts.....	28
IV.4.Méthode connexionistes.....	28
V.Structure générale d'un système OCR.....	28
V.1.Un module d'acquisition.....	29
V.2.Un module de prétraitement.....	31
V.3.Segmentation .....	32
V.4.Extraction des caractéristiques .....	32
V.5.Reconnaissance.....	33
V.5.1.Evaluation des performances d'un classifieur.....	33
V.5.2.Taux de reconnaissance et taux d'erreur.....	33
VI.Conclusion.....	34

## **Chapitre III : Segmentation et extraction des primitives**

Introduction .....	35
I.Prétraitement.....	35
I.1.Filtrage.....	35
I.1.1.Le filtre médian .....	35
I.2.Binairisation.....	36
I.3.Lissage.....	36
I.4.La correction de l'inclinaison.....	38
I.5.Localisation des lignes de référence (ligne de base) .....	39

I.6.Analyse de l'histogramme horizontal.....	39
I.7.Squelettisation .....	40
I.7.1.Algorithme de ZHANG et SUEN.....	41
II.Segmentation .....	42
II.1.Caractéristiques de l'écriture arabe .....	42
II.2.Définitions de base .....	44
II.3.Segmentation du mot.....	45
II.3.1.La segmentation horizontale .....	45
II.3.2.La segmentation verticale .....	46
II.3.3.Segmentation par étiquetage en composantes connexes .....	47
II.3.3.1.Etiquetage à l'aide d'une table d'équivalence «l'algorithme de Rosenfeld»...	47
II.3.3.2.Amélioration de la méthode d'étiquetage .....	52
II.3.4.La segmentation en caractère .....	54
II.3.5.La segmentation après squelettisation .....	54
II.3.6. Le problème de la sur-segmentation et la solution proposée.....	56
III.Extraction des primitives .....	56
III.1.La problématique de l'extraction de l'information de l'extraction des primitives.....	57
III.2.Les approches de l'extraction des primitives.....	58
III.3.Les catégories de primitives.....	59
III.3.1.Les primitives topologiques ou métriques.....	59
III.3.2.Les primitives structurelles.....	59
III.3.3.Les primitives globales.....	59
III.3.4.Les primitives statistiques.....	60
III.4.Présentation des méthodes d'extraction des primitives.....	60
III.4.1.Les mesures topologiques.....	60
III.4.2.Les mesures d'orientation.....	60
III.4.3.Les mesures de forme.....	61
III.4.4.Extraction des profils .....	62
III.4.5.Les moments invariants.....	63
III.4.6.Détection des boucles.....	65
III.4.7.Remplissage des coins .....	66
III.4.8.description de formes par descripteur de fourier.....	67
III.4.9.Point caractéristique .....	67
III.4.10.Position et nombre des points diacritiques.....	68
III.4.10.1.La position des points (y compris la « Hamza »).....	69



III.4.10.2.Le nombre de points.....	69
IV.Conclusion .....	71
<b>Chapitre IV : Réseaux de neurones</b>	
I.Introduction .....	72
II.Neurones biologiques.....	72
II.1.Structure des neurones biologiques .....	72
II.2.Fonctionnement d'un neurone .....	73
III.Mémoire et apprentissage .....	73
IV.Historique .....	73
V.Les réseaux de neurones artificiels .....	73
V.1.Modélisation.....	75
VI.1.1.Définition .....	76
VI.1.2.Le neurone formel .....	77
VI.1.3.Modélisation générale .....	78
VI.Tour d'horizon des réseaux de neurones.....	78
VI.1.Définition .....	78
VI.2.Les réseaux entièrement connectés (réseau bouclé) .....	79
VI.2.1.Les réseaux les plus célèbres .....	80
VI.2.1.1.Le modèle de Hopfield.....	81
VI.2.1.2.Les cartes auto-organisatrices de KOHONEN.....	81
VI.3.Les réseaux à couches (réseau non bouclé) .....	81
VI.3.1.Les réseaux les plus célèbres .....	82
VI.3.1.1.Perceptron de ROSENBLATT.....	82
VI.3.1.1.1.Structure générale du PERCEPTRON .....	83
VI.3.1.1.2.Limite du PERCEPTRON .....	84
VI.3.1.2.ADALINE : Adaptive Linear Element .....	84
VI.3.1.2.1.Structure générale de l'ADALINE .....	85
VI.3.1.3.Perceptron Multicouche (PMC) .....	87
VI.3.1.3.1.Structure du réseau .....	87
VI.3.1.3.2.Algorithme de RETROPROPAGATION .....	87
VI.3.1.3.3.Limite de l'algorithme de la retropropagation .....	87
VII.Apprentissage .....	88
VII.1.Apprentissage supervisé ou apprentissage avec maître .....	88
VII.2.Apprentissage semi-supervisé .....	88

VII.3.Apprentissage non supervisé ou apprentissage sans maître .....	88
VIII.Les règles d'apprentissage .....	89
VIII.1.Règle de Hebb.....	89
VIII.2.La règle de Widrow -Hoff .....	90
IX.Applications des réseaux de neurones .....	91
X.Conclusion .....	91

## **Chapitre V : Résultat et discussion**

Introduction.....	92
I.Architecture du système realize.....	92
I.1.Acquisition.....	92
I.2.Prétraitement.....	94
I.3.Segmentation.....	94
I.4.Extraction des primitives.....	94
II.Classification et Reconnaissance .....	94
III.Validation des résultats .....	96
IV.Conclusion.....	98
<b>Conclusion générale</b>	<b>99</b>

# *Introduction Générale*

L'idée de construire des machines capables de simuler des êtres humains afin de les aider dans certaines tâches, voire de les remplacer, était antérieure aux ordinateurs.

Leur apparition a permis d'étendre le spectre des tâches à simuler en ajoutant celles dont l'exécution relève de facultés mentales comme la perception et le raisonnement.

Parmi ces tâches, on cite la reconnaissance des formes qui est un domaine de l'informatique et qui a pour but de percevoir et de connaître automatiquement les formes présentées dans les signaux et les images ; il s'agit de mettre au point des programmes et des logiciels capables de traiter la multitude des cas concrets : reconnaissance de caractères, de symboles spéciaux,... des applications sont donc nombreuses en inspection automatique, dans la communication homme-machine, écrite ou parlée.

En effet, la reconnaissance des formes est une discipline essentiellement constructive où il faut étudier la faisabilité, analyser les problèmes, mettre au point les solutions et réaliser la mise en œuvre informatique.

Depuis les quarante dernières années la reconnaissance automatique de l'écriture fait l'objet de recherche intensive, les études dans ce domaine se sont multipliées de nombreuses nouvelles techniques ont été développées. Ces techniques sont loin d'atteindre des performances comparables à l'être humain. En absence de toute contrainte, la reconnaissance de l'écriture s'avère un problème extrêmement complexe qui n'a pas de solution satisfaisante jusqu'à ce jour.

Il est certain que malgré le développement accéléré des modes de communication, l'écriture manuscrite constitue toujours un lien étroit et privilégié entre les êtres humains, leur permettant d'échanger de façon naturelle des sentiments, des informations, des idées. De manière plus prosaïque, elle relie les citoyens à leur administration, les clients à leur organisme bancaire, les consommateurs aux sociétés de distribution. Son importance peut être évaluée de façon permanente, par le volume de courrier manuscrit acheminé chaque jour par la poste, par le nombre de feuilles de déclaration d'impôts, ... etc.

La reconnaissance et l'interprétation de l'écriture manuscrite, sont des techniques relativement récentes, qui s'inscrivent dans le cadre général de la communication écrite homme-machine.

La lecture automatique de l'écriture manuscrite présente un intérêt indéniable dans l'accomplissement des tâches fastidieuses comme celles que l'on recentre dans certains domaines : le tri postal, la lecture des chèques bancaires, la lecture de bordereaux, des bons de commande, des feuilles de déclaration..., elle offre aujourd'hui un surcroît d'intérêt avec le développement des nouvelles méthodes permettant de communiquer directement avec la machine de manière plus naturelle, grâce à l'essor des agendas, des bloc-notes électroniques et des ordinateurs sans clavier.

A priori, on pourrait concevoir la reconnaissance du manuscrit comme une émanation de la reconnaissance de l'écriture imprimée ou dactylographiée, pour laquelle les principaux obstacles semblent avoir été surmontés. Mais le passage de la reconnaissance de l'imprimé à celle du manuscrit n'est pas aussi évident que certains avaient pu le supposer initialement. Sa mise en œuvre est beaucoup plus complexe. Les difficultés rencontrées proviennent principalement de la nature même de l'écriture manuscrite, de son caractère cursif et de son extrême variabilité.

Cependant, malgré ces difficultés, la particularisation des problèmes à résoudre et la spécification des tâches à accomplir ont permis de parvenir aujourd'hui à un nombre de plus en plus important de réalisations tangibles.

Ces applications décrites précédemment montrent clairement les spécificités du domaine de la reconnaissance de l'écriture manuscrite par rapport à celui de la reconnaissance optique des caractères (OCR : Optical Character Recognition) qui concerne les caractères imprimés ou dactylographiés .Il est nécessaire de distinguer également la reconnaissance en ligne (on-line) de l'écriture manuscrite ,qui relève plutôt de l'interfaçage entre l'homme et l'ordinateur ( un stylo spécial est connecté à la machine et ne fonctionne que sur une tablette sensible ), de la reconnaissance hors-ligne (off-line) . Seule la reconnaissance hors-ligne sera considérée dans ce travail.

En effet, en se basant sur l'écriture arabe, qui a bénéficié de quelques recherches, on constate que sa difficulté réside dans sa cursivité, sa large gamme de styles ainsi que les différentes formes d'un même caractère selon sa position dans le mot. Le manuscrit arabe pose encore quelques problèmes et reste partiellement résolu.

Tous les chercheurs sont confrontés à un problème difficile et incontournable, celui de la segmentation. La segmentation fait partie du processus de découpage de texte en mots ou mots en caractères mais utilise une capacité mémoire importante due aux dictionnaires volumineux .Dans notre cas, nous utilisons une segmentation en caractères plus complexe mais utilisant une base de données moins volumineuse.

En pratique, notamment dans le cadre de notre travail, la procédure de reconnaissance proprement dite doit procéder généralement en quatre étapes :

1. L'acquisition des données à lire (mot de la wilaya).
2. Prétraitement (filtrage, binarisation et correction de l'inclinaison).
3. Segmentation et extraction des caractéristiques.
4. La reconnaissance des caractères en utilisant les réseaux de neurones.

Dans le premier chapitre nous donnons un aperçu général sur la reconnaissance des formes, et nous décrivons le schéma général d'un système de reconnaissance ainsi que les méthodes de classification et leurs domaines d'utilisation.

Le second chapitre traite des systèmes de reconnaissance de l'écriture arabe manuscrite où nous présentons les aspects généraux de la reconnaissance de l'écriture, et nous détaillons les étapes nécessaires au développement d'un système de reconnaissance de caractères.

Le troisième chapitre est consacré aux segmentation et extraction des caractéristiques. Cette dernière est l'une des deux fonctions essentielles d'un système de reconnaissance d'écriture. Nous décrivons les caractéristiques pertinentes à la classification et nous présentons les quatre catégories principales de primitives ; les primitives topologiques, statistiques, structurelles et globales.

Le quatrième chapitre est consacré à l'analyse des notions de base, et des outils nécessaires pour l'étude et la compréhension des réseaux de neurones. Nous décrivons la modélisation des réseaux de neurones artificiels, qui consiste à décrire le modèle du neurone, le modèle des connexions, et la règle d'adaptation, et nous définissons la notion et les différents types d'apprentissage. Nous présentons quelques modèles de base des réseaux de neurones, leurs principes, leurs domaines d'applications, leurs capacités, et leurs limites.

Le dernier chapitre est consacré à la présentation des résultats de notre application.

Enfin, nous terminons ce mémoire par une conclusion générale et perspective.

# CHAPITRE I

## *Reconnaissance Des Formes*

## **I. Introduction**

L'énorme quantité d'informations apportées par les systèmes informatiques de saisie et de stockage de mesures que l'homme est amené à traiter ces dernières années, a favorisé le développement de branches comme la reconnaissance des formes (RDF) pour l'interprétation automatique des données.

Si les machines informatiques sont désormais acceptées comme indispensables en tout ce qui concerne le calcul numérique et la gestion, on a quelques réticences à les reconnaître aptes au traitement des données non purement chiffrées.

Pourtant l'ordinateur est par définition l'outil universel capable d'effectuer n'importe quel type de calcul numérique ou symbolique. Cette universalité provient du fait qu'on a le choix du type de représentation pour les données dans les machines, et que Ton peut sur ces représentations programmer n'importe quel algorithme.

Aucun argument fondamental n'empêche de tenter de faire réaliser à de telles machines des algorithmes qui mènent à des comportements traditionnellement réservés à l'être intelligent : reconnaître une personne à sa voix, comprendre un texte, se déplacer dans un environnement non connu a priori, etc..

Le problème peut se résoudre en représentant d'une façon adéquate les données et en les modélisant sous forme d'un programme.

La difficulté provient du fait que l'homme effectue généralement ses tâches d'une manière inconsciente et inanalysable, après un long apprentissage et selon des structures de représentation et de décision mal connues : nous sommes bien en peine d'expliquer comment nous lisons, comment nous reconnaissons un objet, une parole, encore mieux nous sommes incapable de ne pas les reconnaître.

Bien qu'elles soient loin d'avoir les mêmes capacités, les mêmes facilités de perception et d'adaptation que l'homme, les machines informatiques rentrent en compétition avec lui parfois avec beaucoup de succès.

C'est pourquoi, des recherches sont entreprises pour comprendre et approcher ce que l'homme peut faire, aussi bien dans le domaine de la perception (Reconnaissance des Formes) que celui de la compréhension (Intelligence Artificielle).

## **II. Qu'est ce que la Reconnaissance des Formes ?**

La reconnaissance des formes (RF) s'intéresse au traitement par ordinateur des données naturelles, cette discipline vise donc à se placer dans le champ de la perception automatique. La reconnaissance des formes peut être définie par l'ensemble des méthodes et moyens permettant au



moins de reproduire sur machine si possible de dépasser, l'ensemble des moyens de perception de l'homme.

En effet, la reconnaissance de formes vise donc à équiper les ordinateurs des organes de sens ; cela ne signifie pas simplement la mise en place de systèmes physiques de capteurs, mais aussi et surtout l'extraction d'information pertinente, la catégorisation des phénomènes perçus (y compris l'apprentissage des catégories) et l'utilisation des connaissances des experts humains dans les domaines concernés.

La reconnaissance des formes est définie comme étant l'ensemble des Techniques informatiques de représentation et de décision permettant aux machines de simuler un comportement sensible.

Le but de la reconnaissance des formes est d'obtenir des programmes permettant de reproduire efficacement les phénomènes de perception de l'être humain. Elle consiste donc à doter la machine d'organes de sens (capteurs physiques tels que : scanner, caméra, microphone) captant l'information extérieure sous forme variée.

Elle a pour objectif d'une part de décrire ces informations captées, et d'autre part, de prendre sur la représentation ainsi obtenue, une décision d'identification par référence à un ensemble d'apprentissage.

De ces définitions surgissent les difficultés suivantes.

- Saisir l'information.
- La représenter convenablement.
- Prendre une décision d'identification dans cet espace de représentation.

La reconnaissance des formes est un domaine qui regroupe plusieurs sous-domaines tels que la reconnaissance de la parole, la vision par ordinateur...etc. La vision par ordinateur est un domaine qui s'intéresse au traitement et à l'interprétation de l'image. L'entrée d'un système de vision est constituée d'une image. Le système apporte un certain nombre de connaissances qui interviennent à tous les niveaux. La sortie est une description de l'entrée en termes d'objets et de relations entre ces objets [1].

La vision par ordinateur permet, à partir de l'image d'un objet ou d'une scène réelle, d'en déduire des données exploitables par une machine. C'est aussi l'ensemble des techniques et méthodes qui participent à l'acquisition et au traitement des images en vue de prendre une décision de façon automatique. Elle fait intervenir les techniques d'intelligence artificielle (IA) telles que le traitement des connaissances et réseaux de neurones [2].

### **III. Domaine d'application**

La diversité des organes de sens dont dispose la R.F donne une idée sur ses domaines ; nous allons citer quelques uns des plus importants.

#### **III.1. Reconnaissance des formes d'origine humaine**

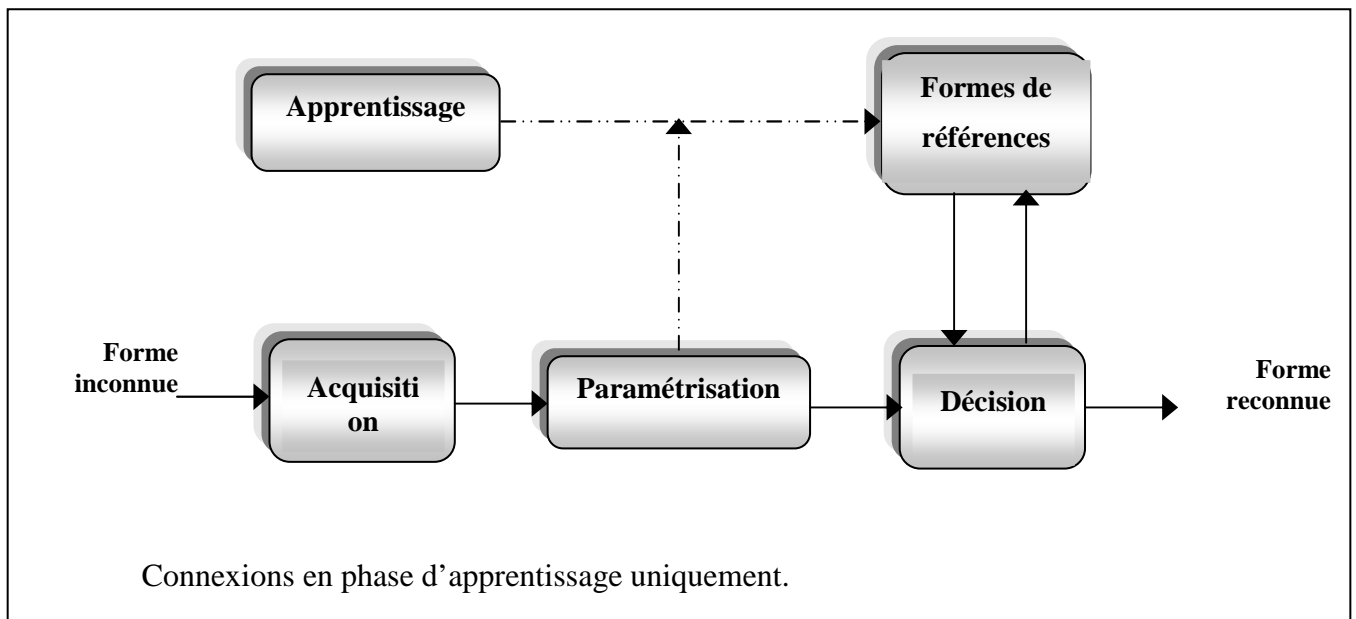
Le traitement du signal fournit des paramètres très utilisables pour pousser plus loin l'analyse, et décider quel est le signal émis, en fonction d'un répertoire de signaux possibles. Un exemple très spectaculaire, promis à un grand développement, est celui de la reconnaissance de la parole. Pour l'instant, les machines commercialisées reconnaissent des mots (ou suite de quelques mots), pris dans un vocabulaire limité à quelques dizaines d'éléments... des prototypes de laboratoires sont capables de tenir compte de la syntaxe et de maîtriser un vocabulaire raisonnablement étendue (quelques milliers de mots).on estime que d'ici une dizaine d'années, les machines seront capables de comprendre la parole humaine et auront des performances suffisantes pour être utilisables dans des taches complexes (intégration de bases de données, acquisition de connaissance spécialisées.. etc.).

Un autre domaine extrêmement utile est celui des signaux biomédicaux; les techniques de la RF peuvent apporter une aide appréciable aux médecins pour les diagnostics et dépouillements des signaux médicaux, tels l'électrocardiogramme, l'électrorétinogramme, l'analyse des chronogrammes [3]; la mesure des débits sanguins. Cependant les applications concernent les images biomédicales: l'analyse des cellules sanguines [4]. La cytologie (étude des segments d' **ADN**, comptage globulaire, étude des structures de virus), etc.

### **IV. Système général de reconnaissance de formes**

Le but d'un système de reconnaissance de formes est l'identification des formes plus au moins complexes qu'on lui présente.

La structure générale d'un tel système est la suivante :



**FigureI.1. Schéma de principe d'un système de reconnaissance des formes**

Donc le système comprend trois grands modules :

- 1- Un système de saisie.
- 2- Un étage de paramétrisation.
- 3- Un étage de décision.

Dont les fonctions sont respectivement :

#### **IV.1.Acquisition de la forme**

Pour capter la forme (par exemple, microphone pour la parole, capteur optiques pour la reconnaissance d'images : caractères écrits, objet d'une scène) et pour la traiter de façon à la rendre manipulable par un ordinateur.

Le scanner.

Appareils photos numériques [5].

Caméra classique reliée a une carte d'acquisition.

Les capteurs CCD.

## IV.2. Prétraitement

Les étapes de prétraitement consistent à sélectionner dans l'espace de représentation l'information nécessaire à l'application. Cette sélection passe souvent par l'élimination du bruit du aux conditions d'acquisition. Toutes les techniques élaborées dans ce sens se gardent de modifier les propriétés essentielles des formes, ce qui pourrait conduire, dans le cas contraire à de graves erreurs d'analyse et plus tard de reconnaissance [6]. Donc, le rôle du prétraitement est de préparer les données reçues du capteur à la phase suivante d'analyse consacrée à l'extraction des paramètres [6].

Le prétraitement peut donc être considéré comme étant constitué de deux phases le filtrage et la segmentation.

### IV.2.1. Le Filtrage

On peut considérer une image comme étant constituée de plusieurs zones homogènes représentant les objets. Dans la réalité, des fluctuations des niveaux de *gris* sont présentes à cause du bruit. On cherchera donc à diminuer l'amplitude de ces fluctuations sans toucher aux transitions [7]. Donc, le bruit est l'ensemble des perturbations *qui* affectent la qualité de la segmentation des régions de l'image [8]. Il se manifeste par la présence d'informations résiduelles venant perturber les données propres à la forme [6].

Le bruit est dû principalement [6]:

- Au dispositif d'acquisition : capteur.
- Aux conditions de prise de vue : éclairage, positionnement.
- A la scène elle-même : poussière, rayure.
- A l'environnement de la scène : fond, composition des formes, nature de la matière. Pour palier à ces dégradations on utilise le filtrage. On distingue essentiellement deux types de filtres :
  - **Filtres linéaires** : la transformation sur le pixel est le fruit d'une combinaison linéaire des niveaux de gris des pixels voisins.
  - **Filtres non linéaires** : les pixels voisins interviennent dans la transformation suivant une loi non linéaire.

Dans ce qui suit, nous allons décrire trois exemples de filtres.

#### IV.2.1.1. Filtre moyenneur [7]

Il s'agit d'effectuer une moyenne des niveaux de gris autour du pixel central. C'est un filtre linéaire qui peut être mis sous la forme d'un masque que l'on va déplacer sur toute l'image.

On peut utiliser des masques du type :

$$1/25 \times \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Masque 5x5

$$1/9 \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Masque 3x3

**Figure I.2. Quelques types de masques**

La taille du masque est un paramètre variable, plus la taille du masque est grande, plus le filtrage est important. L'inconvénient du filtre moyenneur est l'introduction de flou dans l'image.

#### IV.2.1.2. Filtre Gaussien

L'expression de la gaussienne en deux dimensions est donnée par :

$$h(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right]$$

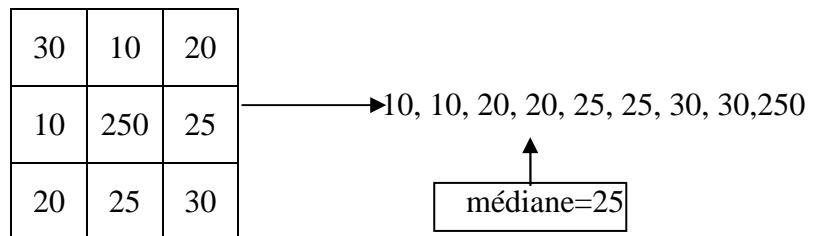
L'avantage d'un tel filtre est que l'on règle très facilement le degré de filtrage à travers le paramètre  $\sigma$ .

Par rapport au filtre moyenneur, il accorde plus d'importance aux pixels proches du pixel central, mais dégrade lui aussi les contours.

#### IV.2.1.3. Filtre médian

Ce filtre est *non linéaire*. Il consiste à affecter au pixel central le niveau de gris séparant la population en deux effectifs égaux.

*Exemple :*



**Figure I.3. Principe du filtre médian**

L'avantage de ce filtre est qu'il donne de très bons résultats sur le bruit impulsionnel (poivre et sel).

Ses inconvénients sont le fait, qu'il supprime les détails fins et qu'il conduit à des temps de calcul élevés.

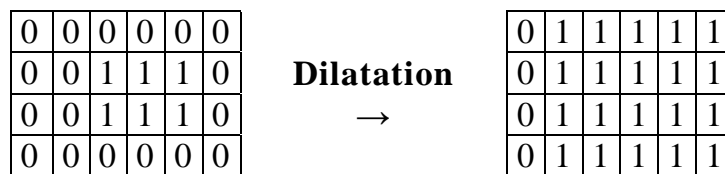
#### IV.2.1.4. Filtrage morphologique [9]

Ce type de filtrage est utilisé pour éliminer des pixels isolés dans des images binaires (2 niveaux de gris noir « 0 » et blanc « 1 ») qui sont considérés comme du bruit. Il met en correspondance chaque pixel et son voisin par une fonction logique (ET, OU, XOR). Parmi les opérateurs morphologiques, nous citons :

- **La dilatation**

Elle permet d'éliminer les pixels blancs isolés. On effectue le ET logique des huit voisins du pixel considéré.

La dilatation élimine les tâches (les trous) blanches dans les zones noires mais ajoute des pixels noirs au contour des objets présents dans l'image.

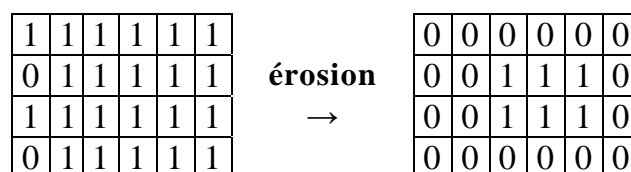


**Figure I.4. Principe de dilatation**

- **L'érosion**

Elle permet d'éliminer les pixels noirs isolés au milieu des parties blanches de l'image, on effectue le OU logique des huit voisins du pixel considéré.

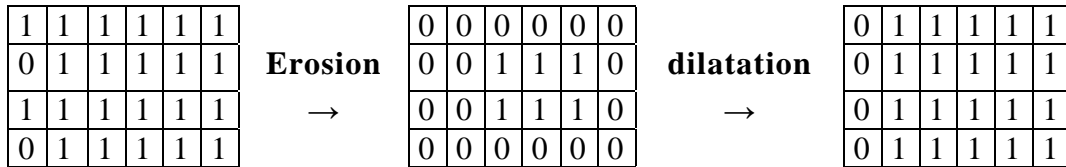
En appliquant une érosion, ces tâches noires peuvent être éliminées mais la taille des objets présents dans l'image diminue car l'érosion enlève des pixels du contour, entraînant une déformation de certains objets.



**Figure I.5. Principe de l'érosion**

- **L'ouverture**

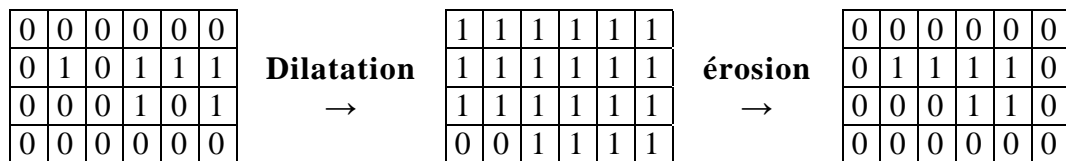
L'ouverture est constituée par une opération d'érosion suivie d'une dilatation. Elle permet de retrouver la taille normale des objets de l'image.



**Figure I.6. Principe de l'ouverture**

- **La fermeture**

La fermeture est une opération morphologique qui consiste à faire subir à l'image une opération de dilatation suivie d'une érosion. Elle permet de retrouver la taille normale des objets de l'image.



**Figure I.7. Principe De La Fermeture**

#### IV.2.2. La Segmentation

Une image constitue une représentation d'un univers composé d'entités (objets). Le but de toute méthode de segmentation est l'extraction d'attributs caractérisant ces objets. Ces attributs correspondent à des points d'intérêt (contours) ou des zones caractéristiques de l'image (régions) [1]. D'où la naissance de deux approches *Frontière et Région*.

- **Approche frontière** : les frontières correspondent aux contours des objets, qui possèdent l'essentiel des caractéristiques de la forme d'une région. La segmentation dans ce cas s'intéresse aux pixels composant le contour.

- **Approche région** : chaque région correspond à la surface de l'objet. La segmentation dans ce cas s'intéresse aux pixels composant ces régions.

#### IV.2.2.1. Approche frontière

Un contour est un point de l'image qui appartient à la frontière entre deux régions voisines. C'est le lieu où l'on observe une discontinuité de la fonction des niveaux de gris [6]. Détecter les contours dans une image, revient à localiser ces discontinuités. Pour cela, deux approches sont utilisées : *Gradient et Laplacien*. Ces approches se basent sur la différentiation de l'image en la dérivant une fois, on obtient alors le *Gradient* ou bien deux fois on obtient ainsi le *Laplacien*.

#### IV.2.2.2. Approche région

La segmentation en régions homogènes vise à segmenter l'image I en se basant sur des propriétés intrinsèques de la région qui déterminent le critère de segmentation [10].

- **Segmentation par utilisation des critères** [7]

La construction des régions de l'image se fait en utilisant un ou plusieurs critères d'homogénéité. Les principaux critères d'homogénéité d'une région utilisés sont :

- Niveau de gris.
- La couleur pour les images en couleur.
- Le mouvement

Trois méthodes de segmentation en régions sont utilisées :

- **Méthodes par fusion**

Au départ, il s'agit de considérer chaque pixel comme une région. Ensuite on essaie de les regrouper à partir d'un double critère :

- critère de similarité des niveaux de gris : exemple : variance des niveaux de gris dans une région R.
- critère d'adjacence.

**Principe** : on choisit un germe et on fait croître ce germe tant que les pixels de son voisinage vérifient les tests d'homogénéité. Lorsqu'il n'y a plus de candidat on choisit un autre germe.

- **Méthodes par agrégation de pixels (par séparation)**

Cette méthode, contrairement à la précédente suppose au départ que tous les pixels appartiennent à la même région. Si les pixels de cette région ne vérifient pas le critère d'homogénéité, la région est divisée. Le processus est réitéré sur chaque région et s'arrête lorsque le critère d'homogénéité est vérifié (obtention d'une région homogène). La variance des niveaux de



gris (inférieure à un certain seuil), peut être utilisée comme critère d'homogénéité. On peut également rajouter un critère sur la taille des régions.

- **Méthodes mixtes (split and merge)**

Ces méthodes combinent les méthodes précédentes. A partir d'une partition arbitraire de l'image en régions, à chaque étape de l'algorithme les régions sont soit divisées si elles ne sont pas homogènes, soit fusionnées si elles sont similaires.

Ces méthodes nécessitent l'utilisation de structures de données compliquées (arbre). Par conséquent, le temps de calcul (parcours de l'arbre) et l'espace mémoire utilisé seront considérables.

- **Segmentation par seuillage** [10]

Cette segmentation consiste à associer à chaque pixel la classe (région) à laquelle il appartient, et cela en comparant son attribut (par exemple niveaux de gris) à un ou plusieurs seuils. Elle est utilisée pour segmenter une image composée de régions dont chacune est caractérisée par un **attribut différent** de l'autre. Le problème réside dans le choix des seuils.

- **Segmentation par étiquetage en composantes connexes** [7,8]

Une image d'étiquettes est une image où les pixels d'une même région ont tous la même valeur appelée étiquette de région. La plupart des méthodes citées nécessitent un algorithme d'étiquetage en composantes connexes : tous les pixels connexes ayant le même attribut doivent être affectés à une même région. Il existe plusieurs algorithmes d'étiquetage. La méthode la plus classique repose sur un balayage séquentiel de l'image. Son principe est le suivant :

Rendre l'image binaire Partant d'un pixel, si une étiquette a été attribuée aux pixels voisins alors affecter la même étiquette à ce pixel sinon lui affecter une nouvelle étiquette. Ce processus est répété jusqu'à ce que tous les pixels de l'image soient étiquetés. Cette méthode est utilisée pour segmenter une image qui possède plusieurs objets similaires en niveau de gris. Puisque cette méthode opère sur une image binaire, nous allons expliquer brièvement *la binarisation*.

### **La binarisation** [9]

La binarisation est une procédure de représentation d'images sur deux niveaux (0 pour noir et 255 pour *blanc* dans le cas où le pixel est représenté par un octet). Cette opération permet *de créer une image* binaire à partir de l'image d'origine. La binarisation peut être la première étape d'isolement des objets par rapport au fond. Elle nous permet soit d'isoler les objets en noir sur fond blanc *ou* l'inverse, soit pour isoler une partie de l'image (masquage) ou pour y effectuer des transformations morphologiques.

### IV.3. Paramétrisation

Chargé d'extraire d'une forme un ensemble de paramètres bien représentatifs (par exemple, l'énergie dans des zones fréquentielles pour la parole, des ensembles de points noirs ou blancs pour décrire un caractère imprimé, ou encore des lignes représentant le « *squelette* » d'un objet dans une image).

La sélection et la définition des primitives lors de l'extraction sont la partie la plus importante dans le système de reconnaissance, car le choix de celles-ci dépend du type de problème à résoudre [11].

L'objectif de l'extraction des caractéristiques dans le domaine de la reconnaissance consiste à exprimer les primitives sous une forme numérique ou symbolique appelée *codage*.

A ce stade, l'image est représentée par un ensemble de caractéristiques numériques afin de réduire sa taille tout en gardant l'information essentielle et en supprimant la redondance. Cette étape est très importante puisque, toute mauvaise extraction se propage vers les étapes ultérieures et influe négativement sur la décision finale du système.

De bonnes caractéristiques sont celles qui satisfont deux conditions :

*Petite variance intra classe* : c'est à dire que des formes légèrement différentes avec des caractéristiques générales similaires, doivent avoir des valeurs numériques proches.

*Grande variance inter classe* : c'est à dire que des caractéristiques de différentes classes doivent être bien différentes numériquement.

#### Extraction des paramètres [6]

Il existe différentes méthodes d'extraction de primitives citées dans la littérature.

Trier et al [11], décrivent des méthodes d'extraction en fonction de la représentation du caractère. Cette extraction se fait sur des images en niveaux de gris, en binaire, en contour binaire et selon une représentation vectorielle.

Selon le type d'images qu'on traite ainsi que la nature des objets, on peut distinguer plusieurs types de paramètres :

#### • Les paramètres morphologiques et métriques

*Surface et périmètre* : La *surface* **s** et le *périmètre* **p** sont certainement les attributs métriques les plus simples pour caractériser la forme d'une région (**s** : égale par exemple au nombre de pixels situés à l'intérieur d'une région et **p** : égale au nombre de pixels situés le long du contour d'une *squelette* ou région).

*Compacité* : la compacité, définie à partir de la surface **s** et du périmètre **p** par :  $C = 4\pi s/p^2$  est certainement le plus utilisé de ces attributs morphologiques. Un théorème bien connu indique que

C prend sa valeur maximale 1 pour le cercle.

**Elongation** : L'élongation fournit un autre attribut morphologique pour apprécier la forme globale d'une région pour un rectangle, l'élongation  $e$  est le rapport entre la longueur  $L$  et la largeur  $I$ .

$$e = L/I$$

#### • Les paramètres statistiques

Ces mesures permettent d'analyser la distribution statistique des pixels dans l'image, ce qui permet de s'en servir ultérieurement pour une pré-classification en sous-ensembles.

Parmi ces paramètres, citons l'histogramme qui est un outil très important puisqu'il permet de définir au mieux l'information globale que contient l'image.

Les mesures de textures appartiennent, elles aussi, à ces paramètres, on peut citer :

$$\text{La moyenne } \overline{g}_R = \frac{1}{S} \sum_{g=0}^{g=n-1} g \times Fr(g) \quad \text{I.1}$$

$$\text{La variance } \sigma_R^2 = \frac{1}{O} \sum_{g=0}^{g=n-1} (g - \overline{g})^2 \times Fr(g) \quad \text{I.2}$$

Où  $Fr$  est la fréquence d'apparition d'un niveau de gris  $g$  dans une région  $R$  de  $n$  niveaux de gris et  $S$  sa surface.

Plus le nombre des caractéristiques est réduit, plus le temps de calcul sera réduit pendant la phase d'analyse et de leurs extractions.

Cependant, il est nécessaire d'effectuer pour chaque problème de reconnaissance une évaluation expérimentale de quelques méthodes d'extraction de primitives les plus prometteuses. Ces expériences permettront de faire un choix judicieux des primitives à extraire car souvent, l'utilisation d'une seule méthode d'extraction de caractéristiques n'est pas suffisante pour obtenir de bonne performance du système de reconnaissance.

La solution évidente est de combiner plusieurs méthodes d'extraction afin de donner une meilleure description de la forme (caractère, chiffre, mot) à classer.

#### IV.4. La Reconnaissance

La reconnaissance regroupe les deux étapes d'apprentissage et de décision qui jouent des rôles assez proches dans les systèmes de **RDF**. La classification est l'élaboration d'une règle de décision qui transforme les attributs caractérisant les formes en appartenance à une classe (passage de l'espace de codage vers l'espace de décision) [12]. Avant qu'un modèle de décision ne soit intégré

dans un système de reconnaissance de l'écriture, il faut avoir procédé auparavant à deux étapes : l'étape *d'apprentissage* et l'étape de *test*.

#### **IV.4.1. Apprentissage**

L'étape d'apprentissage consiste à caractériser les classes de formes de manière à bien distinguer les familles homogènes de formes. C'est une étape clé dans le système de reconnaissance. On distingue deux types d'apprentissage : apprentissage *supervisé* et apprentissage *non supervisé* [12].

##### **• Apprentissage supervisé**

Un échantillon représentatif de l'ensemble des formes à reconnaître est fourni au module d'apprentissage. Chaque forme est étiquetée par un opérateur appelé professeur, cette étiquette permet d'indiquer au module d'apprentissage la classe dans laquelle le professeur souhaite que la forme soit rangée. Cette phase d'apprentissage consiste à analyser les ressemblances entre les éléments d'une même classe et les dissemblances entre les éléments de classes différentes pour en déduire la meilleure partition de l'espace des représentations. Les paramètres décrivant cette partition sont stockés dans une table d'apprentissage à laquelle le module de décision se référera ensuite pour classer les formes qui lui sont présentées.

##### **• Apprentissage non supervisé**

On l'appelle aussi classification automatique ou encore apprentissage sans professeur. Il s'agit, à partir d'échantillons de référence et de règles de modélisation, de construire automatiquement les classes ou les modèles sans intervention de l'opérateur. Ce type d'apprentissage nécessite un grand nombre d'échantillons et des règles de construction précises et non contradictoires. Ce mode évite l'assistance d'un opérateur mais n'assure pas toujours une classification correspondant à la réalité attendue par l'utilisateur.

#### **IV.4.2. L'étape de test**

L'étape de test permet d'évaluer la performance du classifieur pour un apprentissage donné. C'est une étape importante car elle peut mettre en cause le choix des primitives ou le choix de la méthode d'apprentissage.

#### **IV.4.3. La Décision [6]**

La décision ou classement est l'étape de reconnaissance proprement dite, son rôle est d'identifier la forme test à partir de l'apprentissage réalisé.

La méthode de décision est souvent celle utilisée par l'apprentissage, ce qui veut dire que les

critères utilisés pour la comparaison sont les mêmes que ceux utilisés pour l'apprentissage. En effet, il est évident que le choix de critères différents pour la décision, n'assurant pas l'uniformité de la description, ne peut pas conduire avec certitude à un résultat cohérent.

Parmi les techniques utilisées, certaines sont fondées sur la notion de proximité et nécessitent de calculer une distance ou une probabilité de ressemblance avec les modèles définis. D'autres sont fondées sur l'analyse de la structure de la forme et essayent plutôt de vérifier une certaine cohérence dans les relations entre les différentes sous structures.

La réponse de la décision peut être, selon le cas, le nom de la forme en cas de bonne reconnaissance. Plusieurs noms en cas d'ambiguïté ou bien le rejet de la forme en cas d'incompatibilité de description avec les formes de références. Dans les deux premiers cas, la réponse peut être accompagnée d'un taux ou score de confiance.

## **V. Les méthodes de classifications**

Il existe plusieurs méthodes de classification des données, mais le point commun à toutes ces techniques est l'extraction d'un petit nombre de paramètres caractéristiques.

### **V.1. Les méthodes statistiques**

Se fondant sur une caractérisation statistique des paramètres des formes étudiées, ces méthodes permettent de prendre une décision de classification d'une forme inconnue suivant un critère de « *probabilité maximale d'appartenance à une classe* ».

Les exemplaires des classes correspondent aux observations d'une variable aléatoire  $X$ . Chaque réalisation  $x$  est représentée par un vecteur de  $R_n$ . Chaque composante du vecteur correspond à un descripteur. L'objectif est toujours d'assigner une réalisation inconnue à sa classe d'appartenance en minimisant l'erreur de décision. Ce problème peut être résolu de nombreuses manières, le choix de la méthode dépend en partie des connaissances a priori que l'on a sur les distribution de probabilités des exemplaires des classes.

#### **V.1.1. Méthode de BAYES [13]**

L'approche bayésienne consiste à choisir parmi un ensemble des formes, celle pour laquelle la suite de primitives (observations) extraite à la plus grande probabilité a posteriori par rapport aux formes préalablement apprises.

Dans cette approche, on suppose que les mesures  $X_i$  (suite d'observations) faites sur une forme peuvent s'exprimer sous la forme d'un vecteur  $X = (X_1, X_2, \dots, X_n)$ , de l'espace  $R_n$ . On

dispose d'un ensemble  $V$  d'apprentissage, c'est-à-dire d'un jeu de tels vecteurs dont on connaît en plus la classe d'appartenance. Le problème peut alors se résumer sommairement de la façon suivante : étant donné un vecteur inconnu obtenu par mesure sur une forme, trouver la classe à laquelle on doit l'affecter. La probabilité que la vraie forme «  $c$  », représentée par le vecteur  $c = (c_1, c_2, \dots, c_n)$  ; ait pu générer  $X$  est donnée par la formule de Bayes:

$$P\{C/X\} = \frac{P(C)P(X/C)}{P(X)} \quad \text{I.3}$$

$$P(X/C) = \text{MAX} P(X/C) * P(C) \quad \text{I.4}$$

Cette formule met en évidence les termes fondamentaux de l'approche statistique, qu'il faudrait déterminer pour réduire la classe optimale :

$P(X/C)$  : est la probabilité d'observer  $X$  sachant que la vraie suite d'observation est  $C$ , (représente le vrai caractère «  $c$  »). Ce terme est lié à la variation des observations correspondant à la classe de  $C$ .

La valeur de  $C$  qui maximise la probabilité conditionnelle  $P(C/X)$  est appelée *solution de maximum de vraisemblance* (MV), celle qui maximise la probabilité conditionnelle  $P(C/X)$  est appelée *solution au sens du maximum a posteriori* (MAP),  $P(C)$  est la probabilité a priori que  $C$  soit la suite d'observations choisie aléatoirement parmi l'ensemble  $V$  des vecteurs d'observations et  $P(X)$  est la probabilité de la suite d'observations  $X$ .

Puisque  $P(X)$  est indépendant de  $C$ , le vrai caractère (en occurrence  $c$ ) qui fournit la meilleure interprétation de la suite d'observation, est celui qui maximise la probabilité conjointe  $P(X, C)$ .

$$P(X/C) = \text{MAX} P(X/C) * P(C) \quad \text{I.5}$$

Pour estimer  $P(C)$ , il suffit de totaliser  $P(X/C)$  pour toutes les suite observations (représentant, des caractères)  $C \in V$ .

$$P(X) = \sum_{C \in V} P(X/C) * P(C) \quad \text{I.6}$$

Or ce n'est pas raisonnable de stocker toutes les combinaisons de  $P(X/C)$ .

Nous pouvons simplifier la formule en supposant l'indépendance conditionnelle des caractéristiques, ce qui nous amène à :

$$P(X/C) = \prod_{i=1}^n P(x_i/c_i) \quad \text{I.7}$$

La probabilité  $P(x_i/c_i)$  est simplement la probabilité de confusion entre  $x_i$  et  $c_i$  c'est-à-dire

trouver la primitive  $x_i$  à la place de la primitive  $c_i$ .

## V.1.2. Classificateurs paramétriques

### V.1.2.1. Classificateur euclidien [13]

Il s'agit de l'un des plus simples classificateurs qui puissent être conçus. La classe dont le vecteur de caractéristiques moyen est le plus proche, au sens de la distance Euclidienne, du vecteur de caractéristiques de l'objet à classifier est assignée à ce dernier. Les fonctions discriminantes utilisées sont donc de la forme suivante :

$$\Phi_i(X) = -\frac{1}{2}(X - M_i)^T(X - M_i) \quad \text{I.8}$$

### V.1.2.2. Le classificateur quadratique [13]

Comme le nom l'indique, les frontières de décision fournies par ce modèle de classificateur sont quadratiques. Les fonctions discriminantes s'expriment par :

$$\Phi_i(X) = -\frac{1}{2}(X - M_i)^T(X - M_i) \quad \text{I.9}$$

Où  $M_i = E(X|C_i)$  est le vecteur de caractéristiques moyen des éléments qui appartiennent à la classe  $C_i$ ,  $E\{\cdot\}$  désignant l'opérateur d'espérance mathématique, et  $\{\cdot\}^T$  celui de transposition.

Le terme quadratique  $X^T X$  est indépendant de la classe de l'objet, et les fonctions discriminantes peuvent également s'écrire :

$$\Phi_i(X) = M_i^T X - \frac{1}{2} M_i^T M_i \quad \text{I.10}$$

Les frontières qui séparent les classes dans l'espace  $\mathbb{R}^d$  sont ici linéaires.

$$\Phi_i(X) = \min \left( X_i^T X - \frac{1}{2} X_i^T X \right)$$

## V.1.3. Classification non paramétrique

### V.1.3.1. La méthode du plus proche voisin

Ce classificateur est une extrapolation du classificateur Euclidien décrit précédemment. Au lieu d'utiliser le vecteur de caractéristiques moyen  $M$ , comme unique prototype d'une classe, la méthode du plus proche voisin fait intervenir tous les exemplaires des vecteurs de caractéristiques disponibles. La distance Euclidienne entre chacun de ceux-ci et celle de l'objet à classifier est

calculée, et la classe assignée à l'objet est alors celle du prototype le plus proche de celui-ci. Les fonctions discriminants sont donc de la forme :

$$\Phi_i(X) = -\min_{x_j \in \omega_i} \frac{1}{2} (X - M_i)^T (X - M_i) \quad \text{I.11}$$

Le terme quadratique pouvant être omis, ces fonctions se réduisent à :

$$\Phi_i(X) = \min \left( X_i^T X - \frac{1}{2} X_i^T X \right) \quad \text{I.12}$$

Les frontières de décision entre classes sont linéaires par morceaux, c'est-à-dire constituées de nombreux petits polygones convexes, chacun contenant un seul prototype d'une seule classe. Chaque classe est alors délimitée par un polygone très complexe, qui n'est pas nécessairement convexe, ni même d'une seule pièce. Ce classificateur permet ainsi d'établir des frontières de décision relativement complexes, lorsque suffisamment d'exemplaires de chaque classe sont disponibles.

## V.2. Méthode structurelle

Qui s'intéressent à la structure des formes et à leur description en termes d'assemblage de sous formes primitives.

La méthode structurelle à la représentation d'une forme quelconque complexe comme des assemblages structurés de motifs élémentaires qu'on appelle les primitives ; ces derniers peuvent être des arcs, des cercles, segments ..... etc.

Les deux structures les plus courantes dans cette méthode sont :

### V.2.1. Structure des arbres et des graphes

Un graphe est un outil mathématique permettant de décrire des relations dans un ensemble d'objet, ainsi de modéliser une forme. Cette modélisation consiste à la conservation des éléments de base d'une forme. De plus, les relations qui rendent compte de leur organisation. La reconnaissance passe par deux phases :

1. La construction des graphes représentatifs des formes.
2. Comparaison de la représentation obtenue et celle de la référence par un calcul de distance entre graphes.

### V.2.2. Structures syntaxiques (automates et grammaires)

Ces structures sont inspirées de la théorie du langage formel où la forme est analysée syntaxiquement comme une phrase d'un langage à l'aide d'une grammaire appropriée.



### **V.3. Réseaux de Neurones Artificiels [13]**

Un réseau de neurones est un outil d'analyse statistique permettant de construire un modèle de comportement à partir de données qui sont des exemples de ce comportement. Les réseaux de neurones sont essentiellement utilisés pour faire de la classification. Construit à partir d'exemples de chaque classe qu'il a apprise, un réseau de neurones est normalement capable de déterminer à quelle classe appartient un nouvel élément qui lui est soumis.

Pour pouvoir utiliser les capacités de classification d'un réseau de neurones, il faut suivre un certain nombre d'étapes : la construction de la structure du réseau, la constitution d'une base de données de vecteurs représentant au mieux le domaine à modéliser, le paramétrage du réseau par apprentissage et enfin la phase de reconnaissance.

Les neurones peuvent être organisés de différentes manières, c'est ce qui définit l'architecture et le modèle du réseau. L'architecture la plus courante est celle dite du perceptron multicouche.

Le grand avantage des réseaux de neurones réside dans leur capacité d'apprentissage automatique, ce qui permet de résoudre des problèmes sans nécessiter l'écriture de règles complexes, tout en étant tolérant aux erreurs.

Cependant, ce sont de véritables boîtes noires qui ne permettent pas d'interpréter les modèles construits. En cas d'erreurs du système, il est quasiment impossible d'en déterminer la cause.

## **VI. Conclusion**

Dans ce chapitre nous avons fait un aperçu global d'un système de reconnaissance de formes à savoir: les prétraitements, la segmentation, l'extraction des primitives et la classification. Dans le cadre de notre projet, nous nous sommes intéressés à l'ensemble de ces étapes de reconnaissance de formes appliquée à l'écriture arabe. Dans les prochains chapitres, nous étudierons en détail la reconnaissance de l'écriture arabe manuscrite.

# **CHAPITRE II**

## *Reconnaissance De L'Écriture Arabe*

## I. Introduction

La mise en œuvre de la reconnaissance de l'écriture arabe manuscrite est beaucoup plus complexe, les difficultés rencontrées proviennent principalement de la nature même de l'écriture manuscrite, de son caractère cursif et de son extrême variabilité.

La reconnaissance automatique de l'écriture manuscrite consiste à faire lire par une machine, des mots ou des séquences de mots tracés par les êtres humains.

L'écriture est un moyen particulier de codage d'information (idées, concepts, pensées,..). La reconnaissance de l'écriture peut donc être considérée comme un processus inverse, c-à-d un processus de décodage ou de lecture artificielle.

La complexité provient du fait que l'écriture manuscrite est à la fois personnelle et universelle, particulière et générale, elle possède deux caractéristiques fondamentales :

- Un aspect ou signifiant qui caractérise l'auteur de l'écrit.
- Un contenu sémantique qui caractérise le sens de ce qui est écrit.

### I.1. Historique de la reconnaissance des caractères arabes [14]

Un des premiers essais de reconnaissance des caractères est celui de **GRIMSDALE** en 1958, un autre travail pour la même période est celui de **BLEDSOE** et **BROWING** qui décrivent une approche plus générale, utilisant une mosaïque de 10\*15 cellules photoélectriques

Au début des années 60, **MURRAY EDEN** émet un concept, qui connu sous le nom de ((Analyse par méthode de synthèse)) est à l'origine d'une très importante recherche dans le monde. Mais la caractéristique du travail d'**EDEN**, est qu'il a formellement prouvé que tous les caractères manuscrits sont formés par un nombre limité de caractéristiques un point qui a été implicitement inclus dans les travaux précédents. Cette notion a été utilisée plus tard dans toutes les méthodes syntaxiques (structurelles) de reconnaissance des caractères.

Il importe de consacrer un paragraphe spécial à la reconnaissance de l'écriture arabe puisque c'est le sujet qui nous intéresse directement.

Comparativement à d'autres domaines de la reconnaissance des formes et de l'intelligence artificielle, la reconnaissance automatique de l'écriture arabe a suscité malheureusement un nombre relativement restreint de travaux.

Récemment, **SHOUKRY** (1983) a présenté une démarche pour la présentation et la structuration des problèmes de reconnaissance structurelle en général et des traces manuscrites en particulier.

L'outil utilisé, fondé sur la notion de carte topologique, a permis d'obtenir une représentation simple d'un tracé donné. Pour tirer profit de cette représentation pour les

besoins de la reconnaissance , un logiciel a été développé dans le but de reconnaître des ensembles variés de caractères entre autre l'alphabet arabe. Les caractères acquis par tablette graphique, sont représentés sous forme de graphe puis classés en fonction de leur caractéristiques statistiques et topologiques. L'algorithme développé est chargé d'explorer le graphe du caractère et d'extraire des caractéristiques topologiques (point d'intersection, point d'inflexion,...) géométrique (nombre de faces, codage des faces du graphe).

Pour chaque caractère, il est construit un vecteur de référence qui représente les caractéristiques dites principales. Lors de la reconnaissance, ce vecteur est comparé avec le vecteur de référence. S'il y a plusieurs candidats, des caractéristiques annexes sont calculées, puis une distance est évaluée entre le vecteur à reconnaître et ceux possédant les mêmes caractéristiques principales.

D'après **A.AMIN** (1985) qui a testé cette méthode sur les caractères arabes manuscrits , son application à des caractères standards permet d'obtenir un taux de reconnaissance plus élevé que pour les caractères manuscrits .En revanche, l'algorithme développé est général et peut s'appliquer à n'importe quel système d'écriture.

Deux méthodes ont été testées pour la reconnaissance des caractères, la première est dite structurelle, elle est fondée sur l'exploitation de caractéristiques propres à l'écriture arabe , telle que les groupes de points, les zigzag ; la deuxième n'est que l'application de la méthode de **SHOUKRY** .

Différentes applications ont montré qu'il y a des changements dans la forme géométrique du caractère ou encore dans sa position , sa taille , son inclinaison et parfois même son orientation pour faciliter la reconnaissance . **FRANZ** (1962) a développé une technique de normalisation pour la reconnaissance de caractères imprimés.

Le troisième système est fondé sur une méthode syntaxique , déjà employé par **BERTHOD** (1974) pour la reconnaissance des caractères latins : Un analyseur syntaxique effectue la reconnaissance à partir d'une description formelle décrivant les caractères en terme de traits et de courbes. **AMIN** affirme avoir testé avec succès cette méthode pour un vocabulaire d'une centaine de mots, mais aucune précision sur le taux de reconnaissance n'a été donnée.

Aucun système n'est actuellement capable de reconnaître l'écriture manuscrite de façon universelle. Les systèmes existants se répartissent en deux grandes classes de méthode de reconnaissance :

- ⊗ La reconnaissance **hors ligne** " **off -line** " ou « statique ».
- ⊗ La reconnaissance **en ligne** " **on-line** " ou « dynamique ».

La reconnaissance **off-line** est réalisée après que l'écriture ou l'impression soit accomplie. Dans un système OCR typique, les caractères entrés sont lus et digitalisés par un scanner. Chaque

caractère est alors localisé et segmenté et la matrice résultante est introduite dans le module de prétraitement pour effectuer un éventuel filtrage, réduction de bruit, et une normalisation de la taille si ceci est nécessaire.

Plusieurs articles traitent de la reconnaissance des caractères latins [18], chinois et japonais [16,17]. Cependant, bien que presque un tiers de milliard de personnes dans le monde entier, dans plusieurs langues différentes, emploient les caractères arabes pour l'écriture, peu de progrès de recherches, que ce soit en on-line ou en off-line, ont été réalisés en reconnaissance automatique des caractères arabes. Ceci résulte du manque d'appui et naturellement de la nature cursive de ses règles d'écriture. Deux approches ont été appliquées à la reconnaissance de caractères arabes imprimés et manuscrits. Celles-ci peuvent être classées par catégorie comme suit:

1- L'approche analytique dans laquelle les mots sont segmentés en caractères (ou pseudo caractères) [15,19,20,21,22,23,24,25,26,27,28]. Ces caractères sont alors envoyés à un système de reconnaissance de caractères et la combinaison des résultats d'identification produit une liste rangée de mots possibles. Cette approche est théoriquement plus forte en manipulant un grand vocabulaire. En effet avec un nombre constant de classes (par exemple le nombre de lettres dans l'alphabet), on peut manipuler n'importe quelle suite de caractères et donc un nombre illimité de mots.

2-L'approche globale dans laquelle l'identification considère des mots et ne tente pas d'identifier des caractères individuellement. Cette approche a un avantage qui est celui d'éviter l'étape difficile de segmentation. Sa grande faiblesse est la manipulation d'un vocabulaire limité. Néanmoins, l'augmentation de la taille du lexique risque d'augmenter exponentiellement le taux de confusion.

Le mot arabe (imprimé ou manuscrit) est en général cursif et les lettres arabes sont normalement reliées sur la ligne de base. Cette connectivité s'avérera un problème important dans le procédé de segmentation.

## **II. Variabilité de la trace écrite [CRE]**

L'alphabet arabe se compose de 29 lettres. Cependant, la forme d'un caractère pourrait avoir jusqu'à quatre formes différentes voir tableau 1, selon sa position dans le mot: au début du mot, au milieu du mot, à la fin du mot et isolé, de ce fait le nombre de formes de caractères augmente de 29 à environ 60.

On distingue généralement la variabilité des écritures entre des personnes différentes ou *variabilité inter-scripteur* et la variabilité de l'écriture d'une même personne au cours du temps ou *variabilité intra-scripteur*.

Plusieurs facteurs contribuent à créer les différents aspects de la variabilité de l'écriture :

- Les facteurs de variabilité propre au scripteur

L'écriture se modifie lentement au cours du temps avec l'âge et l'état de santé physique et mentale du scripteur.

- Les facteurs de variabilité intentionnels

Suivant la tâche à effectuer : brouillons, notes personnelles, lettres intimes ou officielles, la trace écrite prend un aspect différent. La rapidité plus ou moins grande du tracé peut modifier la forme du caractère.

- Les facteurs de variabilité émotionnels

Une écriture spontanée présente davantage d'inégalités qu'une écriture contrôlée qui paraît plus compacte et moins personnelle.

- Les facteurs de variabilité circonstanciels

La variabilité peut provenir des conditions matérielles et physiques dans lesquelles le texte a été écrit, notamment de la nature du support (ex : papier plastique, etc.) et du type du stylo (plume, feutre, etc.).

- Les facteurs de variabilité contextuels

Les formes d'une lettre varient suivant qu'elle est située au début ou à la fin d'un mot et en fonction de la lettre qui la suit et de celle qui la précède.

## II.1. Domaines d'application

Aucun système n'est actuellement capable de reconnaître l'écriture manuscrite de façon universelle. Les systèmes existants se répartissent en deux grandes classes de méthode de reconnaissance :

✕ La reconnaissance **hors ligne** ou « statique ».

✕ La reconnaissance **en ligne** ou « dynamique ».

Dans le cas « **hors ligne** » : Il s'agit de reconnaître le texte manuscrit à partir d'un document écrit au préalable. L'image du texte écrit est numérisée à l'aide d'un scanner, les informations recueillies se présentent sous la forme d'une image discrète constituée d'un ensemble de pixels.

Dans le cas « **en ligne** » : Il s'agit de reconnaître l'écriture au fur et à mesure de son tracé, le texte est saisi avec un stylo et une tablette à numériser.

La reconnaissance de l'écriture **hors ligne** connaît un essor important dans les domaines associés au développement d'intérêts économiques.

En outre, certains caractères arabes différents peuvent avoir exactement la même forme et se distinguent les uns des autres par seulement rajout de points diacritiques voir (**Figure II.1**).



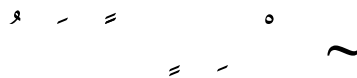
**Figure II.1. Cinq caractères ont les mêmes formes.**

Ces points diacritiques (une partie d'un caractère qui est nécessaire pour compléter un caractère arabe) peuvent être un point, un groupe de points ou une hamza (celle-ci peut apparaître sur, au-dessus de ou au-dessous de la ligne de base) et positionnés selon le cas, au-dessus ou au-dessous du corps du caractère voir (**Figure II.2**).



**Figure II.2. Différentes points diacritiques**

Un mot arabe peut avoir aussi des diacritiques jouent le rôle des voyelles dans les langues latines, qui sont écrits au-dessous ou au dessus



**Figure II.3. Différentes diacritiques voyelles**

L'Arabe est écrit de droite à gauche Chaque caractère, dans un mot, peut être relié a gauche et/ou à droite par un point de raccordement qui, habituellement, se trouve sur une ligne imaginaire, qu'on appelle ligne de base.

### **III. Les différents aspects de l'OCR**

Il existe plusieurs systèmes d'OCR selon le type de données à traiter et de l'application visée. Voici quelques aspects de L'OCR :

- **Reconnaissance de l'imprimé ou du manuscrit**

L'approche n'est pas la même selon qu'il s'agisse de l'imprimé ou du manuscrit. Les caractères imprimés sont dans le cas général alignés horizontalement et séparés verticalement, ce qui simplifie la phase de lecture, bien que certaines fontes présente parfois des accollements qu'il faut repérer. Dans le manuscrit, les caractères sont souvent ligaturés et leur graphisme est inégalement proportionné provenant de la variabilité intra et inter-scripteur.cela nécessite généralement l'emploi de techniques de délimitation spécifique, et souvent des connaissances contextuelles pour guider la lecture.

- **Reconnaissance monofonte, multifonte, omnifonte**

La question se pose pour un texte imprimé, un système est dit monofonte s'il ne traite qu'une fonte à la fois, il est dit multifonte s'il est capable de reconnaître un mélange de fontes préalablement apprises,

enfin un système est dit omnifonte s'il est capable de reconnaître toute fonte sans l'avoir apprise.

- **Reconnaissance en ligne (on-line) ou hors ligne (off-line)**

La première dynamique se fait pendant l'écriture, elle permet de corriger ou de modifier l'écriture de manière directe et instantanée, tandis que la deuxième démarre après la fin de l'acquisition du document entier, elle permet d'analyser un grand nombre de caractères, résultant d'un pré traitement coûteux [30].

## **IV. Les techniques de l'OCR**

### **IV.1.Méthode de corrélation ou de masquage**

Le principe de cette méthode est fondé sur l'examen d'un certain nombre de pixels dans l'image du caractère et l'étude de leur appartenance. Une forme est reconnue si l'ensemble de ses pixels est présent. Cependant à cause de la rigidité de ces méthodes (conçues à partir d'emplacements fixes de pixels) chaque masque de la taille du motif recherché est déplacé sur toute la région où le motif est supposé être et on identifie le motif à partir du masque qui réalise le maximum de correspondances. Cette méthode est la plus ancienne et n'est plus employée aujourd'hui que pour des produits spécifiques (banque, poste), elle représente l'avantage d'être très rapide, relativement fiable et surtout très facile à mettre en oeuvre, mais elle correspond à une série de caractères bien particuliers avec une fonte fixe.

### **IV.2.Méthodes structurelles**

Elles sont fondées sur les structures propres du caractère, la structure est exprimée en terme de composants primitifs correspondant à des formes élémentaires telles qu'un rebroussement, un changement d'orientation, un accroissement ou un décroissement de pente, ces composantes sont appelées primitives, parmi ces méthodes on peut citer :

- **Méthodes de tests**

Elles consistent à appliquer sur chaque caractère entré des tests de plus en plus fins sur l'absence ou la présence de ces primitives. Le processus le plus répandu ou habituel consiste à diviser à chaque test, l'ensemble des choix en deux jusqu'à n'obtenir qu'une seule forme correspondant au caractère entré.

- **Comparaison des chaînes**

Les caractères sont présentés par des chaînes de primitives, la comparaison du caractère test avec un modèle de référence consiste à mesurer la ressemblance entre les deux chaînes en calculant des distances.

- **Méthode syntaxique**

Chaque caractère est représenté par une phrase dans un langage où le vocabulaire est constitué



de primitives. Les caractères d'une même famille sont représentés par une grammaire, la reconnaissance consiste à déterminer si la phrase de description du caractère peut être générée par la même grammaire, on définit la distance d'une phrase à une grammaire comme étant la distance minimale de modification à faire subir à la phrase pour que le caractère (ou la phrase) soit accepté par la grammaire.

### **IV.3.Méthode des experts**

Chaque expert a la responsabilité de son caractère, l'avantage de cette technique est la rapidité; dès qu'un caractère est identifié, on passe à l'autre, le seul problème de cette méthode est la complexité de programmation d'un expert. Un système expert est un logiciel issu des recherches en intelligence artificielle (IA) destiné à remplacer l'être humain dans toutes les tâches d'expertise sur un domaine ou un problème en tentant de reproduire le raisonnement d'un ou plusieurs experts humains [31].

### **IV.4.Méthodes connexionnistes**

Les réseaux de neurones ont fait l'objet de plusieurs recherches dans différents domaines durant ces dernières années. Leur utilisation dans les systèmes OCR comme classifieur se justifie par le fait qu'ils sont très appropriés aux problèmes n'ayant pas une solution déterministe (cas de la reconnaissance de l'écriture) et qu'ils sont facilement entraînaibles. Par conséquent, ils permettent de s'adapter facilement à la grande variabilité inhérente à l'écriture.

Un réseau de neurone est un ensemble d'unités linéaire (unités de calcul) interconnectées entre elles. Un réseau de neurone est caractérisé par le nombre de neurones et leur configuration de connexion.

Dans un système d'OCR, par exemple, le réseau de neurone est entraîné pour apprendre le résultat correct de la classification de chaque exemple d'apprentissage. Les connaissances a priori et les connaissances détaillées au réseau sont minimales au début. Mais il s'auto organise lors de l'apprentissage pour atteindre les relations entre les entrées et les données.

## **V. Structure générale d'un système OCR**

L'objectif de la reconnaissance d'un document est l'obtention d'une description syntaxique des différents éléments qui le constituent à partir de la masse d'informations qu'il contient à l'état brut ; c'est d'abord un moyen économique qui permet de baisser considérablement le coût de la saisie et augmenter sa vitesse.

La reconnaissance de caractères permet à une machine d'approcher le comportement humain de la lecture, au départ le document est sur un support papier et après passage par un système OCR, ce dernier va créer une copie numérisée du document à traiter.

De manière générale, un système OCR voire (**Figure II.4**) est composé de:

- Un module d'acquisition et de prétraitement.
- Un module de segmentation.
- Un module de reconnaissance.

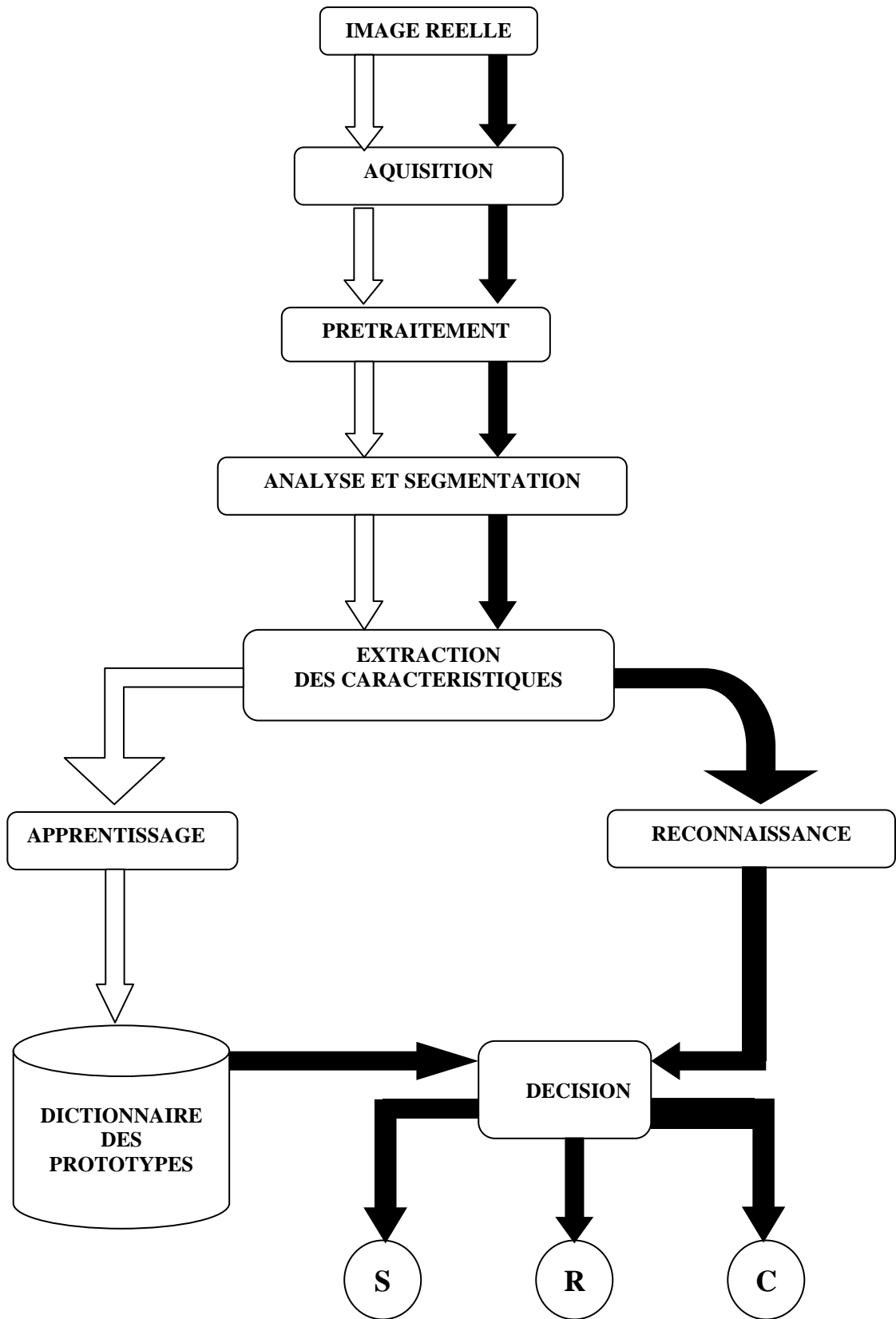
### **V.1. Un module d'acquisition**

L'opération d'acquisition permet le passage de l'information du monde réel au monde numérique de l'ordinateur. Ce passage nommé aussi numérisation ou digitalisation, n'est autre que la transformation d'un signal analogique délivré à travers un dispositif photosensible (scanner) en un signal numérique.

L'acquisition de l'image passe par deux étapes, l'une matérielle et l'autre logicielle.

#### **a) Acquisition physique**

Le système d'acquisition, le scanner en l'occurrence, est relié à l'ordinateur et est piloté par un logiciel. Ce dernier permet de digitaliser l'information analogique et stocke les données numériques dans un fichier graphique.



⇨ : Mode apprentissage.

➡ : Mode reconnaissance.

S : Succès. R : Rejet. C : Confusion.

Figure II.4. Structure générale d'un système OCR

## **b) Acquisition logique**

Le deuxième niveau d'acquisition, consiste à exploiter le fichier graphique pour en extraire l'image. Elle se présente alors sous forme d'un tableau où chaque élément représente un pixel dont l'intensité est une valeur numérique couramment appelée niveau de gris.

## **c) Fichier graphique**

Les principaux formats utilisés sur PC sont : BMP, TIF, GIF, PCX, TGA, JPEG. La différence entre un format et un autre réside dans l'organisation de la structure, remplacement de l'image au sein du fichier, sa taille, etc. L'information est inscrite dans un fichier selon des règles qui régissent la structure de chaque format.

La structure du fichier que nous utilisons dans notre application est le format BMP.

## **V.2. Un module de prétraitement**

L'information brute de l'opération d'acquisition peut engendrer des bruits dus à l'orange de capture lui-même, ce qui altère la qualité de l'image.

Le prétraitement consiste à remédier à cet inconvénient en filtrant la forme saisie. Plusieurs techniques sont employées elles dépendent totalement du mode d'acquisition, du type d'image et des informations que l'on veut retirer.

Le prétraitement peut comprendre des dizaines d'opérations différentes effectuées sur l'image numérisée. Cette dernière peut présenter des distorsions ou des perturbations provoquées par divers facteurs, et qui sont principalement l'outil d'acquisition (qualité du digitaliseur) et l'état du document (vieillesse, froissement...). Cette étape regroupe toutes les techniques visant à améliorer la qualité d'une image. Parmi ces techniques, on peut citer par exemple le problème de la réduction de bruit d'acquisition ou celui de l'élimination du fond dans le cas de la lecture d'un texte écrit sur un fond texture.

On inclut aussi généralement dans les prétraitements des opérations de type redressement des écritures penchées, détection et éventuellement redressement des lignes de base, etc. [32]. Dans des images à plusieurs niveaux de gris, on peut recourir à une binarisation qui est considérée comme un traitement préliminaire, et ce, tant pour la facilité de manipulation d'images binaires que pour le faible espace mémoire requis pour les stocker. Le principal problème de la binarisation réside dans le choix d'un seuil et dans la détermination de la zone de validité de ce seuil de décision. Si la zone de validité représente l'image entière, le seuil est dit «global»; dans ce cas, très courant en pratique, le risque que l'image comporte des zones pour lesquelles ce choix est mal adapté est grand. Si la zone de validité est locale, les techniques de binarisation sont alors à seuil adaptatif [29].

L'étape de prétraitement, située en amont de la segmentation, a pour but de faciliter cette dernière en faisant une correction de l'inclinaison dans l'image de document numérisée.

### **V.3. Segmentation**

La segmentation est une opération qui consiste à séparer un document en ses entités élémentaires. Ces entités correspondent à des groupements de pixels délimités par un séparateur physique. L'entité de base est la composante connexe, le séparateur est alors la transition noir blanc ou blanc/noir. Dans le cas de l'écriture, l'entité de base est le caractère qui ne correspond pas directement à l'entité physique pour les lettres attachées.

Deux approches de segmentation différentes existent; celle qui tient compte du contour des objets ; on parle alors de segmentation par extraction de contours. Et celle qui privilégie le contenu de la région à segmenter, donc les pixels qui la composent. On ne peut pas dire qu'une approche est meilleure que l'autre, leur différence réside dans l'information qu'elles mettent en valeur et qui est le contour pour l'une et le contenu pour l'autre.

Comme nous le verrons plus loin, la segmentation d'un paragraphe de texte en caractères s'effectue en trois étapes : La segmentation du texte en lignes puis en parties connexes (qui peuvent être des mots ou des sous-mots) et enfin en caractères.

### **V.4. Extraction de caractéristiques**

Pour simplifier l'image du mot, il est préférable de la paramétrer grâce à l'extraction de certaines de ces caractéristiques, l'ensemble de ces caractéristiques constitue une sorte de signature du caractère ou bien un vecteur d'attribut, et au lieu d'alimenter le module de reconnaissance avec l'image elle-même, on va l'alimenter par ce vecteur. Cette étape est primordiale car l'image initiale est ignorée pour ne considérer que ces paramètres, et les performances du classifieur dépendent directement du choix des vecteurs d'attributs. Un bon choix de ces derniers se présente comme un vecteur de nombre minimal de paramètres pour éliminer la redondance et l'ambiguïté.

Comme nous l'avons vu dans le chapitre I, cette étape appelée aussi étape d'analyse, est l'une des étapes les plus difficiles et plus importantes dans la construction d'un système de reconnaissance. La difficulté revient principalement au manque de règles préalablement établies pour obtenir les traits pertinents de l'objet à reconnaître. Le but de cette étape est la représentation de la forme à reconnaître par un ensemble d'informations manipulables par un ordinateur.

Le choix des paramètres est une tâche difficile et exige certaines contraintes tel que :

- Les caractéristiques ne demandent pas des calculs complexes.
- Compter le nombre d'intersection de la lettre avec des droites déterminées.
- Calculer les moments invariants et les moments invariants de projection.
- Calculer les paramètres dans un espace transformée (Fourier, Hoogh...).
- Détecter les boucles, les concavités et les pentes.

- Les caractéristiques des histogrammes.
- Effectuer une squelettisation et en extraire les points caractéristiques

## V.5.Reconnaissance

Elle constitue l'étape Classification et décision. La reconnaissance est l'identification d'un caractère après extraction de ses caractéristiques. Un système ne peut reconnaître un caractère que s'il le connaît déjà. L'apprentissage est donc une phase importante de la chaîne de reconnaissance.

L'apprentissage consiste à construire un dictionnaire qui contiendra un certain nombre de prototypes obtenus en consultant toutes les formes possibles des caractères pouvant être soumis à ce système [33]. Plus le dictionnaire est riche, meilleure sera la reconnaissance.

Donc dans cette étape, tout caractère saisi en vue de sa reconnaissance est représenté sous une forme analogue à celle utilisée pour décrire ceux du dictionnaire; il s'agit donc de trouver le prototype qui lui ressemble le plus et de décider si le caractère est identifié ou non. Le résultat de l'identification pourrait être :

- **Caractère ou mot reconnu** : le caractère est effectivement reconnu (succès).
- **Caractère ou mot rejeté** : le caractère n'est pas reconnu. Si l'apprentissage est dynamique, il permettra de rajouter les caractères non reconnus si ceci s'avère nécessaire.
- **Caractère ou mot ambiguïté** : le caractère est confondu, par le système, avec un autre (erreur de reconnaissance). Dans ce cas, il faudra rajouter des caractéristiques supplémentaires pour lever cette ambiguïté.

### V.5.1. Evaluation des performances d'un classifieur

Nous décrivons dans ce paragraphe les critères les plus souvent employés pour évaluer la performance des systèmes de reconnaissance en phase de généralisation. L'objectif est d'obtenir une estimation la plus fidèle possible du comportement du système dans des conditions réelles d'utilisation. Pour cela des critères classiques comme les taux de reconnaissance et taux d'erreurs sont presque systématiquement utilisés [34].

### V.5.2. Taux de reconnaissance et taux d'erreurs

Les taux de reconnaissance et d'erreurs permettent d'évaluer la qualité du classifieur  $f_M$  par rapport au problème pour lequel il a été conçu. Ces taux sont valus grâce une base de test qui contient des formes d'écritures dans le même espace de représentation  $E$  que celles utilisées pour l'apprentissage. Pour que l'estimation du taux de reconnaissance soit la plus fiable possible il est primordial que le reconnaiseur n'ait jamais utilisé les échantillons de cette base pour faire son apprentissage (la base de test ne doit avoir aucun individu en commun avec la base d'apprentissage

et les éventuelles bases de validation). De plus cette base de test doit être suffisamment représentative du problème de classification.

Les performances en terme de taux de reconnaissance sont alors déterminées en présentant au classifieur chacun des exemples  $e_j$  de la base de test et en comparant la classe donnée en résultat  $f_M(e_j) = s$ .

À la vraie classe de  $e_j$ . En considérant que la base de test contient  $N$  individus et que sur ceux-ci  $N_{corrects}$  sont bien classés par le système le taux de reconnaissance  $\tau_{rec}$  est simplement défini par :

$$\tau_{rec} = \frac{N_{corrects} \times 100}{N}$$

Le taux d'erreur  $\tau_{rec}$ , est défini à partir du nombre d'individus  $N_{corrects}$  mal classés.

$$\tau_{rec} = \frac{N_{err} \times 100}{N}$$

## VI. Conclusion

Dans ce chapitre, nous avons fait une brève synthèse sur la reconnaissance de l'écriture largement entamée dans la littérature en faisant le lien avec le domaine de la RDF. Nous avons présenté aussi une description générale d'un système de reconnaissance de l'écriture. Les différents modules d'un système OCR y sont décrits brièvement, à savoir les modules d'acquisition, de prétraitement, de segmentation et de reconnaissance. L'étape de segmentation représente la partie la plus difficile du processus car tout erreur dans cette phase se propage à la phase de reconnaissance.

Nous implémenterons dans le chapitre suivant les modules de segmentation et extraction des primitives.

# **CHAPITRE III**

*Segmentation Et*

*Extraction Des*

*Primitives*



## Introduction

Le prétraitement est une technique qui sert à préparer les données reçues du capteur à la phase suivante d'analyse consacrée à l'extraction des paramètres. Cette phase n'est possible et surtout fiable que si les données du capteur sont dénuées de bruit, corrigées de leurs erreurs éventuelles, homogénéisées, normalisées et réduites à l'essentiel. Nous allons développer dans la suite les principales opérations de prétraitement en les illustrant par des exemples dans le cas des images.

## I.Prétraitement

### I.1.Filtrage

Cette étape consiste à éliminer les bruits dans l'image obtenue à la sortie du scanner. Plusieurs filtres ont été implémentés dont les filtres linéaires (filtre moyen, filtre de Dérivée) et le filtre médian non linéaire. Le filtre médian est celui qui s'adapte le mieux à notre application (puisque le mot présente dans la majorité des cas un type de bruit impulsif) et c'est donc lui que nous avons implémenté dans notre système

#### I.1.1. Le filtre médian

Le principe est de calculer en un point non pas une combinaison linéaire des niveaux de gris de ses voisins, mais une valeur médiane dépendant d'un tri des niveaux de gris des points voisins. (Figure III.1)

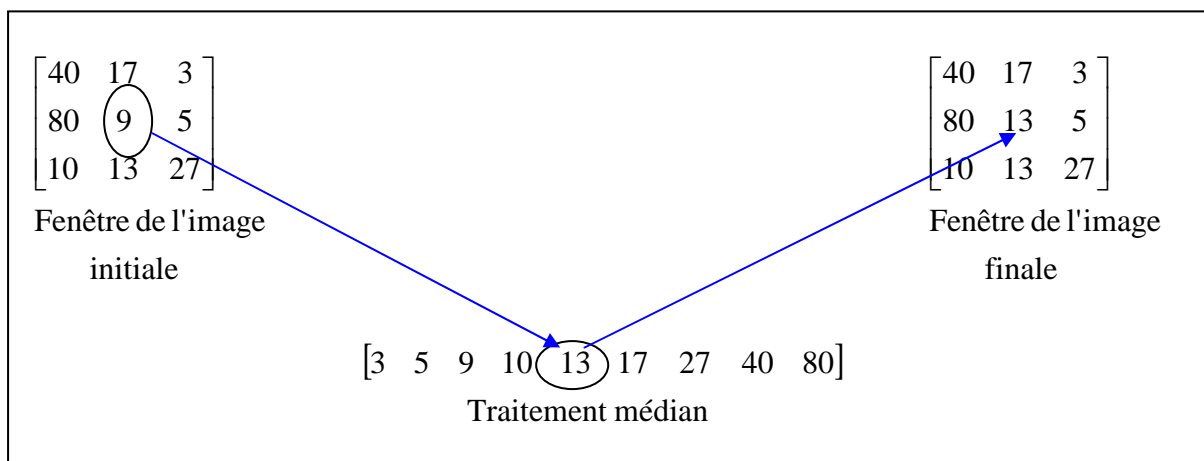


Figure III.1.Principe du filtre médian

### Algorithme

**Pour** chaque **pixel** de l'image **Faire**

- 1- On classe les pixels voisins du pixel courant (compris dans la fenêtre) par valeurs croissantes.
- 2- On prend la valeur médiane des pixels classés et on l'affecte au pixel courant.
- 3- Elimination des points noirs isolés.

## I.2. Binairisation

Pour les images acquises en niveaux de gris, la binairisation devient nécessaire avant d'attaquer la phase du traitement. La binairisation permet de mieux distinguer les caractères du fond, elle consiste à attribuer à chaque pixel de l'image une valeur de '0' ou '1' : « 0 » qui représente le noir (le mot), et le « 1 » représente le blanc (le fond de la page). Pour cela, elle applique en premier lieu l'opération de seuillage.

### Seuillage

Il consiste à déterminer la valeur du seuil à partir duquel tous les pixels ayant un niveau de gris inférieur à cette valeur sont représentés par un zéro « 0 » le noir, et tout les pixels de niveau de gris supérieur auront la valeur un « 1 » (le blanc).

La valeur du seuil est déterminée à partir de l'histogramme du niveau de gris se trouvant dans la vallée entre les deux pics de l'histogramme.

Le seuil doit être calculé d'une manière adéquate, car les composantes du texte liées des traits fins peuvent se déconnecter, ce qui modifie la forme originale du texte.

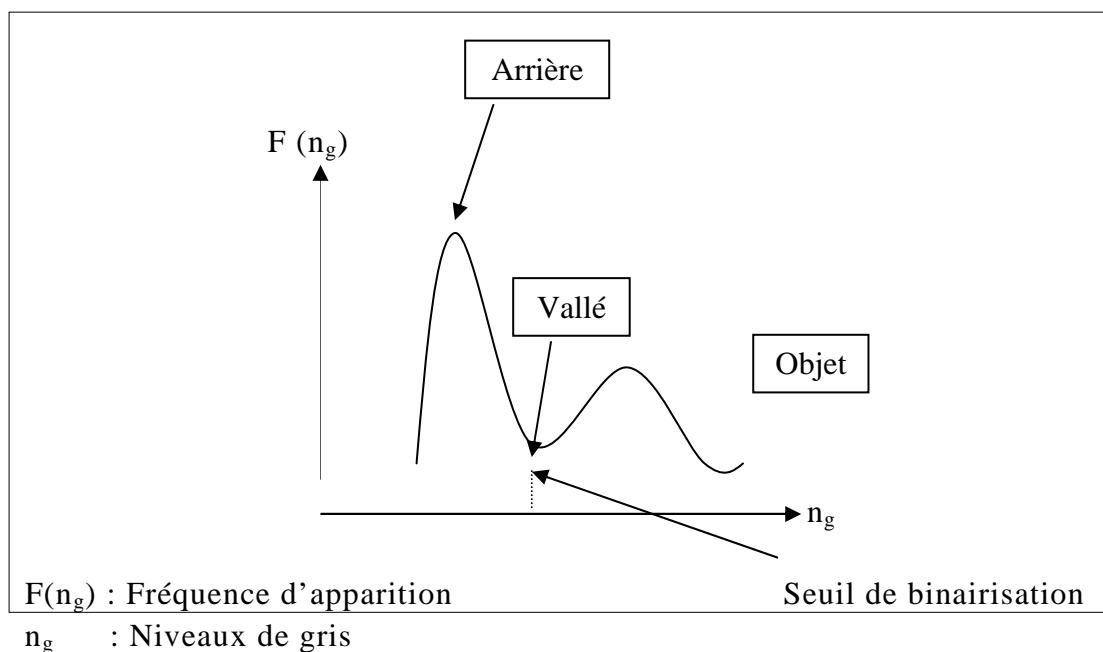


Figure III.2. Histogramme des niveaux de gris

## I.3. Lissage

L'image du mot peut être entachée de bruit dû aux artefacts de l'acquisition et souvent à la qualité du document, conduisant soit à des absences de points (trous) soit à des empâtements ou des excroissances et donc à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales appelées nettoyage et comblement.

Nous utilisons les notions suivantes pour la description des masques de transformation de voisinage : « 1 » pour un point de la forme et « 0 » pour un point du fond et X pour un point quelconque.

L'opération de nettoyage conduit à supprimer les petites tâches et les excroissances de la forme. Elle est réalisée de différentes manières suivant le type de bruit à enlever :

- Par élimination des points de la forme isolés ou situés à l'extrémité des contours, en appliquant sur l'image le masque suivant dans les huit directions :

0	0	0
X	1	0
0	0	0

- Par élimination des points formant des angles droits ou des excroissances du contour, en appliquant le masque suivant dans les huit directions:

X	0	0
1	1	0
X	0	0

- Par élimination des points formant des coins, en appliquant le masque suivant dans les quatre directions principales:

0	0	0
X	1	0
0	X	0

Pour le comblement, il s'agit d'égaliser les contours et de combler les trous internes la forme du caractère en lui ajoutant des points noirs. On distingue :

- Le comblement de trous isolés. Si le voisinage d'un point du fond correspond au masque suivant, alors ce point est mis à 1 :

1	1	1
1	0	1
1	1	1

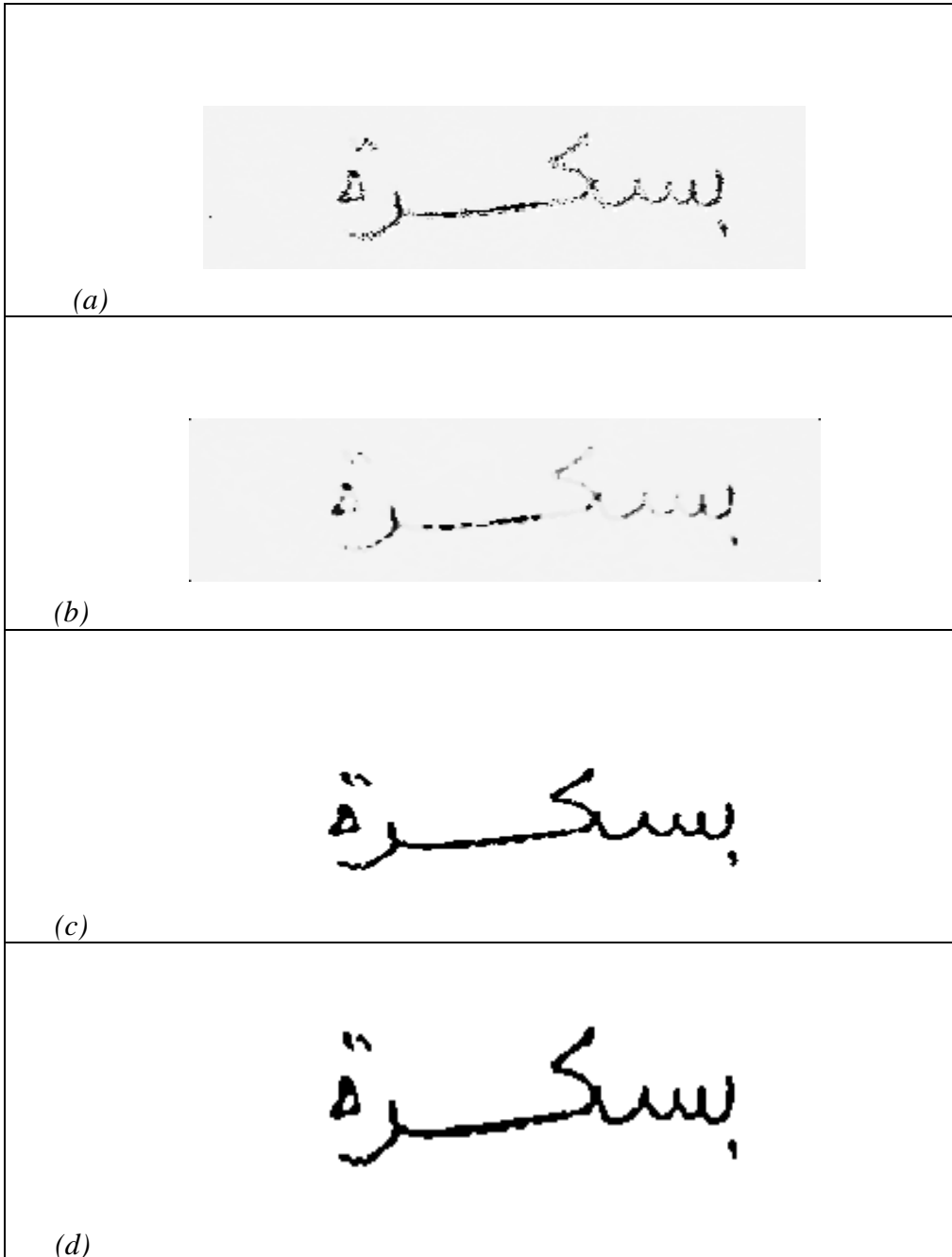
- La correction des irrégularités des contours dues à un effritement de la forme. On applique sur l'image le masque suivant dans les huit directions :

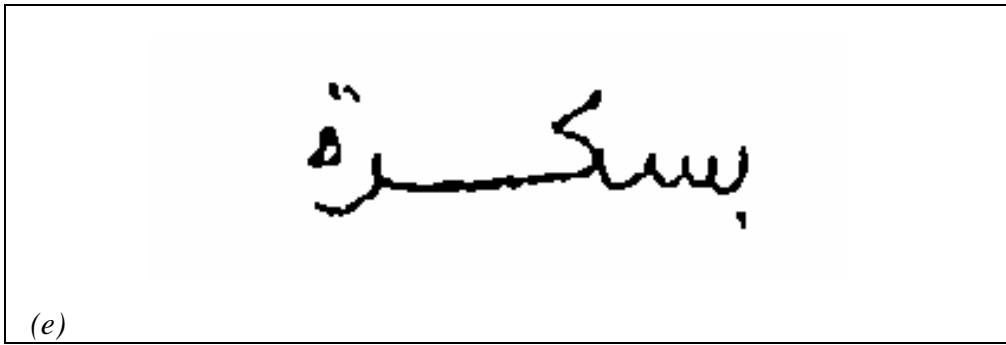
X	1	1
0	0	1
X	1	1

#### I.4. La correction de l'inclinaison

Cette opération consiste à corriger la pente d'un mot ou à redresser l'inclinaison des lettres dans un mot afin de faciliter la segmentation. L'opération est effectuée à l'aide d'une transformation ligne par ligne où chaque pixel noir de coordonnées  $(x, y)$  est remplacé par les coordonnées  $(x', y')$

données par: 
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & -\tan \theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

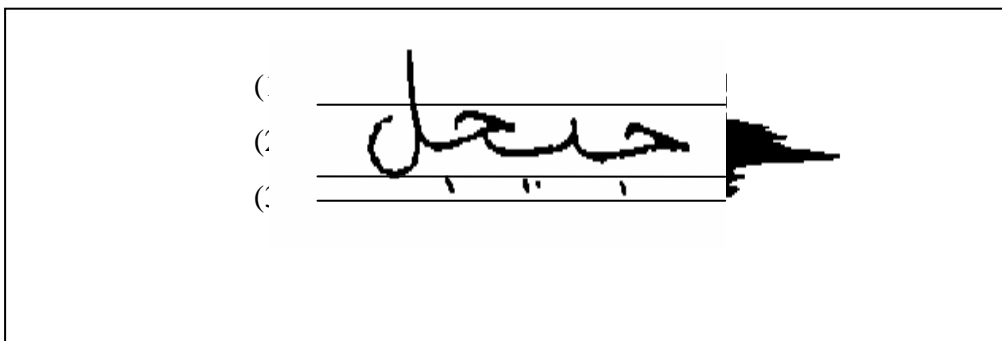




**Figure III.3. (a) image bruitée, (b) image filtrée, (c) Image binarisée (seuil = 242), (d) Résultat de lissage, (e) Correction de l'inclinaison**

### I.5. Localisation des lignes de référence (ligne de base)

D'un point de vue local, on peut résumer les règles de disposition des caractères entre eux par la connaissance de trois zones d'écriture distinctes (Figure III.4) la zone de corps des minuscules (2), la zone des extensions hautes (1) et la zone des extensions basses (3). De plus, ces zones sont normalement horizontales et identiques sur toute une ligne d'écriture [35].



**Figure III.4. Les différentes zones d'écriture**

- (1) : Zone des extensions hautes.
- (2) : Zone médiane.
- (3) : Zone des extensions basses.

### I.6. Analyse de l'histogramme horizontal

L'histogramme permet de mettre aisément en évidence la zone médiane du mot car la contribution des minuscules sans hampe ni jambage y est déterminante. C'est une fonction  $h$  :

$$h : \begin{array}{l} [1, n] \rightarrow \mathbb{N}^+ \\ i \rightarrow h(i) \end{array}$$

Où l'indice  $i$  représente l'indice des lignes et  $n$  le nombre de lignes de l'image.

On recherche dans un premier temps une ligne de l'image appartenant à la zone médiane, quelle que soit la combinaison des lettres constituant le mot. Pour cela, on suppose que la hauteur

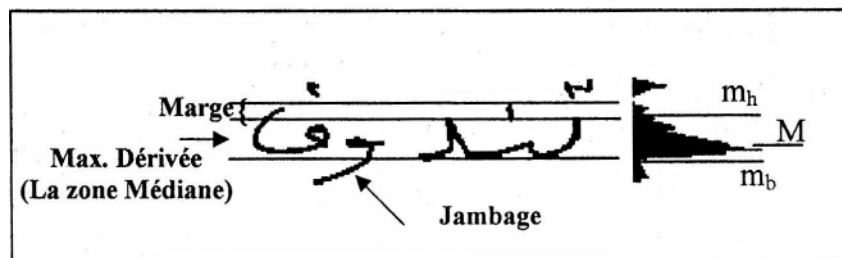
attendue de la zone médiane soit comprise aux alentours d'une valeur  $h_M$  fixée à l'avance, car il n'est pas envisagé ici de reconnaître des mots de taille quelconque. On calcule, pour  $h_M/4 < i \leq n - h_M/4$ , la somme  $S(i)$  suivante :

$$S(i) = \sum_{j=i-h_M/4}^{i+h_M/4} h(j)$$

On opère ainsi un important lissage. L'indice  $i$  correspondant à la ligne où la somme  $S(i)$  est maximum, est noté  $M$  : dans la plupart des cas, cette ligne d'indice  $i$  se trouve à l'intérieur de la zone médiane du mot, même si elle est parfois plus près d'un bord de la zone que de l'autre.

Dans un deuxième temps, on recherche dans la partie supérieure à la ligne d'indice  $M$  ainsi que dans la partie inférieure, les indices des minimums de l'histogramme respectivement  $m_h$  et  $m_b$ . Dans le cas idéal, ces deux minimums délimitent la zone médiane.

"تندوف" La (Figure III.5) illustre la détection de ces paramètres sur le mot "



**Figure III.5. Les différentes zones d'écriture**

Lorsque la zone médiane du mot est obtenue, les lettres à hampe ou à jambage sont distinguées des lettres médianes quand elles dépassent les limites de la zone médiane d'une grandeur supérieure à une marge fixée.

### I.7. Squelettisation

Le but de la squelettisation est de simplifier l'image du caractère, la squelettisation consiste à rechercher une ligne médiane ayant le même ordre de connexions que le caractère lui-même.

Les critères retenus pour les méthodes de squelettisation sont les suivants :

- L'épaisseur du squelette doit être de 1.
- Le squelette doit conserver les propriétés topologiques de la forme comme le nombre de parties, de trous et de connexités.
- Le squelette doit conserver les propriétés métriques de la forme comme la longueur totale et la distance entre les parties de la forme.

Les différents algorithmes de squelettisation tentent de conserver les propriétés topologiques et métriques de la forme liées à la courbure, à l'orientation et à l'échelle.

Parmi les algorithmes de squelettisation le plus utilisée est l'algorithme de ZHANG et SUEN.

### I.7.1. Algorithme de ZHANG et SUEN

Cet algorithme conserve l'essentiel de l'information dans le squelette, de plus il est rapide sachant qu'un contour est tout pixel allumé qui possède au moins un 8-voisin « éteint ».

Dans cet algorithme, on considère que 1 correspond au point allumé et 0 correspond au point éteint.

L'algorithme de ZHANG et SUEN applique une série d'itérations sur chaque point de contour suivant deux étapes

1<sup>ère</sup> étape :

Elle est appliquée à tous les points du contour. On force un point à 0 si les conditions suivantes sont toutes vérifiées :

- 1)  $2 \leq N(P1) \leq 6$  ou  $P1$ : point de contour et  $N(P1)$ : nombre de pixels allumés de  $P1$ .
- 2)  $S(P1) = 1$  ou  $S(P)$ : nombre de transitions 0-1 dans l'ordre  $P2, P3, P4, \dots, P8, P9$  ( $P2, P3, \dots, P8, P9$  sont les voisins de  $P1$ ).
- 3)  $P2 * P4 * P8 = 0$ .
- 4)  $P2 * P6 * P8 = 0$ .

$$\begin{array}{|c|c|c|} \hline P9 & P2 & P3 \\ \hline P8 & P1 & P4 \\ \hline P7 & P6 & P5 \\ \hline \end{array}$$

2<sup>ème</sup> étape :

On applique à tous les points du contour de l'image résultante de l'étape 1, les conditions suivantes :

- 1-  $2 \leq N(P1) \leq 6$ .
- 2-  $S(P1) = 1$ .
- 3-  $P2 * P4 * P8 = 0$ .
- 4-  $P2 * P6 * P8 = 0$ .

Si toutes les conditions sont vérifiées alors on force le point  $P1$  à 0. Cette procédure de base est appliquée itérativement à l'image jusqu'à ce que plus aucun point ne soit supprimé ce qui termine l'algorithme produisant le squelette de l'image.

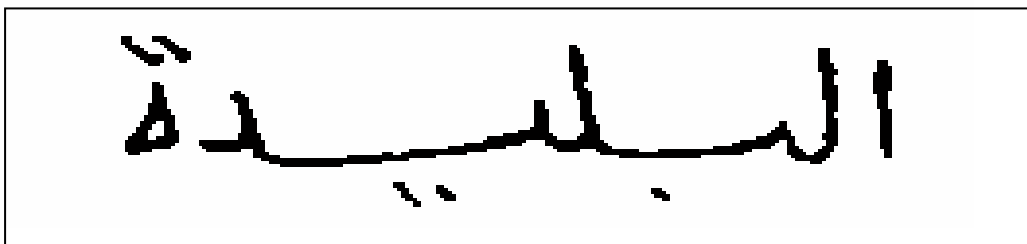


Image non squelettisée



**Figure III.6. Résultat de squelettisation utilisant l'algorithme de ZHANG et SUEN**

## **II. Segmentation**

Après le prétraitement, on passe à la segmentation qui permet de localiser les éléments de l'écriture arabe, à savoir les lignes, les parties connexes et de séparer ses dernières en caractères isolés, la segmentation d'une partie connexe en lettres est une opération naturelle mais elle est extrêmement complexe.

La segmentation est une étape critique et décisive dans plusieurs systèmes de reconnaissance.

En d'autre terme, l'efficacité d'un système de reconnaissance en dépend fortement.

Elle est définie , comme étant l'opération qui cherche à décomposer une image de texte en pseudo image de symboles individuels , le résultats de cette opération , est une forme isolé à partir d'une image et qui pourrait être une composante connexe ou non.

### **II.1. Caractéristiques de l'écriture arabe**

La langue arabe est très riche en mots , ce qui la reconnaissance de mot de manière globale impossible .on opte alors pour la reconnaissance des caractères séparément ,ces entités de base n'étant pas initialement isolées , elle font partie d'un ensemble souvent structure en mots, les lignes et paragraphes . Ce qui explique la nécessité d'un module de segmentation pour pouvoir effectuer la séparation des paragraphes en lignes les lignes en mots ou sous mots et ces derniers en lettres[36].

Nous pouvons résumer les caractéristiques de l'écriture arabe par les points suivants :

- L'alphabet arabe comporte 28 caractères de base qui peuvent prendre quatre formes différentes suivant leur position dans le mot et (au début, au milieu, à la fin, isolé), plus la hamza qui a un rôle très important dans la phonétique et le LAM-ALIF qui composé de LAM et ALIF.

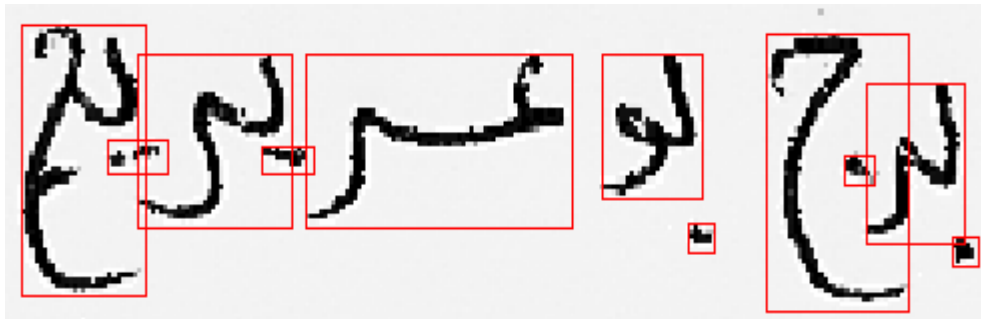


**Figure III.7. La lettre "ع" dans cet exemple a quatre formes**



Contrairement à l'écriture latine, l'écriture arabe s'écrit de droite à gauche de manière cursive (liées entre elles).

- Certains caractères différents ont la même forme mais se distinguent par la position et le nombre de points qui leur apparaissent.
- Les voyelles ne sont pas systématiquement utilisés dans l'écriture arabe .des signes qui correspondent aux voyelles sont employés pour éviter les erreurs de prononciation, on distingue deux type de textes :
  - Avec les signe de voyelles (le Coran, les livres d'apprentissage pour enfant).
  - Sans les signes de voyelles (les livres, les journaux, les publications, etc....).
- Certains caractères se chevauchent (Figure III.8) c-à-d qu'il est impossible d'encadrer un caractère dans un rectangle sans croiser son successeur.



**Figure III.8.Le chevauchement**

- Plusieurs caractères sont composés de boucles et de courbes, ces courbes sont souvent écrites dans le sens horaire.
- De plus, un mot arabe peut avoir un ou plusieurs sous mots.

Nom caractère	Position début	Position milieu	Position finale	Position isolée
<b>Alif</b>	أ	ا	آ	ا
<b>Ba</b>	ب	ب	ب	ب
<b>Ta</b>	ت	ت	ت	ة, ت
<b>Tha</b>	ث	ث	ث	ث
<b>Djim</b>	ج	ج	ج	ج
<b>Haa</b>	ح	ح	ح	ح
<b>Kha</b>	خ	خ	خ	خ
<b>Dal</b>	د	د	د	د
<b>Dhal</b>	ذ	ذ	ذ	ذ
<b>Ra</b>	ر	ر	ر	ر
<b>Zey</b>	ز	ز	ز	ز
<b>Tâ</b>	ط	ط	ط	ط

Dhá	ظ	ط	ظ	ظ
Sad	ص	ض	ص	ص
Dhad	ض	ظ	ض	ظ
Sin	س	س	س	س
Chin	ش	ش	ش	ش
Ain	ع	ع	ع	ع
Ghain	غ	غ	غ	غ
Fa	ف	ف	ف	ف
Káf	ق	ق	ق	ق
Kaf	ك	ك	ك	ك
Lam	ل	ل	ل	ل
Mim	م	م	م	م
Noon	ن	ن	ن	ن
Ha	ه	ه	ه	ه
Waw	و	و	و	و
Ya	ي	ي	ي	ي

Tableau III.1.Les différentes formes de la lettre arabe

## II.2. Définitions de base

Les projections verticales et horizontales sont définies par les équations suivantes:

$V_j = \sum g(i, j)$ : La projection verticale.

$H_i = \sum g(i, j)$ : La projection horizontale.

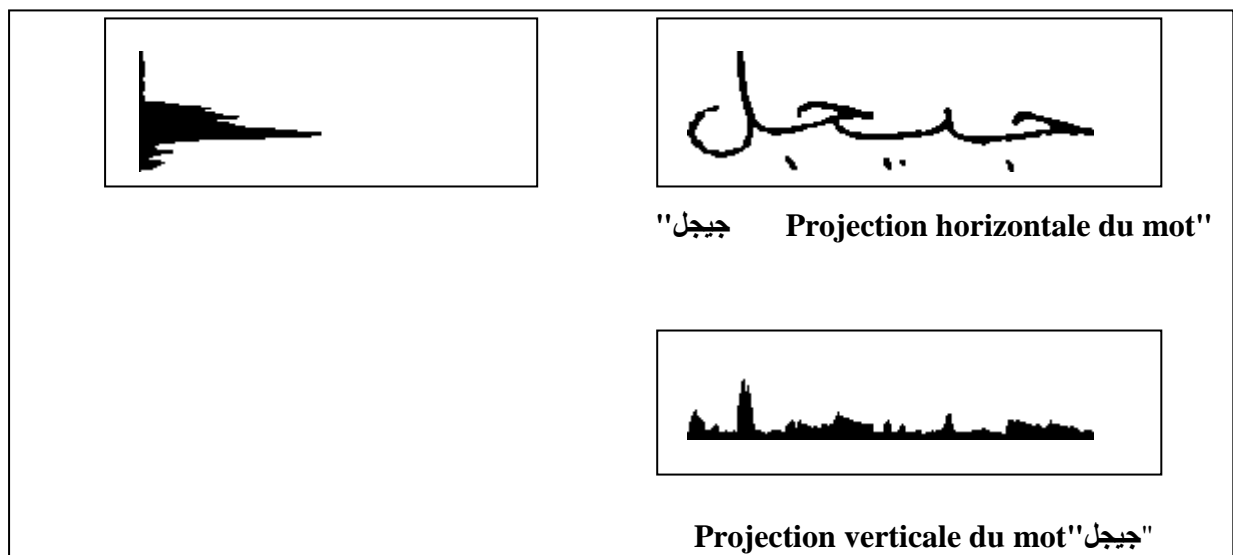
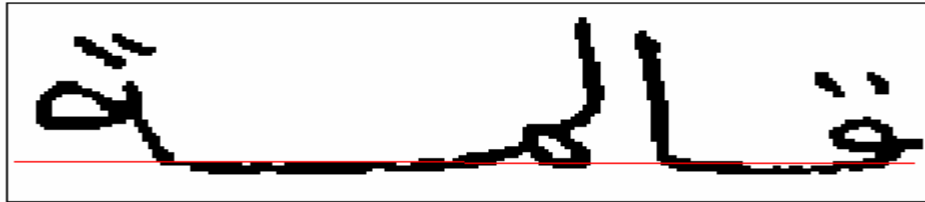


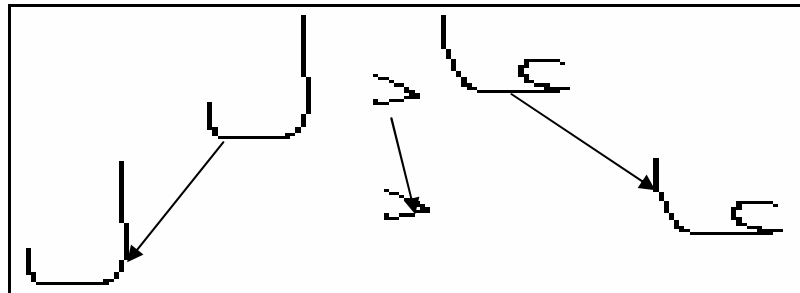
Figure III.9.Projection verticale et horizontale du mot "جيجل"

✚ La ligne médiane (de jonction) : C'est la ligne où la concentration de pixels noirs est plus forte.



**Figure III.10. La ligne médiane**

- ✚ *La ligne pixel* : C'est l'ensemble de pixels adjacents horizontalement pouvant prendre chacun la valeur 0 ou 1.
- ✚ *La ligne de texte* : C'est un ensemble de lignes de pixels se trouvant les uns au dessous des autres.
- ✚ *Partie connexe* : C'est ensemble de pixels noirs adjacents les uns aux autres elle peut représenter un mot ou une partie du mot.



**Figure III.11. Les parties connexes d'un mot**

- ✚ *Le nombre de transitions verticales (0,1) ou (1,0)*

C'est le nombre de transitions verticales (0,1) ou (1,0) pour chaque colonne de la partie connexe (1 : noir, 0 : blanc).

### **II.3. Segmentation du mot**

#### ***Description de la méthode***

Elle est subdivisée en trois étapes :

- La segmentation horizontale.
- La segmentation verticale.
- La segmentation en caractère.

#### **II.3.1. La segmentation horizontale**

Elle a pour but de localiser les lignes de texte en projetant horizontalement la page de mot sur un axe vertical. L'histogramme ainsi obtenu présente des zones non nulles qui représentent les

lignes pixels qui rentrent dans la constitution de la ligne de texte et des zones nulles qui représentent les interlignes.

Le début de la ligne correspond à la première ligne de la matrice binaire qui possède au minimum un pt noir.

La fin de la ligne de texte correspond à la ligne de la matrice binaire qui ne possède aucun point noir (le balayage se fait de droite à gauche). (Figure III.12)

### Algorithme

Début

Tant que non fin ligne de l'image (matrice binaire)

Faire

Tant que la ligne de l'image ne contient aucun point noir

Faire

Saut vers la ligne suivante ;

Fait ;

Sauvegarder le début de la ligne de l'image (ligne pixel) ;

Tant que la ligne de l'image contient un point noir au minimum

Faire

Saut vers la ligne suivante ;

Fait ;

Sauvegarder la fin de la ligne de l'image (ligne pixel) ;

Fait ;

Fin.

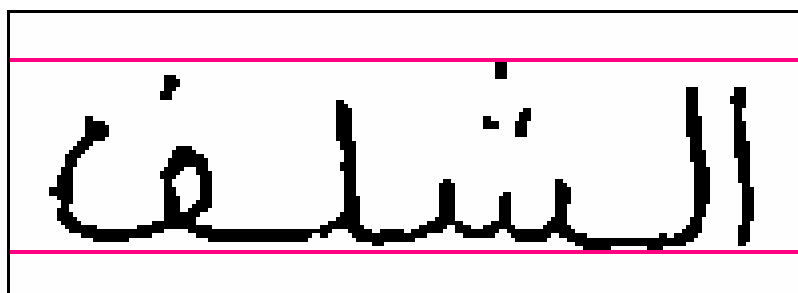


Figure III.12.Segmentation horizontale

### II.3.2. La segmentation verticale

Cette méthode a pour but la séparation du mot en parties connexes en parcourant l'histogramme vertical colonne par colonne.

Le début de la partie connexe correspond à la première colonne de la matrice binaire qui contient au minimum un point noir, la fin de la partie connexe correspond à la première colonne de la matrice binaire qui contient aucun point noir. (Figure III.13)

### Algorithme

Début

Tant que la colonne de l'image ne contient aucun point noir

Faire

Saut vers la colonne suivante ;

Fait ;

Sauvegarder la colonne début de la partie connexe ;

Tant que la colonne de l'image contient un point noir au minimum

Faire

Saut vers la colonne suivante ;

Fait ;

Sauvegarder la colonne fin de la partie connexe ;

Fin



Figure III.13.Segmentation verticale

### II.3.3. Segmentation par étiquetage en composantes connexes

Cette technique consiste à regrouper les pixels voisins dans un ensemble (partie connexe) chaque ensemble est disjoint des autres.

Nous présentons ici quelques unes des méthodes d'étiquetage en composantes connexes.

#### II.3.3.1. Etiquetage à l'aide d'une table d'équivalence « l'algorithme de Rosenfeld » [37]

Cet algorithme ne nécessite que deux balayages de l'image, mais requiert l'utilisation d'une table d'équivalence. C'est la solution la plus couramment employée de par sa simplicité de mise en œuvre.

Lors d'un premier balayage de l'image, une nouvelle étiquette est assignée à chaque point rencontré non encore étiqueté et sans voisin étiqueté. Si le voisinage du point courant comporte un

ou plusieurs points déjà étiquetés, la valeur minimum de ces étiquettes est choisie et les conflits sont stockés dans la table d'équivalence. Les équivalences sont Résolues à la fin du premier balayage et un second balayage attribue à chaque pixel son étiquette définitive.

En bref, cet algorithme est exécuté en deux phases :

- **la première phase**

Consiste un balayage de l'image ligne par ligne pendant lequel on affecte une étiquette temporaire à chaque pixel.

- **la deuxième phase**

Comporte la résolution des équivalence et un deuxième balayage de l'image pour l'attribution des étiquettes définitives aux pixels.

Avant la description le processus d'étiquetage, voilà quelques définitions :

### **Equivalence**

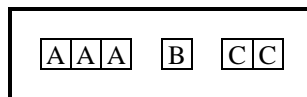
Une étiquette 'a' est dite équivalente à l'étiquette 'b' s'il existe un chemin reliant le point correspondant à 'a' au point correspondant à 'b', tel que ce point ne passe pas par 0.

### **Table d'équivalence**

C'est un tableau dynamique dont la taille augmente pendant le parcours séquentiel de l'image selon le besoin (lors de l'apparition de nouvelles étiquettes et lors de l'insertion des étiquettes équivalentes à d'autres étiquettes).

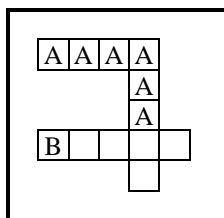
### **Le processus d'étiquetage**

En parcourant une ligne horizontale de gauche à droite, on associe un numéro (une étiquette) à chaque pixel de telle sorte que tous les pixels voisins portent le même numéro (le numéro zéro est réservé pour un pixel "vide"). Lorsque sur cette ligne, le voisinage est interrompu, puis reprend plus loin, le numéro est incrémenté de 1. Les étiquettes sont représentées par les lettres A, B et C sur la (Figure III.14).



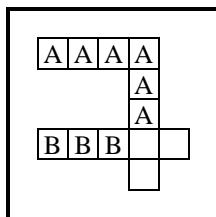
**Figure III.14. Propagation de l'étiquetage des pixels de gauche à droite sur une ligne horizontale**

Lorsqu'une nouvelle ligne est commencée, on propage naturellement l'étiquetage de haut en bas en recopiant le numéro du pixel qui se trouve au-dessus du premier pixel de la nouvelle ligne. S'il n'y a pas de pixel au-dessus, un nouveau numéro est utilisé (Figure III.15).



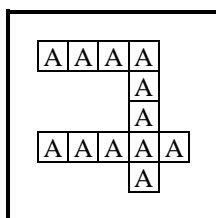
**Figure III.15. Propagation de l'étiquetage des pixels de haut en bas sur une colonne verticale**

Lorsqu'un conflit se présente entre la propagation horizontale et la propagation verticale des étiquettes, deux cas se présentent alors (Figure III.16) :



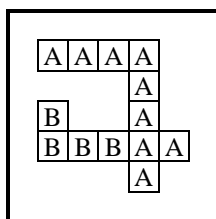
**Figure III.16. Conflit entre la propagation horizontale et verticale**

**1<sup>er</sup> cas :** si le numéro des pixels horizontaux correspond à une nouvelle étiquette, alors il est facile de résoudre le conflit en remplaçant la nouvelle étiquette de tous les pixels à gauche du point de conflit par le numéro prioritaire du pixel de la ligne précédente (Figure III.17). La nouvelle étiquette est alors annulée.



**Figure III.17: Résolution du conflit correspondant au 1<sup>er</sup> cas**

**2<sup>ème</sup> cas :** Si le numéro des pixels horizontaux a déjà été propagé à partir de la ligne précédente, il serait trop long de remplacer tous les pixels correspondants. Aussi, on note dans un tableau que les deux étiquettes en conflit désignent une unique composante connexe (Figure III.18).



**Figure III.18. Résolution du conflit correspondant au 2<sup>ème</sup> cas**

Deux étiquettes en conflit désignent une même composante connexe.

### Algorithme du premier balayage

Début

0 : la couleur blanche ; 1 : la couleur noire ; img : la matrice de l'image binaire ; etiq=0 ;

Si (img(i , j) = 0 ) alors

Si (tous les prédécesseurs de img sont 0) alors

img(i , j) = etiq ++ ;

Sinon

Si tous les prédécesseurs de img(i,j) ont la même etiq « e » alors

img(i,j) = 'e'

Sinon

e1=min ; img ( i , j) =e1 ;

Mise à jour de la table d'équivalence;

Fsi ;

Fsi ;

Fsi ;

Fin.

Dans cet algorithme, on considère pour chaque point p ses voisins déjà traités, on travaille soit par la 4-connexité soit par la 8-connexité.

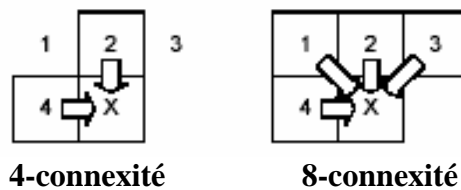


Figure III.19. Voisinages 4-connexité et 8-connexité

Au fur et à mesure qu'on progresse dans le balayage de l'image, de nouvelles étiquettes apparaissent par conséquent on alloue une nouvelle colonne pour chacune d'elles.

### La mise à jour de la table d'équivalence

Lors du premier balayage et en arrivant à un pixel à étiqueter dont les prédécesseurs sont différents on prend la dernière étiquette de la colonne (au niveau de la table d'équivalence)

Correspondant à l'étiquette de l'un des trois prédécesseurs ,ensuite on récupère la dernière étiquette des colonnes correspondant aux étiquettes des autres prédécesseurs après une comparaison entre ces étiquettes , l'étiquette la plus basse est attribuée en pixel courant(min).



### Algorithme de la résolution de la table d'équivalence

Début

Pour i = 0 à max-ligne faire

Pour j=0 à max-eti faire

Si (t [i+1, j] != 0) alors

Pour k =0 à max-eti Faire

Si (T [0, k] == T [0, j]) alors T [0, k] == T [i+1, j]

Fsi ;

Fait ;

Fsi ;

Fait ;

Fait ;

Fin.

A la fin de cette étape les différentes étiquettes à effectuer aux pixels de l'image se trouvant à la première ligne.

### Algorithme du deuxième balayage

Ce second balayage de l'image exploite la table d'équivalence pour affecter à chaque pixel l'étiquette finale qui lui est équivalente.

### Algorithme

Début

Tant que non fin de la ligne courante Faire

Si ( img(j , i) != 1) alors img(i , j) = T[0 , img(i , j) ] ;

Fait;

Fin.

1	1	1	0	1	1	0	1
1	1	0	1	1	0	1	1
1	0	1	1	0	1	1	1
1	1	1	0	1	1	1	1
0	0	0	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

**Image de depart**

1	1	1	0	2	2	0	3
1	1	0	4	2	0	5	3
1	0	6	2	0	7	3	3
1	1	1	0	8	3	3	3
0	0	0	9	3	3	3	3
10	10	10	3	3	3	3	3
10	10	10	3	3	3	3	3
10	10	10	3	3	3	3	3

**Image des étiquettes après premier balayage**

1	2	3	4	5	6	7	8	9	10
0	0	0	2	3	2	3	3	3	3
0	0	0	0	0	1	0	0	0	0

**Table d'équivalence**

1	1	3	1	3	1	3	3	3	3
0	0	0	2	3	2	3	3	3	3
0	0	0	0	0	1	0	0	0	0

**Résolution de la table d'équivalence**

1	1	1	0	1	1	0	3
1	1	0	1	1	0	3	3
1	0	1	1	0	3	3	3
1	1	1	0	3	3	3	3
0	0	0	3	3	3	3	3
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3

**Image finale**

Le problème de cet algorithme est qu'il est nécessaire de parcourir de nombreuses fois la table d'équivalence, aussi bien lors de l'étiquetage que lors de la résolution des équivalences. De plus, le nombre de parcours est inconnu, ce qui ne permet pas de déterminer précisément la complexité de cet algorithme.

### **II.3.3.2. Amélioration de la méthode d'étiquetage**

Cet algorithme peut être vu comme une évolution du précédent. Au lieu de résoudre les équivalences à la fin de l'étiquetage, celles-ci sont en partie résolues durant le premier balayage car les comparaisons ne sont pas faites entre les étiquettes, mais entre les étiquettes pointées par les étiquettes du voisinage du pixel courant. Il en résulte que la relation d'équivalence est construite de « proche en proche ».

La résolution de la classe d'équivalence se fait alors en une seule passe, sans jamais devoir revenir en arrière.

Soient les pixels au voisinage du point considéré  $p_2, p_4$  et leurs étiquettes  $e_2, e_4$ .

Soit *max-etiq* le nombre total d'étiquettes créées à la fin du balayage.

### Algorithme du premier balayage

Début

{0 : la couleur blanche ;

1 : la couleur noire ;

img : la matrice de l'image binaire ;

etiq :=0 ;

{

Si (img(i,j) != 0 ) alors

Si (tous les prédécesseurs de img sont 0) alors

img(i,j) = etiq++ ;

Sinon

Si tous les prédécesseurs de img(i,j) ont la même etiq « e » alors

img(i,j) = 'e'

Sinon

$P_2 = T[e_2]$  ;

$p_4 = T[e_4]$  ;

$e = \text{MIN}(p_1, p_2)$  ;

$T[T[e_2]] = e$  ;

$T[T[e_4]] = e$  ;

fsi;

fsi;

fsi;

}

Fin.

### Algorithme de la résolution de la table d'équivalence

Début

{

Pour k = 1 à *max-etiq* Faire

$T[k] = T[T[k]]$  ;

Fait

}

Fin.

### Algorithme du deuxième balayage

Début

Tant que non fin de la ligne courante Faire

Si  $(img(j,i) \neq 0)$

alors

$img(i,j) = T[img(i,j)]$  ;

fsi;

Fait;

Fin.

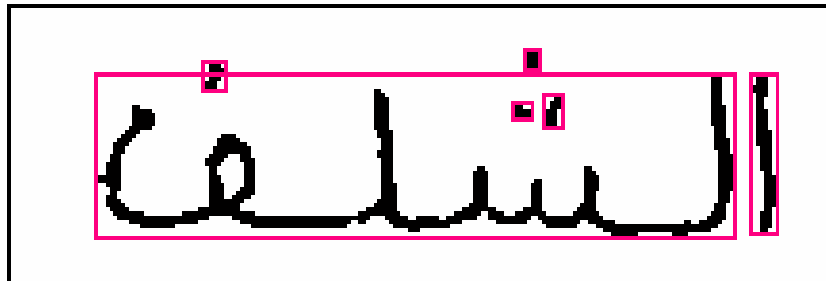


Figure III.20. Etiquetage en composantes connexes

### II.3.4. La segmentation en caractère

Le but de cette étape est l'extraction des caractères de chaque partie connexe obtenue dans la segmentation verticale.

On définit quelques notions :

- **Seuil de segmentation** : Correspond à la valeur la plus répétitive parmi les valeurs d'histogramme des colonnes (verticale), il varie d'une partie connexe à une autre.
- **Ligne de haut** : Elle est déduite par un balayage vertical commençant par la première ligne haute on détermine la ligne qui contient le premier pixel noir.
- **Ligne de bas** : Elle est déduite par un balayage vertical commençant par la dernière ligne de bas, on détermine la ligne qui contient le premier pixel noir.

### II.3.5. La segmentation après squelettisation

Cet algorithme repose essentiellement sur le squelette du texte et la détection des points de branchement ou de croisement qu'on appelle « *les points essentiels* ».

- L'idée générale de ce processus est de faire la segmentation entre chaque deux points de branchement ou croisement et la colonne de segmentation doit contenir un seul pixel.

Après avoir extrait le squelette du texte on applique l'algorithme de segmentation suivant :

### Algorithme

#### Au début de la composante connexe

On parcourt la partie connexe de droite à gauche jusqu'à trouver un point essentiel soit **ColBran1** la colonne contenant ce point.

Si le nombre de pixel maximal entre **ColBran1** et la colonne extrême droite de la composante connexe est supérieur à 1 alors on parcourt la partie connexe de droite à gauche jusqu'à trouver une colonne contenant un seul pixel noir et on fait la segmentation.

#### Au milieu de la composante connexe

On continue le parcourt de la partie connexe de droite à gauche jusqu'à trouver le deuxième point essentiel soit *ColBran2* la colonne contenant ce point.

Si le nombre de pixel minimal entre *ColBran1* et *ColBran2* est égale à 1 alors on parcourt la partie connexe de droite à gauche jusqu'à trouver une colonne contenant un seul pixel noir et on fait la segmentation.

#### A la fin de la composante connexe

Entre le dernier point essentiel et la colonne extrême gauche de la composante :

On calcule la distance *DH* : la distance entre le point de branchement et la fin de la partie connexe.

On calcule la distance *DV* : la distance entre la ligne contenant le point de branchement et la ligne contenant le point extrême.

**Si  $DV > DH$  alors** on fait un retour en arrière jusqu'à trouver une colonne contenant un seul pixel noir, et on coupe au niveau de cette colonne ;

**Si** à la fin de la composante connexe on n'a trouvé aucun point de branchement ou de croisement, on calcule la longueur et la largeur de la partie connexe ;

**Si** la longueur  $>$  la largeur **alors** on parcourt la partie connexe de droite à gauche jusqu'à trouver une colonne contenant un seul pixel noir, et on coupe cette colonne ;

**Si** entre les deux points extrêmes, il existe une colonne contenant plus d'un pixel noir **alors**

On parcourt la partie connexe de droite à gauche jusqu'à trouver une colonne contenant un pixel noir et on coupe la colonne ;

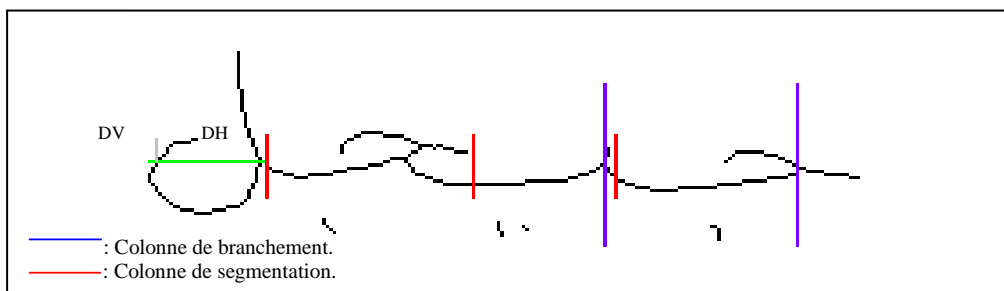


Figure III.21. Le mot "جيجل" segmenté

### II.3.6. Le problème de la sur-segmentation et la solution proposée

La segmentation obtenue segmente les caractères "س" et "ش" en trois parties.

La solution proposée de ce problème:

- On remarque que le caractère "س" contient deux ou trois colonnes de branchement avec l'absence totale des points diacritiques et pas de boucle, on prend ces propriétés en considération pour masquer les deux colonnes de branchements qui sépare le caractère en trois parties.
- Pour le caractère "ش" la même chose que le caractère "س" avec la présence de trois points en haut soit attachées soit séparées.

### III.Extraction des primitives

L'extraction de caractéristiques est de fait l'une des deux fonctions essentielles d'un système de reconnaissance d'écriture. Elle comprend la mesure des caractéristiques de la forme d'entrée (nettoyée) qui sont pertinentes à la classification. Lorsque l'extraction des caractéristiques est terminée, la forme est représentée par l'ensemble des caractéristiques extraites.

Il existe un nombre important de caractéristiques possibles que l'on peut extraire d'une forme finie à deux dimensions. Il faut toutefois ne s'attarder qu'aux caractéristiques qui ont une pertinence possible pour la classification, ce qui suppose qu'au cours de la période de conception, le spécialiste s'attarde aux caractéristiques qui, selon une certaine technique de classification, apporteront les résultats les plus certains et les plus efficaces.

Nous traiterons de l'extraction des primitives dans le cas général, c'est-à-dire quel qu'en soit le niveau (lettre, mot) en présentant une synthèse des méthodes employées que nous avons élaborée à partir d'une étude bibliographique sur plusieurs systèmes de reconnaissance.

Au cours de l'extraction des primitives, plusieurs objectifs, qui précèdent la reconnaissance, peuvent être envisagés. Les principaux objectifs que nous définirons dans la perspective de la reconnaissance de l'écriture manuscrite sont : l'analyse, le paramétrage, la modélisation, le codage et la classification.

- **L'Analyse**, dont la définition littérale est "la décomposition d'un tout en ses parties", consiste généralement en l'extraction d'un ensemble d'attributs caractéristiques du texte. Concrètement, l'analyse d'un texte manuscrit consiste à recueillir des informations statistiques telles que la disposition des lignes d'écriture, leur orientation, leur régularité, l'espacement des mots et des lettres, la régularité et l'inclinaison des lettres, l'épaisseur du trait, ainsi que la ligature des lettres à l'intérieur des mots.
- **Le Paramétrage** consiste à établir une liste d'attributs représentés par une variable binaire (attribut présent ou absent) ou multivaluée (proportionnelle à l'importance de l'attribut), qui

ont été détectés et évalués dans l'image. A la différence de l'analyse, le paramétrage ne concerne qu'un mot ou qu'une lettre en vue de sa reconnaissance.

- **La Modélisation** est la construction d'une représentation approximative de la forme entière. A la différence du paramétrage, l'objectif est la réduction de l'information utile, au minimum nécessaire pour représenter complètement la forme, en particulier son aspect et sa structure. Si la modélisation est une description proche d'aspect de la forme originale, on parle alors d'une schématisation, avec une approximation plus ou moins importante. Sinon, il s'agit d'un codage de l'information.
- **La classification** peut être considérée comme une identification partielle de l'information. L'objectif de la classification est, lorsque l'on ne dispose pas de toute l'information nécessaire à l'identification complète de la forme (cas de la reconnaissance), de déterminer quand même une catégorie à laquelle elle appartient. Le principe est que moins d'information sont nécessaire pour distinguer les caractères que pour les reconnaître.

### **III.1.La problématique de l'extraction de l'information de l'extraction des primitives**

L'objectif de l'étape de l'extraction de l'information est la sélection de l'information pertinente qui se trouve noyée dans la masse de l'information brute acquise. La problématique de cette étape a pour origine le risque de perte d'information signifiante pour la reconnaissance. La minimisation de ce risque est conditionnée par deux dilemmes liés chacun à un paradoxe de causalité. Il s'agit du dilemme de la réduction et du dilemme de la segmentation.

Nous allons développer dans cette partie la nature de ces dilemmes en présentant les différentes origines de perte d'information lors du processus de la reconnaissance de l'écriture manuscrite.

La première étape du processus est l'acquisition de l'information. Il est évident que l'on ne peut pas extraire une information qui n'est pas présente dans le tracé. C'est pourquoi, avant de traiter de la perte d'information pendant la phase d'acquisition, nous considérerons la perte d'information pendant la phase de production de l'écriture.

#### ➤ **La perte d'information pendant la phase de production de l'écriture**

Cette première origine de perte d'information, qui a lieu pendant l'écriture elle-même, est due au facteur humain. Cette question, qui s'écarte un peu de notre sujet, a cependant des conséquences déterminantes sur la reconnaissance.

Le caractère subjectif de l'écriture tient au fait que le lecteur étant aussi le scripteur, il peut ne pas s'apercevoir que son écriture est illisible !

L'auteur a en outre la liberté de s'appliquer à ne bien tracer que les lettres permettant de distinguer les mots. Cette attitude a une conséquence importante sur la stratégie de lecture à adopter

pour la reconnaissance. En effet, elle implique d'avoir la connaissance du vocabulaire parmi lequel l'auteur a estimé pouvoir distinguer les mots.

La limite à considérer pour ce type de perte d'information est que si un humain ne peut pas lire un texte, un ordinateur ne le pourra probablement pas non plus. Cela fixe une borne inférieure à la qualité d'écriture admissible pour un système de reconnaissance.

➤ **La perte d'information pendant la phase d'acquisition**

L'objectif de cette étape est d'acquérir le maximum d'informations, afin d'obtenir des images numériques les plus précises. La perte d'information pendant l'acquisition est liée le plus souvent à la faiblesse des capteurs (scanner ou éventuellement caméra), ou aux conditions d'acquisition.

A ce niveau du processus, il reste à déterminer si l'information signifiante est accessible ou non parmi toute l'information acquise.

➤ **La perte d'information pendant la phase de réduction**

La phase de réduction consiste à extraire l'information signifiante qui est noyée dans le bruit. Dans le cas de la classification, l'objectif est d'éliminer davantage de bruit que d'information afin de faciliter la discrimination. Dans le cas de l'approximation, l'objectif est de modéliser en minimisant la perte d'information.

Une autre perte d'information se produit également au cours de la phase de segmentation.

➤ **La perte d'information pendant la phase de segmentation**

Le problème de la segmentation nous ramène au paradoxe de causalité. Pour pouvoir segmenter le mot en lettres afin de les reconnaître isolément, il est nécessaire au préalable de les localiser, or comment localiser ces lettres sans les avoir reconnues auparavant?

### **III.2. Les approches de l'extraction des primitives**

En fonction de l'objectif fixé et de la méthode d'extraction choisie, l'approche de l'extraction des primitives peut être systématique ou heuristique.

- La modélisation et le codage conduisent à une approche systématique dans la mesure où l'objectif fixé est la détermination d'une représentation complète de la forme, même de façon approximative. Dans la modélisation, les primitives sont obtenues a posteriori, par le résultat de l'approximation, tandis que, en ce qui concerne le codage, les catégories de primitives sont définies a priori. Un test, qui est par exemple réalisé à l'aide d'une sonde, permet de valider la présence de chacune des primitives sur l'ensemble de la forme
- Le paramétrage conduit plutôt à une approche heuristique. Dans ce cas, on ne cherche pas nécessairement une représentation complète mais seulement des indices significatifs. De même que dans le cas du codage, ces indices sont des primitives définies a priori.



### **III.3. Les catégories de primitives**

Nous avons distingué quatre catégories principales de primitives : les primitives topologiques, statistiques, structurelles et globales.

#### **III.3.1. Les primitives topologiques ou métriques**

Le terme métrique désigne la mesure d'une distance. La topologie est "l'étude des propriétés de l'espace (et des ensembles) du seul point de vue qualitatif". Concrètement, la topologie consiste, à l'aide de sondes appliquées directement sur l'image du caractère, à effectuer par exemple sur l'échantillon les mesures et les tests suivants :

- Compter dans une forme le nombre de trous.
- Evaluer les concavités.
- Mesurer des pentes et autres paramètres de courbures et évaluer des orientations principales.
- Mesurer la longueur et l'épaisseur des traits.
- Détecter les croisements et les jonctions des traits.
- Mesurer les surfaces, les périmètres.
- Déterminer le rectangle délimitant l'échantillon, ou le polygone convexe.
- Evaluer le rapport d'élongation (ou allongement) longueur/largeur,...
- Rendre compte de la disposition relative de ces primitives.
- Descripteur de FOURIER ces caractéristiques sont basées sur des informations extraites à partir de l'image du contour.
- Description des profils des caractères.

#### **III.3.2. Les primitives structurelles**

A la différence des primitives topologiques, les primitives structurelles sont généralement extraites non pas de l'image brute, mais à partir d'une représentation de la forme par le squelette ou par le contour. Ainsi, on ne parle plus de trous, mais de boucles. Cependant, pour le reste, les primitives structurelles correspondent à peu près aux primitives topologiques, il s'agit principalement :

- Des segments de droite.
- Des arcs, boucles et concavités, des pentes.
- Des angularités, points extremum et points terminaux, jonctions et croisements.
- Remplissage des coins.

#### **III.3.3. Les primitives globales**

Elles sont naturellement basées sur une transformation globale de l'image. La caractéristique d'une primitive globale est de dépendre de la totalité des pixels d'une image (transformées de Hough et de Fourier).

### III.3.4. Les primitives statistiques

Elles véhiculent une information qui est distribuée sur toute l'image. L'histogramme, qui représente le nombre de pixels sur chaque ligne ou colonne de l'image, en est un exemple classique et simple à calculer. L'histogramme des transitions, comme l'indique son nom, ne retient que le nombre des transitions 0-1 et 1-0.

On utilise aussi une autre primitive statistique basée sur un moyennage des pixels situés à l'intérieur d'un masque rectangulaire. La construction d'une matrice de masque recouvrant la totalité de la forme permet une représentation statistique à partir d'un nombre très réduit de valeurs correspondant à chaque masque.

Les moments invariants appartiennent eux aussi primitive statistique permet d'extraire des paramètres propres à la forme à reconnaître.

### III.4. Présentation des méthodes d'extraction des primitives

La détection des caractéristiques est l'une des étapes les plus délicates dans la construction d'un système de reconnaissance d'écriture arabe manuscrite. Nous présentons dans cette partie les principales caractéristiques et leurs détections.

#### III.4.1. Les mesures topologiques

Parmi ces mesures élémentaires, nous trouvons :

- La surface (égale, par exemple, au nombre de pixels situés à l'intérieur de la frontière de l'objet).
- Le périmètre (égal au nombre de pixels situés le long de la frontière de l'objet).
- Le nombre de cavités (égal au nombre de trous ou zones composées de pixels n'ayant pas les mêmes niveaux de gris que les pixels de l'objet).

#### III.4.2. Les mesures d'orientation

Pour les régions présentant des allongements caractéristiques, il est possible de calculer les directions suivant lesquelles ces allongements ont été réalisés. Ces directions sont calculées à partir des moments d'inertie centrés du deuxième ordre :

$$\begin{aligned}\mu_{xx} &= \sum_{(i,j) \in R} (j - j_G)^2, \\ \mu_{yy} &= \sum_{(i,j) \in R} (i - i_G)^2, \\ \mu_{xy} &= \sum_{(i,j) \in R} (i - i_G)(j - j_G),\end{aligned}$$

Où  $i_G$  et  $j_G$  représentent les coordonnées du centre de gravité de la région  $R$  et  $(i, j)$  les coordonnées des pixels. La direction  $\alpha$  de l'axe d'inertie principal de  $R$  se déduit de la formule suivante :

$$tg(2\alpha) = \frac{2\mu_{xy}}{\mu_{yy} - \mu_{xx}}$$

Cette orientation n'a de sens que si la région ne présente pas de symétrie de révolution, c'est-à-dire si  $\mu_{yy} \neq \mu_{xx}$ .

### III.4.3. Les mesures de forme

En voit ici deux exemples de nature globale :

- La compacité :  $C = \frac{S}{P^2}$ .

Où  $S$  est la surface et le périmètre. En théorie, cette valeur est maximale pour les cercles et plus petite pour les autres formes. Mais, à cause de la discrétisation, cette valeur est également maximale pour les carrés ou les polygones à peu près réguliers.

- L'allongement :  $C = \frac{long}{larg}$ .

Où *long* correspond à la plus grande dimension de la région, et *larg* à la plus petite.

### L'approche VH2D

L'approche VH2D [37] (vertical-Horizontal-2Diagonal), proposée par Cheng et Xia, consiste à projeter chaque caractère sur l'abscisse, l'ordonnée, ainsi que sur les diagonales 45° et 135°. Les projections s'effectuent en calculant la somme des valeurs des pixels  $i_{xy}$  selon une direction donnée.

### Définitions

- **Projection verticale**

La projection verticale d'une image  $I = (i_{xy})$  (de dimension  $N \times N$ ) représentant un caractère  $C$  est dénotée par :

$$P_y^v = \sum_{x=1}^N i_{xy} \quad \text{où} \quad P^v = [P_1^v, P_2^v, \dots, P_N^v, ]$$

- **Projection horizontale**

La projection horizontale d'une image  $I = (i_{xy})$  (de dimension  $N \times N$ ) représentant un caractère  $C$  est dénotée par :

$$P_y^h = \sum_{y=1}^N i_{xy} \quad \text{où} \quad P^h = [P_1^h, P_2^h, \dots, P_N^h, ]$$

- **Projection sur la diagonale 45°**

La projection sur la diagonale 45° d'une image  $I = (i_{xy})$  (de dimension  $N \times N$ ) représentant un caractère C est dénotée par :

$$P^{d1} = [P_1^{d1}, P_2^{d1}, \dots, P_{2N-1}^{d1}, ] \quad \text{où :}$$

$$P_m^{d1} = \begin{cases} \sum_{l=N-m+1}^N \sum_{k=1}^m i_{lk} & 1 \leq m \leq N \quad \text{et } l = k + N - m \\ \sum_{l=1}^{2N-m} \sum_{k=m-N+1}^N i_{lk} & N+1 \leq m \leq 2N-1 \quad \text{et } l = k + N - m \end{cases}$$

- **Projection sur la diagonale 135°**

La projection sur la diagonale 135° d'une image  $I = (i_{xy})$  (de dimension  $N \times N$ ) représentant un caractère C est dénotée par :

$$P^{d2} = [P_1^{d2}, P_2^{d2}, \dots, P_{2N-1}^{d2}, ] \quad \text{où :}$$

$$P_m^{d2} = \begin{cases} \sum_{l=1}^m \sum_{k=1}^m i_{lk} & 1 \leq m \leq N \quad \text{et } k = m - l + 1 \\ \sum_{l=m-N+1}^{2N-m} \sum_{k=m-N+1}^N i_{lk} & N+1 \leq m \leq 2N-1 \quad \text{et } k = m - l + 1 \end{cases}$$

### III.4.4.Extraction des profils

Le caractère est encadré dans une fenêtre minimale. Le profil compte le nombre de pixel entre la boîte de bondissement de l'image de caractère et le bord de caractère. Le profil d'un caractère peut être pris à n'importe quelle position, supérieur, inférieur, gauche, droite, et orienté. . Les profils décrivent bien les formes externes des caractères et laissent distinguer entre un grand nombre de lettres.

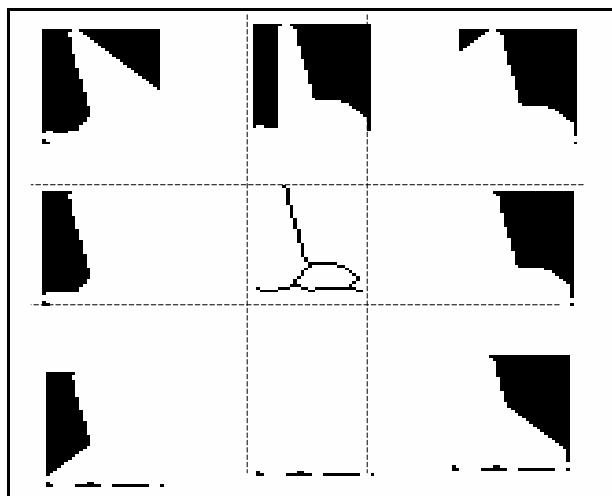


Figure III.22.Les huit profils de la lettre "ط "

### III.4.5 Les moments invariants

La méthode des moments géométriques, comme toutes les méthodes statistiques, permet d'extraire des paramètres propres à la forme à reconnaître. Ces derniers permettent de la distinguer de toutes les autres formes qui lui sont peu semblables. Les moments invariants proposés par Hu [51] ont le mérite de répondre à trois critères qui sont :

- L'invariance par rapport à la translation.
- L'invariance par rapport au changement d'échelle.
- L'invariance par rapport à la rotation.

Si  $f(x, y)$  est une fonction continue et non nulle seulement dans une région du plan  $(x, y)$ , il existe une suite de moments géométriques d'ordre  $(p, q)$  unique définie par :

$$m_{pq} = \int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} x^p y^q f(x, y) dx dy$$

avec  $p$  et  $q \in \mathbb{N}$

Ces moments ne satisfont aucune des trois propriétés citées plus haut. Par conséquence, une deuxième série de moments invariants par rapport à la translation définie par:

$$\mu_{pq} = \int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

Tels que  $(\bar{x}, \bar{y})$  représente les coordonnées du centre de gravité et sont données par :

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}, \quad \text{et } \mu_{pq} \text{ est le moment centralisé d'ordre } (p, q).$$

Pour une image discrète les relations (3.1) et (3.2) deviennent.

$$m_{pq} = \sum_x \sum_y (x)^p (y)^q f(x, y).$$

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y).$$

Dans cette étude on se limitera aux moments d'ordre (III). En effet ils donnent une meilleure séparation des classes par aux moments d'ordre (II) puisque on a plus de paramètres pour décrire la forme à reconnaître.

Les moments centrés d'ordre 3 sont donnés par les relations suivantes :

$$\mu_{10} = m_{10} - \frac{m_{10}}{m_{00}}(m_{00}) = 0.$$

$$\mu_{11} = m_{11} - \frac{m_{10}}{m_{00}}(m_{01}).$$

$$\mu_{02} = m_{02} - \frac{m_{01}^2}{m_{00}}.$$

$$\mu_{20} = m_{20} - \frac{m_{10}^2}{m_{00}}.$$

$$\mu_{30} = m_{30} - 3\frac{m_{10}}{m_{00}}(m_{20}) + 2\frac{m_{10}^2}{m_{00}^2}(m_{10}).$$

$$\mu_{12} = m_{12} - 2\frac{m_{01}}{m_{00}}(m_{11}) - \frac{m_{10}}{m_{00}}(m_{02}) + 2\frac{m_{01}^2}{m_{00}^2}(m_{10}).$$

$$\mu_{21} = m_{21} - 2\frac{m_{10}}{m_{00}}(m_{11}) - \frac{m_{01}}{m_{00}}(m_{20}) + 2\frac{m_{10}^2}{m_{00}^2}(m_{01}).$$

$$\mu_{03} = m_{03} - 3\frac{m_{01}}{m_{00}}(m_{02}) + 2\frac{m_{01}^3}{m_{00}^2}.$$

Les moments centrés et normalisés sont définis par :

$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}.$$

$$\text{Avec : } \gamma = \frac{p+q}{2} + 1.$$

Ces derniers ont la propriété d'être invariants par rapport à la translation et le changement d'échelle. Hu a proposé dans ses travaux, une autre série de moments qui regroupent l'invariance par rapport à la translation, la rotation et le changement d'échelle.

Ils sont donnés par:

$$\phi_1 = n_{20} - n_{02}.$$

$$\phi_2 = (n_{20} - n_{02})^2 + 4n_{11}^2.$$

$$\phi_3 = (n_{30} - 3n_{12})^2 + (3n_{11}^2 - n_{03})^2.$$

$$\phi_4 = (n_{30} - 3n_{12})^2 + (n_{21} - n_{03})^2.$$

$$\phi_5 = (n_{30} - 3n_{12})^2(n_{30} - n_{12})[(n_{30} - 3n_{12})^2 - (3n_{21} - n_{03})^2] + (n_{21} - n_{03})(3n_{11}^2 - n_{12})$$

$$[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2].$$

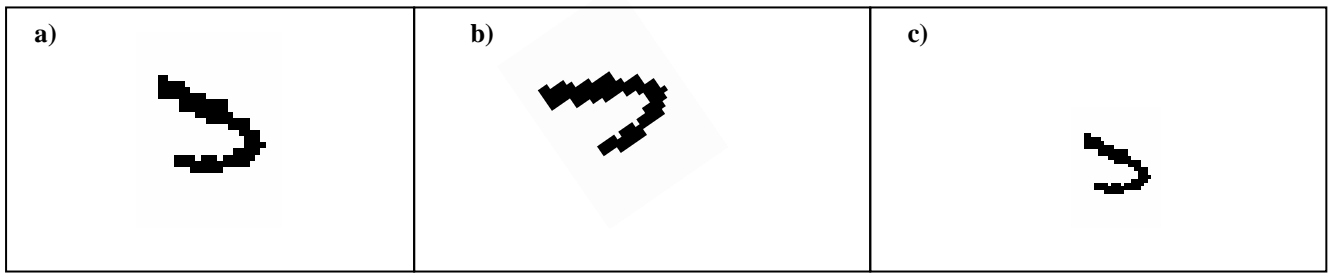
$$\phi_6 = (n_{20} - n_{02})[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03}).$$

$$\phi_7 = (3n_{21} - n_{30})(n_{30} + n_{12})[(n_{30} - 3n_{12})^2 - (3n_{21} - n_{03})^2] + 3(n_{12} - n_{30})(n_{21} - n_{12})$$

$$[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2].$$

On obtient finalement les moments invariants par translation, rotation et changement d'échelle appelé  $\phi_i$ .

La figure suivante montre une image de composant connexe « $\hookrightarrow$ » déformés par rotation, translation et par changement d'échelle ainsi les moments  $\varphi_i$ ,



**Figure III.23. Opération géométrique**

- a) Image originale.
- b) Image tournée de  $45^\circ$ .
- c) Image dont la taille est réduite par 50%.

Moments  $\varphi_i, i=1, \dots, 7$  calculés sur les quatre images décrites à la (Figure III.23).

	<b>Figure (a)</b>	<b>Figure (b)</b>	<b>Figure (c)</b>
$\varphi_1$	0.0049	0.0049	0.0049
$\varphi_2$	0.0532	0.0461	0.0058
$\varphi_3$	-0.0014	-0.0014	-0.0014
$\varphi_4$	0.0001	0.0001	0.0001
$\varphi_5$	0.0043	0.0024	-0.0181
$\varphi_6$	0.0196	0.0144	0.0062
$\varphi_7$	0	0	0

**Tableau III.2. Résultat de calcul de moments invariants d'un exemple**

En observant les moments décrits dans le (Tableau III.2), on constate une assez bonne propriété d'invariance.

### III.4.6. Détection des boucles

La détection des boucles se fait par un remplissage des endroits fermés car une boucle est considérée comme un espace pouvant contenir un liquide sans lui permettre de sortir.

- **Etape de remplissage**

Pour chaque ligne de l'image du caractère :

Colorer le segment entre chaque paire de pixels noirs ( $p_{ij}, p_{ij'}$ ) /  $j \neq j'$

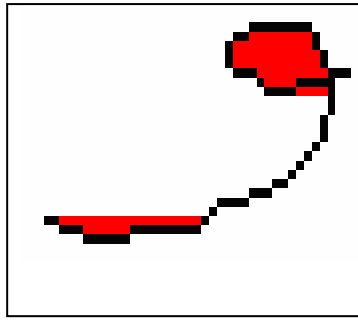


Figure III.24. Remplissage

▪ **Nettoyage**

Pour tout pixel coloré  $p_{ij}$  :

Si au moins un des quatre voisins ( $p_{i+1j}$ ,  $p_{i-1j}$ ,  $p_{ij+1}$ ,  $p_{ij-1}$ ) est blanc

Alors le pixel coloré  $p_{ij}$  devient blanc ;

L'existence des pixels colorés après le nettoyage indique la présence de boucles.

Pour affirmer qu'il existe plus qu'une boucle, il faut seulement trouver un pixel noir n'ayant que des voisins noirs ou colorés.

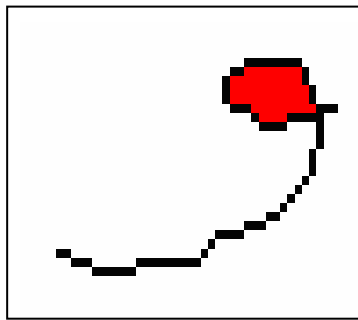


Figure III.25. Nettoyage

**III.4.7. Remplissage des coins**

On considère les cadres des quatre coins comme suit :

Et on définit les variables  $THG$   $THD$   $TBD$   $TBG$  représentant respectivement le remplissage des coins (haut gauche, haut droit, bas droit, bas gauche).

Si au moins un pixel noir se trouve à l'intérieur de l'un des quatre coins

Alors on force le variable correspondant à 1.

Pour le caractère «KEF»  $THG = 0$ ,  $THD = TBD = TBG = 1$ .

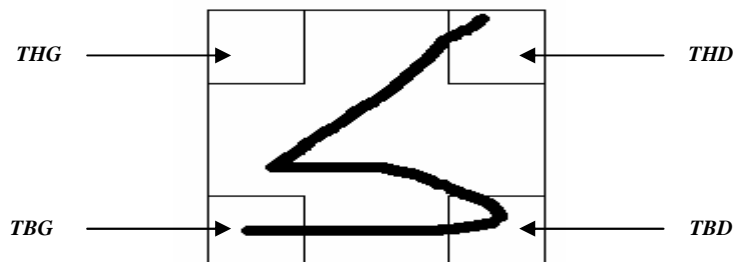
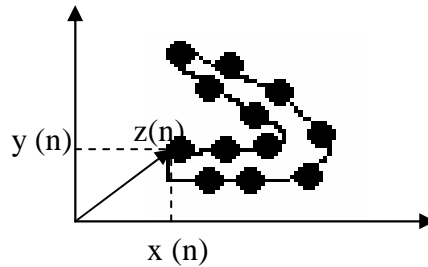


Figure III.26. Remplissage de coins pour le caractère «S»



### III.4.8.description de formes par descripteur de fourier

Le descripteur de fourier correspond à une suite de nombre dont le but est de caractériser au mieux la forme à décrire. Après une opération de suivi de contour extérieur d'un objet, on obtient la suite ordonnée de ses N points représentée sous la forme complexe FigureIII.27.



FigureIII.27. Représentation complexe d'un contour

On remarque que cette fonction est périodique de période T, ce qui formalise par la relation suivante :

$$z(n + kT) = z(n) : \text{avec } k \text{ entier.}$$

Compte tenu de la nature périodique de cette suite, on peut la représenter en utilisant la TFD.

$$z(n) = \left(\frac{1}{N}\right) \cdot \sum_{k=0}^{N-1} Z(k) \cdot e^{2j\pi kn / N}, \text{ pour } 0 \leq n \leq N - 1.$$

$$Z(k) = \sum_{n=0}^{N-1} z(n) \cdot e^{-2j\pi kn / N}, \text{ pour } 0 \leq k \leq N - 1.$$

les coefficients  $Z(k)$  ( $k=0,1,\dots,N-1$ ) désignent les descripteurs de fourier du contour. Après application de la FFT, le contour est décrit par les coefficients (descripteurs) de fourier que l'on visualise (en module) sous forme de raies. La reconstruction du contour peut être réalisée par l'application de la transformée de fourier inverse (IFFT) aux coefficients.

### III.4.9.Point caractéristique

C'est un pixel noir dans l'image amincie ayant un nombre de voisins noirs différents de 0 ou 2.

Les types de points caractéristiques sont *les points extrêmes* et *les points de jonctions*.

1. **Le point extrême** : Un point extrême correspond au début / fin d'un segment de ligne.
2. **Le point de jonction** : Un point de jonction connecte trois branches ou plus.

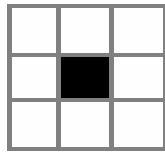
On y trouve 2 types :

Les points de *branchement* et de *croisement*.

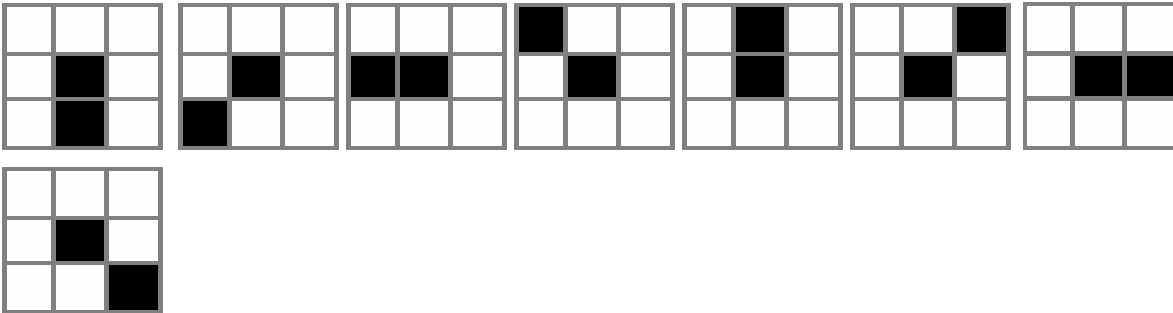
- *Point de branchement* : Correspond à un branchement de trois traits dans l'image amincie.
- *Point de croisement* : C'est un point où quatre traits se croisent.

L'extraction des points caractéristiques se fait dans un cadre de 3 x 3 pixels en calculant le nombre  $N$  de voisins noirs du pixel central pour chaque pixel de l'image amincie.

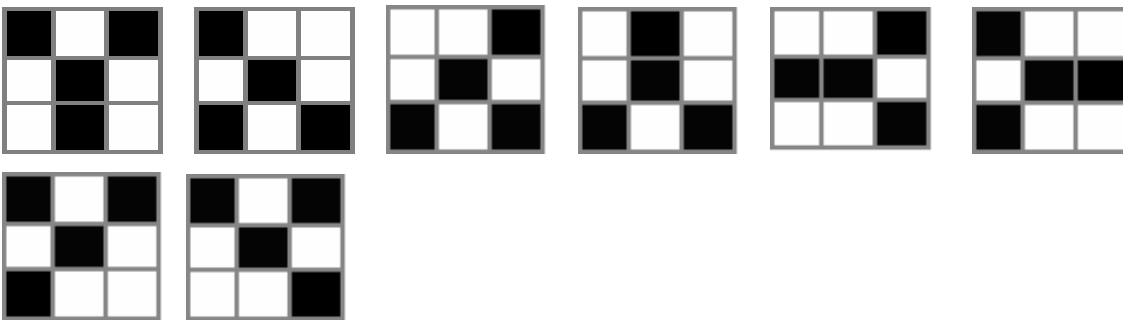
$N=0$  : alors le pixel est un point diacritique.



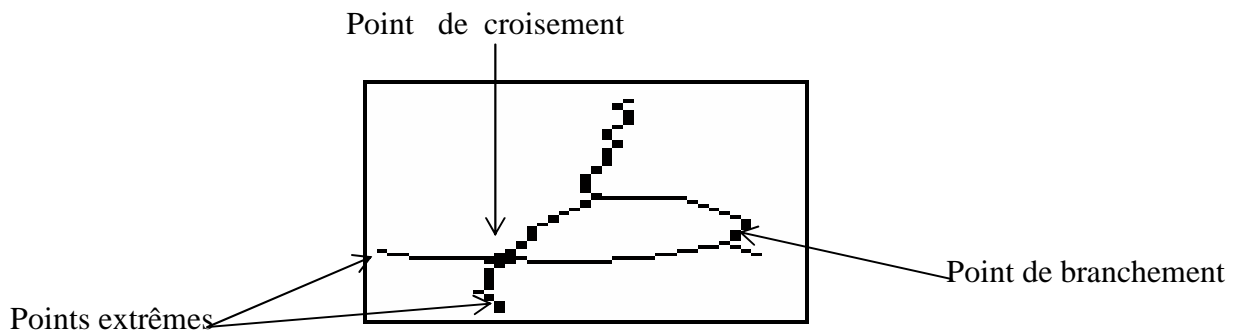
**N=1** : alors le pixel est un point extrême.



**N=3** : et le nombre de transitions circulaires (1,0) ou (0,1) dans le cadre 3\*3 du pixel sélectionné égal à 3 alors le pixel est un point de branchement.



**N=4** : et le nombre de transition circulaire (1,0) ou (0,1) dans le cadre 3\*3 du pixel sélectionné égal à 4 alors le pixel est un point de croisement.

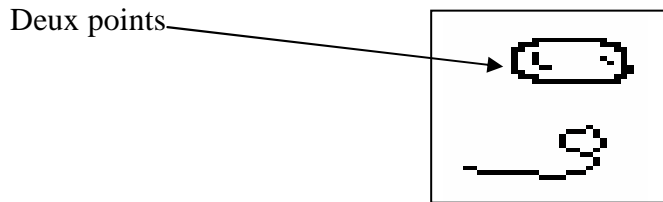


**Figure III.28. Détection des points caractéristiques**

### III.4.10. Position et nombre des points diacritiques

L'écriture arabe est riche en points diacritiques. L'écriture de ces points est généralement, respectée par les scripteurs car ces derniers permettent de distinguer entre les caractères ayant le même corps principal. Ces points peuvent être un seul point, deux, ou trois. De plus ces points sont

simples à détecter et distingués par leurs positions par rapport à la ligne de base (au-dessus ou au-dessous) (Figure III.28). Cette propriété nous a amené à détecter les parties diacritiques séparément des parties primaires, en vu d'une description globale efficace sélectionné égal à 4 alors le pixel est un point de croisement.



**Figure III.29. Détection des points diacritiques**

#### **III.4.10.1. Position des points** (y compris la « Hamza »)

Nous définissons pour cela une variable *pospoint* qui prend les valeurs suivantes :

*Pospoint* =1 pour la position des points en haut du segment.

*Pospoint* =2 pour la position des points en bas du segment.

*Pospoint* =0 si le segment n'est pas ponctué.

#### **III.4.10.2. Nombre de points**

Défini par la variable *Nombpoint* qui prend les valeurs suivantes :

*Nombpoint* = 1 le segment possède 1 point.

*Nombpoint* = 2 le segment possède 2 points.

*Nombpoint* = 3 le segment possède 3 points.

*Nombpoint* = 4 le segment possède une Hamza.

##### ▪ **Extraction et classification des tracés secondaires**

Dans les mots arabes, les parties « diacritiques » (Hamza, point,...) font partie intégrante des caractères. On ne peut pas les omettre, leur nombre et leur position au-dessus ou au-dessous du caractère, change le sens de ce dernier. Cette propriété nous a amené à détecter les parties diacritiques séparément des composantes connexes, en vu d'une description globale efficace.

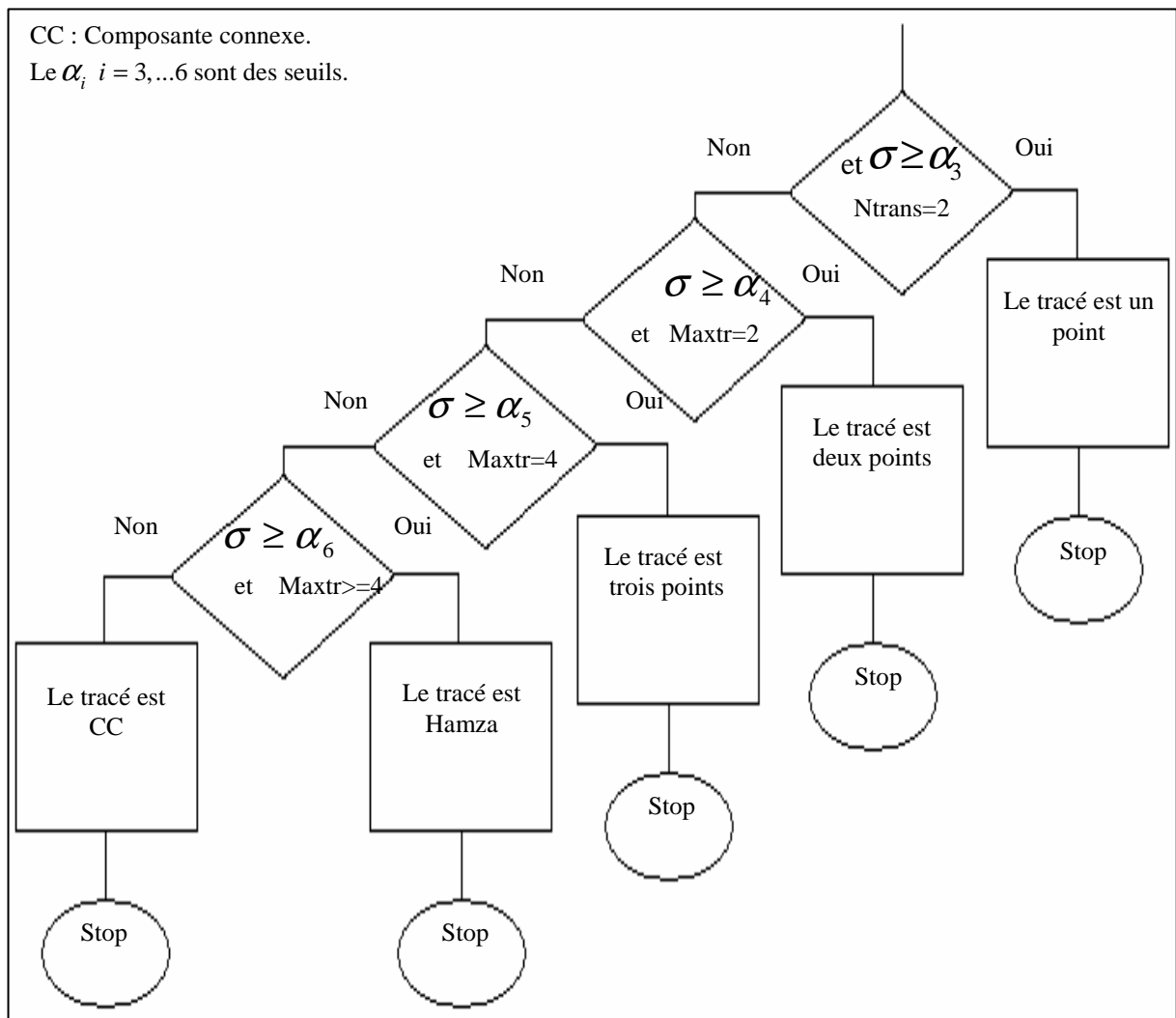
Nous avons effectué dans un premier temps, une discrimination qui sépare le mot en deux types de tracés : le premier représente les points diacritiques, le deuxième groupe les composantes connexes qui peuvent être formées d'un ou plusieurs caractères. Les tracés de types diacritiques rassemblent tout ce qui peut être considéré comme trait secondaire, c'est-à-dire Hamza, point, deux points liés ou trois points liés.

A cet effet, nous avons adopté le paramètre de compacité défini par  $\sigma = \frac{A.4\pi}{p^2}$  tel que : *A* est l'aire de l'entité connexe, et *p* son périmètre.

On définit aussi une variable **Maxtr** qui représente le nombre maximal de transitions verticales (0-1 ou 1-0) de l'entité.

La discrimination entre les tracés secondaires se fait selon un algorithme représenté par la figure

Par conséquent, la discrimination entre les tracés secondaires et principaux d'une part, et entre les tracés secondaires eux-mêmes, peut se faire par application de l'organigramme suivant sur chaque tracé dans l'image de mot :



**Figure III.30. Organigramme de classification des tracés**

Les tracés secondaires (point, deux points liés) sont examinés pour déterminer leurs positions vis-à-vis de la ligne de base. Donc, un tracé secondaire est considéré au-dessus de la ligne de base, si le centre de gravité du Bounding-Box de son contour est situé au-dessus de cette ligne. Sinon, il est considéré au-dessous.

## **IV. Conclusion**

Dans ce chapitre nous avons cité les différentes méthodes de segmentation. La méthode de segmentation horizontale et verticale n'a pas d'effet à cause du chevauchement qui pose l'écriture arabe manuscrite.

Dans notre étude, le but de la segmentation est de localiser les parties connexes afin de les segmenter en éléments plus simples qui sont les caractères. Pour cela, nous avons utilisé l'étiquetage en composantes connexes sur les mots arabes pour localiser et déterminer leurs bornes. L'algorithme de segmentation que nous avons développé se base sur l'image squelette du mot et les points de branchement ou de croisement entre caractère qu'on appelle « points de liaison ou essentiels ».

L'efficacité de cet algorithme de segmentation est liée à l'efficacité de l'algorithme de squelettisation utilisé. Dans notre cas l'algorithme de Zhang et Suen donne des bons résultats, car il conserve le maximum des propriétés topologiques et métriques de la forme.

Ensuite nous avons présenté un ensemble de caractéristiques topologiques et géométriques qui permet une description d'un caractère. Pour qu'un système de reconnaissance soit performant il faut considérer le maximum de caractéristiques décrivant l'objet à reconnaître, tout en tenant compte du problème de temps d'apprentissage qui est considérable dans le cas des réseaux de neurones. Dans notre travail nous avons essayé de décrire les caractères par leurs caractéristiques pertinentes qui permettent de les classer au mieux. Nous avons utilisé dans le chapitre suivant les réseaux de neurones comme méthode de classification.

# **CHAPITRE IV**

## *Réseaux De Neurones*

## **I.Introduction**

Les recherches en informatique s'ouvrent aujourd'hui sur des disciplines nouvelles. Les difficultés rencontrées en intelligence artificielle (IA) ont montré que les tâches intelligentes ne se réduisaient pas à de simple enchaînement d'inférences. La tendance actuelle conduit à une révision des options d'hier. Dès lors, un nouveau courant apparaît qui exploite les recherches en neurophysiologie pour produire des modèles de machines dont le fonctionnement pourrait mimer le fonctionnement des systèmes neuronaux [38].

Le cerveau est constitué de milliards de cellules (environ 100 milliards) appelées **neurones**, fortement interconnectées. En moyenne chaque neurone est connecté à 10.000 autres. Les neurones sont des éléments simples, dont la seule fonction modélisée jusqu'à présent se réduit à une simple prise de décision.

Les interconnexions, appelées synapses détiennent la connaissance emmagasinée par le réseau. Leur rôle consiste à diffuser de l'information en excitant ou inhibant les neurones auquel elles sont connectées [39], [40].

En régime dynamique, un neurone reçoit des signaux d'entrée, les intègre et prend une décision. Si la somme des excitations et des inhibitions dépasse un seuil fixé, il renvoie un signal de fréquence constante.

Le signal est présenté vers les neurones auxquels il est connecté, et sa décision influera ainsi sur eux. Le fonctionnement global d'un tel système permet la réalisation de tâches complexes qu'on ne peut pas aujourd'hui traiter par les ordinateurs [39], [38].

## **II.Neurones biologiques**

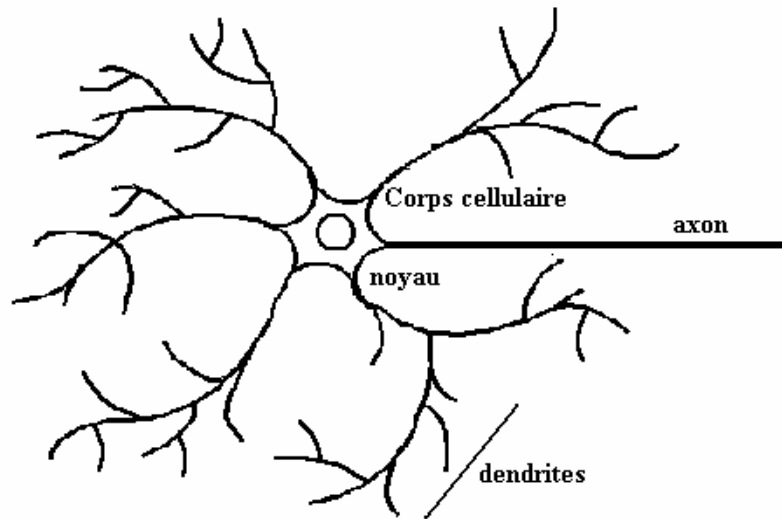
Le neurone est la cellule nerveuse qui constitue l'élément de base du système nerveux central, celui ci en possède environ 100 milliards. Les neurones possèdent de nombreux points communs dans leurs organisations générales et leur système biochimiques avec les autres cellules.

Il présente cependant des caractéristiques qui leur sont propre et se retrouve au niveau de cinq fonctions spécialisées qu'il assure :

- Recevoir des signaux en provenance des neurones voisins.
- Intégrer ces signaux.
- Engendrer un influx nerveux.
- Le conduire.
- Le transmettre à un autre neurone capable de le recevoir.

### **II.1. Structure des neurones biologiques**

Les neurones sont des cellules nerveuses décomposables en quatre parties principales (Figure IV.1) : Les dendrites, le corps cellulaire, l'axone et les synapses.



**Figure IV.1. Un neurone biologique**

- *Les dendrites* : Sont les récepteurs principaux des signaux qui parviennent au neurone.
- *Le corps cellulaire (appelé aussi soma)* : Contient le noyau et effectue les transformations biochimiques assurant la vie du neurone.
- *L'axone* : C'est une fibre nerveuse, elle transporte les signaux émis par le neurone.
- *Les synapses* : Le rôle des synapses est fondamental pour permettre aux cellules nerveuses de communiquer entre elles. Les signaux qui se propagent dans les neurones sont de nature électrique. Cependant, il n'existe pas de liaisons directes entre deux cellules nerveuses. Celles-ci sont séparées par un espace appelé fente synaptique que l'influx électrique, ne peut traverser. Le relais s'effectue à ce niveau par l'intermédiaire d'un médiateur chimique [39].

### **Description**

Les synapses se rencontrent plus particulièrement entre les axones et les dendrites. On peut cependant en trouver entre deux dendrites, entre un axone et un corps cellulaire ou entre deux axones. Dans ce dernier cas les deux configurations suivantes sont possibles soit l'extrémité d'un axone vient modifier l'activité synaptique entre un axone et une dendrite, soit elle vient se brancher directement sur son comportement [39].

### **II.2. Fonctionnement d'un neurone**

D'une façon simple, le neurone capte les signaux (transportés par l'axone) par ses dendrites, et ceci par l'intermédiaire de la synapse connectant l'axone au neurone en question, et par l'aide de substance chimique correspondantes.

Ces signaux sont ensuite acheminés le long des dendrites vers le soma du neurone (corps cellulaire), qui les traite et transmet le courant électrique résultant aux neurones auxquels il est connecté par son axone.



Le schéma classique présenté par les biologistes est celui d'un soma effectuant une sommation des influx nerveux transmis par ses dendrites.

- Si la sommation dépasse un certain seuil, le neurone répond par un influx nerveux qui se propage le long de son axone.
- Si la sommation est inférieure à ce seuil, le neurone reste inactif.

Le système nerveux central est un système de commande central et hiérarchisé, il assure la stabilité (résiste aux changements du monde extérieur), l'adaptabilité (mécanismes régulateurs engendrés suite aux perturbations), l'organisation et la communication entre les membres du corps en entier [39], [41].

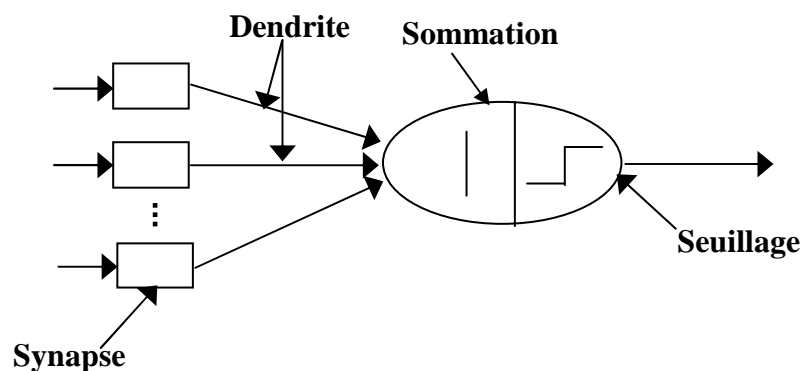


Figure IV.2. Corps cellulaires

### III. Mémoire et apprentissage

On peut définir l'apprentissage comme l'acquisition de propriétés associatives stables ou « *toute modification du comportement relativement durable, que l'on peut attribuer à l'expérience sensorielle passe de l'organisme* », et la mémorisation comme l'engagement et le rappel de ces propriétés aux modifications.

### IV. Historique

Traditionnellement, on fait remonter le début des travaux sur les réseaux de neurones aux travaux de Mc Culloch et Pitts, ils inventèrent le premier neurone formel, un produit scalaire entre un vecteur d'entrée et un vecteur poids, suivi d'un élément à seuil répondant 0 ou 1 selon que le résultat du produit scalaire est supérieur ou non au seuil. Un peu plus tard (1949), Donald Hebb essayait d'expliquer les effets d'apprentissage et de mémoire à partir de groupes de cellules, il propose que les cellules apprennent à modifier l'intensité des connexions qui les relient en fonction de leur activité simultanée [42].

Vers la même époque (1951), Minsky construit ce qui fut probablement le premier modèle réel d'un réseau de neurones, son « *ordinateur neuronal* » appelé Smarl. Ce dernier fut en quelque sorte l'ancêtre de Mark1, le premier perceptron créé par Rosenblatt dans les années soixante, le perceptron de Rosenblatt constitue le premier essai de faire apprendre des catégorisations perceptives à un réseau de neurones artificiels.

Dans les mêmes années, Widrow produisait une machine nommée Adaline (pour adaptative digital linear unit) qui apprendra plus tard à atténuer le bruit des lignes téléphoniques. Le modèle du perceptron suscita beaucoup de recherches et sans doute trop d'espoir, quand deux mathématiciens, Minsky et Papert, démontrèrent en 1969 les limites théoriques et pratiques du perceptron, l'effet fut presque immédiat : les chercheurs et investisseurs se désintéressèrent des réseaux de neurones pour se tourner vers l'approche symbolique de l'intelligence artificielle qui semblait beaucoup prometteuse.

Le renouveau actuel des réseaux de neurones, est dû à des contributions originales, comme celle de Hopfield en 1982, qui en montrant l'analogie des réseaux de neurones avec certains systèmes physiques, a permis de leur appliquer un formalisme riche et bien maîtrisé.

## **V. Les réseaux de neurones artificiels**

### **V.1. Modélisation**

#### **V.1.1. Définition**

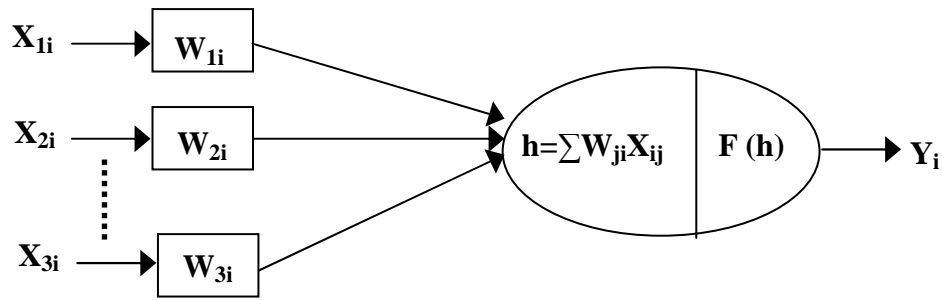
Un automate est caractérisé par :

- Un état interne.
- Un ensemble de connexions avec les autres automates.
- Une fonction de transition qui lui permet de calculer son état interne en fonction des signaux reçus sur ses connexions.

#### **V.1.2. Le neurone formel**

La première modélisation d'un neurone présentée par MAC CULLOCH et PITTS en 1943 [39] s'inspirant de leurs travaux sur les neurones biologiques, ce modèle a deux états possibles : Actif ou inactif. Stimulé par d'autres neurones analogues via des connexions pondérées, lui-même est actif ou inactif selon que la somme des stimulations reçues dépasse ou non un certain seuil ce neurone formel est modélisé donc par un automate booléen c'est à dire que ses entrées et sa sortie sont booléennes.

La structure générale du neurone formel (i) est comme suit :



**Figure IV. 3. Structure générale du neurone formel**

Avec :

- $(X_{ij})$  : Signaux en entrée au neurone.
- $(W_{ij})$  : Poids de connexions.
- $(Y_i)$  : Sortie calculée du neurone.
- $F$  : Fonction d'activation.
- $h$  : Fonction d'entrée totale.

La fonction  $F$  est une fonction de seuillage :

$$\begin{cases} F(h)=0, & \text{si } h < \theta_i \\ F(h)=1, & \text{si } h > \theta_i \end{cases}$$

Ou  $\theta_i$  est le seuil du neurone  $i$ .

**L'Algorithme général de ce neurone est le suivant**

**1-**Fait la somme pondérée  $\sum W_{ji} X_i$  des potentiels d'action qui lui parviennent, puis s'active suivant la valeur de cette sommation pondérée,

**2-Si** cette somme dépasse un certain seuil,

**Alors** le neurone est active et transmet une réponse (sous forme de potentiel d'action) dont la valeur est celle de son activation ;

**Sinon** le neurone n'est pas active et ne transmet rien.

**V.1.3.Modélisation générale**

Le modèle du neurone a peu évolué depuis les travaux de Mc Culloch et Pitts [39]. Il reste utilisé aujourd'hui parfois légèrement modifié, dans la plupart des réseaux de neurones artificielles.

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

- La nature de ses entrées ( $X_i$ ) binaires ou réelles.
- La fonction d'entrée totale ( $h$ ) définissant le prétraitement effectué sur les entrées.

Elle est :

- Booléenne.
  - Polynomial de degré supérieur à 2.
  - Linéaire ( $h(X_1, \dots, X_n) = \sum W_i X_i$ ).
  - Ou affine ( $h(X_1, \dots, X_n) = \sum W_i X_i - a$ ).
- La fonction d'activation, ou d'état,  $F$  définissant l'état interne du neurone en fonction de son état total,  $F$  : peut être :
    - Une fonction binaire

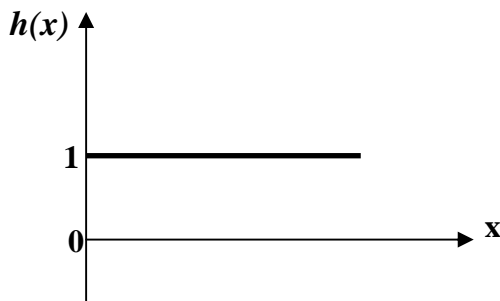


Figure IV.4. Fonction de Heaviside

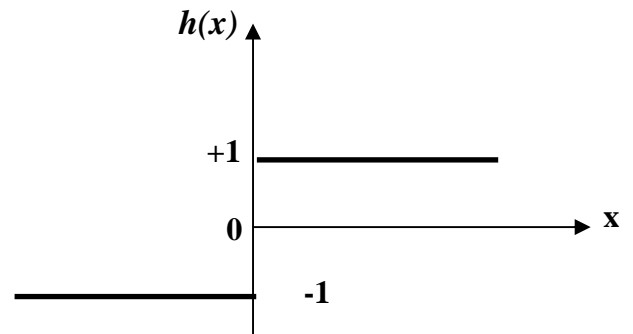


Figure IV.5. Fonction signe

- A seuil ou à multi-seuils

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

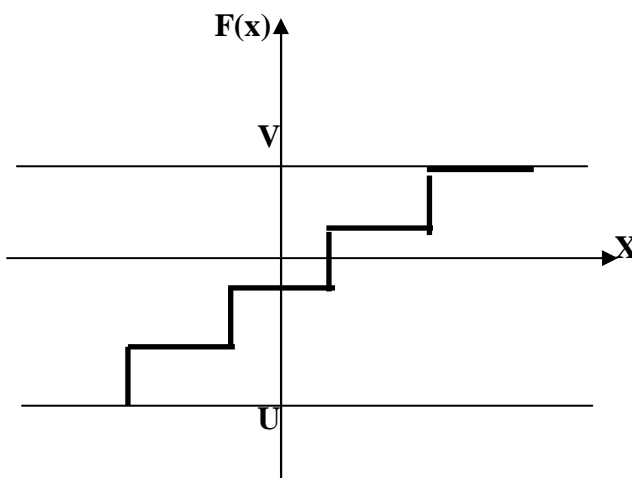


Figure IV.6. Fonction linéaire à multi-seuils

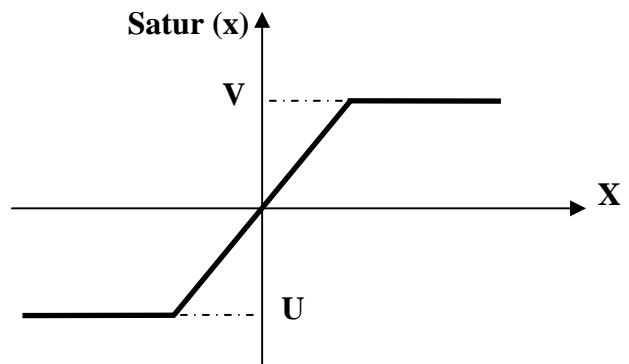
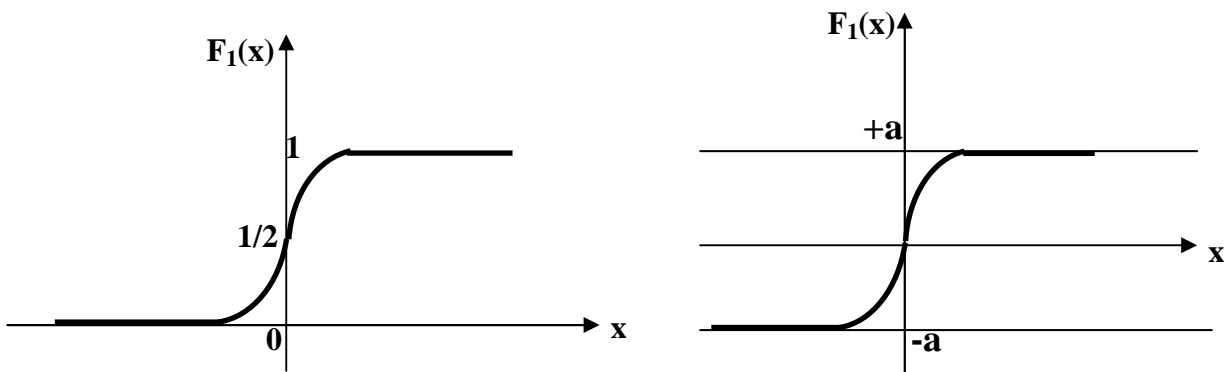


Figure IV.7. Fonction à seuils

$$\square \text{ Sigmoide } \begin{cases} F1(x)=a(e^{kx}-1)/(e^{kx}+1) \\ F2(x)=1/(1+e^{-x}) \end{cases}$$



**Figure IV.8.Fonction sigmoïde**

- La fonction de sortie  $g$  calculant la sortie du neurone en fonction de son état d'activation généralement  $g(x) = x - \theta$ .
- La nature de la sortie du neurone.

## VI.Tour d'horizon des réseaux de neurones

Sous le terme de réseaux de neurones artificiels on regroupe aujourd'hui un certain nombre de modèles inspirés du concept de neurone.

Actuellement, la plupart, pour ne pas dire la totalité, des modèles connexionnistes existants sont fondés sur la notion d'assemblée de neurones très primitifs (réalisant des opérations mathématiques très simples) reliés par des connexions affectées de poids dont la modification se fait en fonction du temps, selon une règle choisie appelée règle d'apprentissage. Ils diffèrent seulement par leur topologie (multicouches ou totalement interconnecté, nombre de neurones d'entrée, nombre de neurones de sortie), leur type de connexions et par la règle d'apprentissage utilisée.

### VI.1.Définition

Un réseau de neurones peut être alors défini par [43] :

- Un ensemble de neurones de même type (automates reliés entre eux) : Nœuds.
- Nombre de connexions pondérées entre les nœuds : Connaissance.
- Un seuil associé à chaque nœud.
- Une fonction de transfert  $F$ , généralement  $F(\sum W_{ij} X_{ij} - \theta_i)$ .

Pour chaque nœud qui détermine l'état du nœud en fonction de :

- Poids des entrées  $W_{ij}$ .
- Valeurs des entrées  $X_{ij}$ .
- Le seuil  $\theta_i$ .

Un réseau de neurones se caractérise premièrement par sa topologie c'est à dire la façon dont les neurones sont reliés. D'une manière générale, l'architecture des réseaux de neurones formels peut aller d'une connectivité totale à une connectivité locale.

La deuxième caractéristique est l'évolution du réseau, on peut dire entraîner le réseau revient à modifier les liaisons inter-neuronales. Chaque modèle totalement interconnecté ou à couches, possède des règles d'apprentissage qui ne font que des dérivées de la règle de Hebb citée ci-dessus.

Utiliser ces règles s'appelle *phase d'apprentissage* du réseau, pendant laquelle on lui fournira des exemples qu'il devra mémoriser, on parle alors d'algorithmes et de types d'apprentissage.

En général, l'architecture des réseaux de neurones se décline grossièrement de deux manières [44] :

- Les réseaux entièrement connectés.
- Les réseaux à couche.

## VI.2. Les réseaux entièrement connectés (réseau bouclé)

Est tel que chaque cellule est reliée à toutes les autres et possède même un retour sur elle-même. Les signaux peuvent être propagés en avant ou rétropropagés, on parle alors de réseaux *feed\_back* (Fig.IV.9).

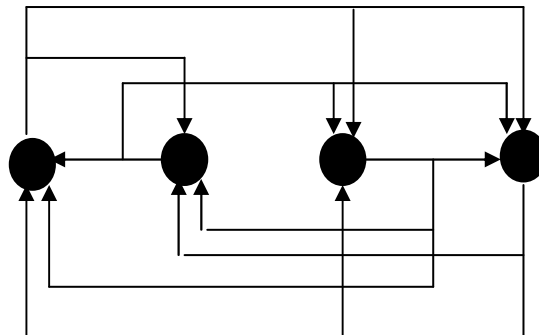
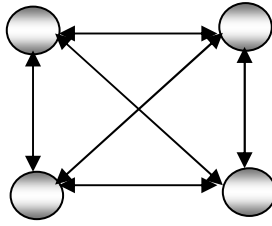


Figure IV.9. Réseaux entièrement connectés

### VI.2.1. Les réseaux les plus célèbres

#### VI.2.1.1. Le modèle de Hopfield

Ce modèle a été proposé par Hopfield en 1982 afin de mémoriser des formes et des motifs. Il est typique d'un réseau totalement connecté. 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). Ce réseau comporte  $N$  neurones binaires, chaque neurone pouvant être dans l'état (1) ou (-1), soit  $2^N$  états possible pour le réseau complet [39] (Fig.IV.10).



**Figure IV.10. Structure d'un réseau de Hopfield à 4 neurones**

La fonction de transfert de chaque neurone est une fonction à seuil. Chaque neurone est relié à tous ses voisins par une connexion de poids notés  $t_{ij}$  symétrique soit  $t_{ij}=t_{ji}$  pour tout  $(i, j) \in [1, N]^2$ , et  $(t_{ii}=0)$  pour chaque neurone(i), mais n'agit pas sur lui-même la première phase dans l'algorithme de Hopfield consiste à initialiser les connexions entre neurones : Il s'agit de mémoriser M exemples. On présente ensuite une forme au réseau, par la suite on itère jusqu'à ce que deux outputs consécutifs ne diffèrent qu'aisément plus.

### Algorithme De Hopfield

**1-Initialiser les connexions**

$$t_{ij} = \sum_{k=1..M} X_i^k X_j^k \quad 1 \leq i, j \leq M$$

La classe K est composée des éléments :  $(X_1^k, X_2^k, \dots, X_n^k)$  qui valent 1

**2- Proposer une forme en entrée**

$$O_i(0) = X_i \quad \text{pour } 1 \leq i \leq N$$

Au temps 0, la sortie du neurone i est  $X_i$ , la forme à déterminer étant  $(X_1, X_2, \dots, X_N)$ .

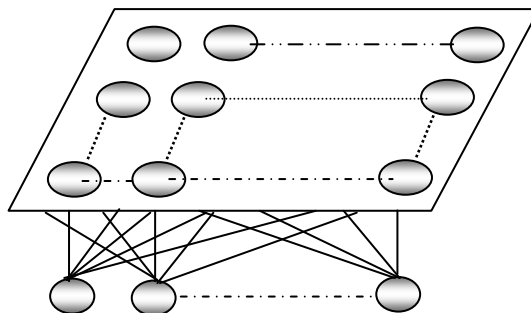
**3-Iterations**

$$O_j(t+1) = F\left(\sum_{i=1..N} t_{ij} O_i(t)\right) \quad \text{pour } 1 \leq i \leq N$$

Au temps (t+1) la sortie du neurone i est la fonction à seuil binaire des sorties des autres neurones. On itère ce processus jusqu'à ce que deux outputs consécutifs ne soient que très peu différents.

#### VI.2.1.2. Les cartes auto-organisatrices de KOHONEN

Les travaux de chercheur Teuvo Kohonen ont une place à part dans la communauté connexionniste. Les méthodes qu'il développe reposent sur des modèles biologiques [45].



**Figure IV.11. Réseaux de Kohonen dont la topologie est fixée à deux dimensions**

Second modèle d'apprentissage non supervisé, l'algorithme de Kohonen reprend cette idée en dégageant la notion de « *neurone leader* » et surtout au voisinage où sont regroupés les neurones qui répondent de façon semblable aux formes proposées en entrée. La topologie du réseau est, elle aussi, spécifique. Elle est formée de deux couches, une d'entrée classique, et une de sortie où les neurones sont entièrement connectés dans un plan.

On définit autour de chaque neurone un voisinage qui évoluera dans le temps en se rétrécissant [44].

Après avoir initialisé les connexions aléatoires, on calcule la distance euclidienne qui s'étend entre l'entrée et chaque neurone de sortie, puis on choisit le neurone de distance minimale. Il ne reste alors qu'à modifier les poids des connexions des neurones appartenant à son voisinage et d'itérer le processus suivant une règle de Hebb après avoir réduit les voisinages [46],[44],[43].

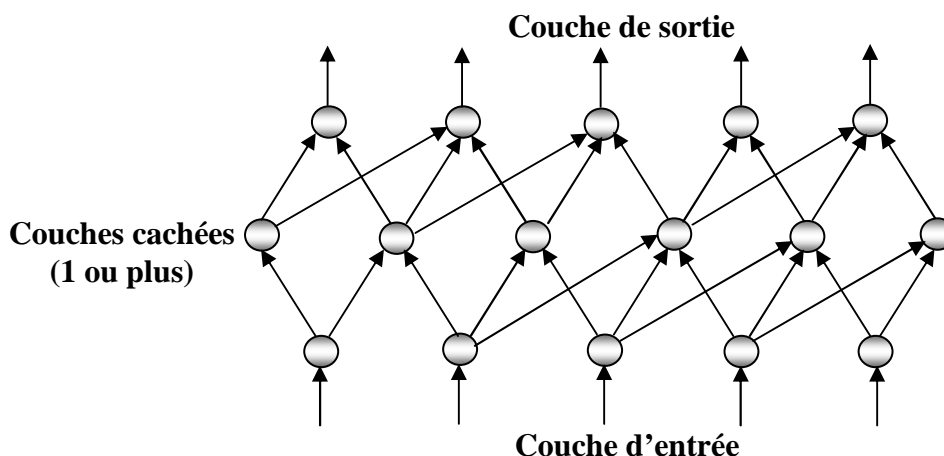
### Algorithme De Kohonen

D'une façon générale, l'algorithme se déroule en six étapes :

- 1- Choix du nombre de neurones et de la topologie du réseau, initialisation aléatoires des poids, et initialisation du voisinage et du pas d'adaptation.
- 2- Présenter aléatoirement un exemple (une entrée) de la base d'apprentissage.
- 3- Rechercher le neurone *i*. « **le plus proche** » (le gagnant).
- 4- Modification des poids  $W_i$ .
- 5- Calcul de l'entendue du voisinage et du pas d'adaptation.
- 6- Si le test d'arrêt est vérifié alors fin, sinon aller à l'étape 2.

### VI.3. Les réseaux à couches (réseau non bouclé)

Ce sont des réseaux possédant une organisation en couches successives (2,3 ou rarement plus), telle que Les neurones d'une même couche ne sont pas connectés entre eux. Chaque couche reçoit alors des signaux de la couche précédente et en envoie à son tour à la couche suivante. Les signaux ne sont alors propagés qu'en avant, on parle alors de réseaux *feed\_forward* (Fig. IV.13).



FigureIV.13. Réseaux à couches



### VI.3.1. Les réseaux les plus célèbres

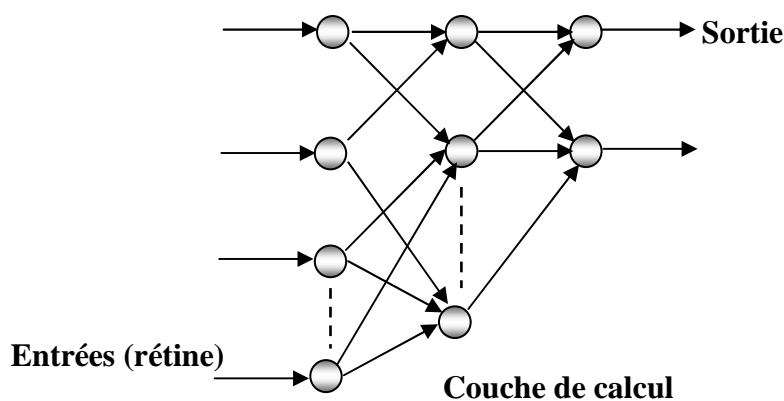
#### VI.3.1.1. Perceptron de ROSENBLATT

Le perceptron peut être considéré comme le premier des réseaux de neurones conçu en 1958 par Rosenblatt. Son but est d'associer des configurations (des formes) présentées en entrée à des réponses, il est linéaire et mono couche, inspiré du système visuel.

Le perceptron est un réseau à trois types de cellules, agencées pour accomplir des fonctions de perceptron de bas niveau [45], [39] :

- Les cellules détectrices de phénomènes extérieurs sont rassemblées dans la **rétine** : où sont présentés les exemples. Les cellules de cette première couche répondent en **oui/non** ; La réponse « **oui** » correspond à une valeur « **1** » et la réponse « **non** » correspond à une valeur « **0** » à la sortie du neurone.
- Les cellules d'association intermédiaires (prés-processeurs) : Connectées à des cellules de la rétine et à des cellules de décision. Chacune de ces cellules calcule une fonction  $f_i$  fixées à l'avance,  $f_i : \mathbf{R}^n \rightarrow \mathbf{R}$
- Des cellules de décision en sortie : reçoivent leurs entrées des cellules d'association. Elles représentent la sortie du perceptron, dans le cas du perceptron simple ces dernières sont réduites à une seule. Seules ces dernières
- cellules renvoient des signaux à la couche de calcul (partie intermédiaire) jusqu'à ce que les connexions se stabilisent.

##### VI.3.1.1.1. Structure générale du PERCEPTRON



FigureIV.14. Structure du Perceptron

Les neurones d'une même couche ne sont pas connectés entre eux. La fonction de transfert utilisée est une fonction à seuil binaire. Seules les connexions entre les unités d'association et les unités de réponse sont affectées d'un poids qui est modifié par la règle d'apprentissage.

L'apprentissage utilisé est un apprentissage supervisés, ce qui signifie qu'on fournit au réseau des exemples mémoriser.L'apprentissage peut être basé sur la loi de Widrow-Hoff , ou la règle de Hebb.

### **Algorithme Générale De Perceptron**

N neurones en entrée, M en sortie :

**1- Initialiser**

Pour  $i=1..N, j=1..M$

$W_{ij}=\text{random}$

$O_i=\text{random}$

**2- Présenter une forme( $x_1,x_2,\dots,x_n$ ) et la sortie souhaitée( $S_1,S_2 ,\dots,S_m$ )**

**3- Calculer la sortie du neurone j**

$O_j(t)=f(\sum_{i=1..N} W_{ij} x_i - \theta_i)$  pour  $1 \leq j \leq M$

F est une fonction de transfert à seuil. Le vecteur de sortie est  $(O_1, O_2,\dots,O_n)$ .

**4 - Modifier les poids**

$W_{ij}(t+1)=W_{ij}(t) - \alpha (O-S)x_i$

$\alpha$  : est le coefficient d'apprentissage,  $0 < \alpha < 1$

$O-S = \sum_{i=1..N}(O_j- S_j)$

**5- Revenir à 2 jusqu'a stabilisation du système.**

#### **VI.3.1.1.2.Limite du PERCEPTRON**

- Les sorties du réseau seront limitées à deux valeurs (vu ses fonctions de transfert)
- Il ne peut traiter que les problèmes linéairement séparables

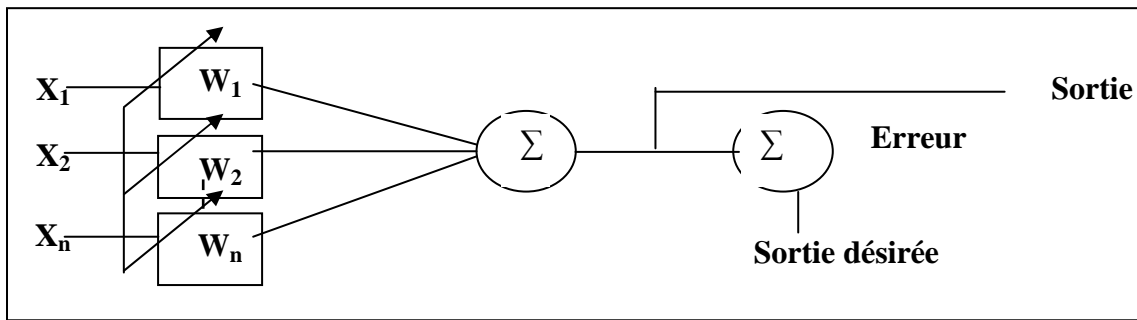
Cependant, il a été démontré que si le problème est linéairement séparable, le perceptron donne toujours des résultats satisfaisants.

#### **VI.3.1.2. ADALINE : Adaptive Linear Elément**

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, 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 soient sa taille et son orientation [44].

##### **VI.3.1.2.1.Structure générale de l'ADALINE**

La structure générale de l'ADALINE est donnée par la figure suivante (Figure IV.15)



**Figure IV.15. Adaline**

Dans ce modèle, la fonction  $f$  est  $f(z_i) = z_i$ , les poids des connexions sont déterminés par la méthode du coût quadratique. On cherche, comme pour l'exemple du perceptron classique, à construire un modèle d'apprentissage supervisé à l'aide de la règle LMS ( Least Mean Squate ) qui n'est qu'une variante de la règle d'apprentissage Delta normalisée par Widrow et Hoff [90]. Cette règle consiste à trouver les coefficients qui permettent de minimiser la fonction qui représente l'erreur quadratique entre le résultat obtenu et le résultat désirée. On peut procède de deux façons, l'un déterministe, l'autre probabiliste.

### Algorithme de l'ADALINE

Etant donné:  $B$ : gain d'apprentissage

$X$ : entrées

$E$ : erreur

1- L'entrée et la sortie souhaitée sont présentées

2- Calcule le signal d'erreur.

$$E = Y_c - Y_s \text{ ou } \begin{cases} Y_c = W_i X_i : \text{Sortie calculée.} \\ Y_s = \text{Sortie souhaitée.} \end{cases}$$

3- ajuste les pondérations par conséquent :

$$\Delta W = W_{nouv} - W_{anc} = \frac{\beta E X}{\|X\|^2}$$

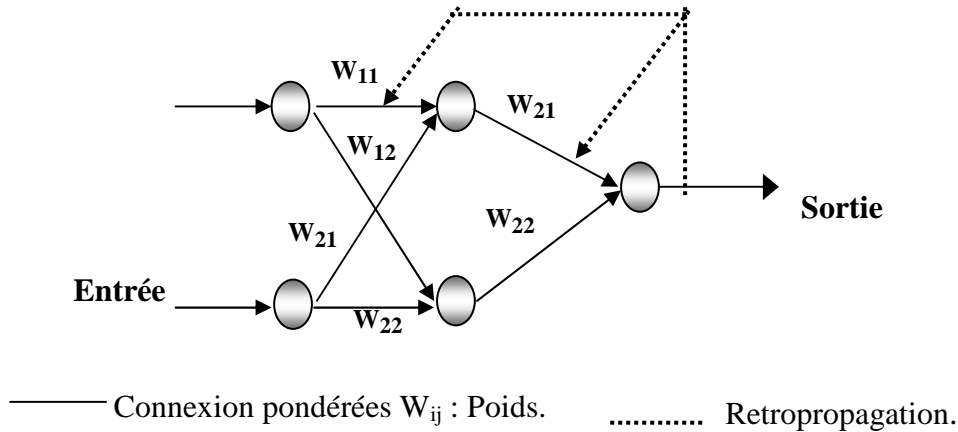
- Les valeurs de bonne pratique de  $\beta$  sont dans l'intervalle  $[0,1]$ .
- Le choix de  $\beta$  influence :
  - La taille des poids, donc le taux d'apprentissage.
  - La convergence de l'algorithme.

Ce modèle de neurone a donné naissance à une autre version du perceptron utilisé dans les réseaux de neurones dont l'algorithme d'apprentissage est donné par la retropropagation (présente ci-dessous), la forme de ce neurone est donnée par la figure suivante [44].

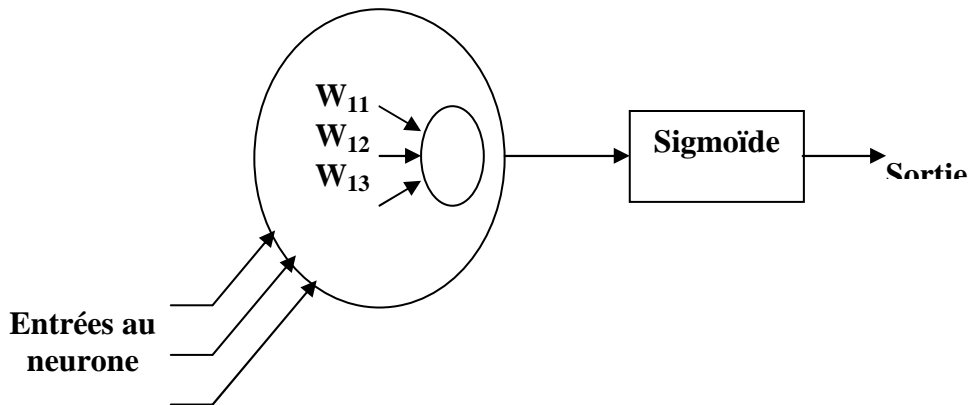
### VI.3.1.3. Perceptron Multicouche (PMC)

C'est une amélioration du perceptron comprenant une ou plusieurs couches intermédiaires dites couches cachées. Ils utilisent, pour modifier leur poids, un algorithme de retropropagation du gradient, qui est une généralisation de la règle de Widrow\_Hoff (voir) les PMC agissent comme un séparateur non linéaire et peuvent être utilisés pour la classification, le traitement de l'image, l'aide de la décision ou la commande d'un processus.

#### VI.3.1.3.1. Structure du réseau



La structure générale de chaque neurone de la couche cachée ou de sortie est:



#### VI.3.1.3.2. Algorithme de RETROPROPAGATION

L'algorithme de la retropropagation ou que l'on désigne couramment par back\_propagation ou Generalized Delta Rule, est une généralisation de la règle de Widrow\_Hoff pour un réseau multicouche. Il a été mis au point simultanément par deux équipes indépendantes en France (Le cun, Gallinari, Fogelman\_Soulie) et aux états unis (Rumelhart, Hinton, williams) et ce dans les années 1980 [39].

Mathématiquement, cet algorithme utilise simplement les règles de dérivation composée et ne présente aucune difficulté particulière. Dans cet algorithme, de même que l'on soit capable de propager un signal provenant des cellules d'entrée vers la couche de sortie, on peut, en suivant le chemin inverse, retropropager l'erreur commise en sortie vers les couches internes.

Les neurones utilisés dans ce modèle sont fondamentalement de même nature que le neurone linéaire à seuil du perceptron.

La fonction d'activation est une fonction sigmoïdale (continue, monotone, croissante et différentiable, limites finies à  $\pm \infty$ ) [48].

$$\text{Par exemple : } \begin{cases} f(x) = 1/(1+e^{-(x+t)}) \\ x = \sum w_i x_i \end{cases}$$

Où  $t$  est le seuil

La topologie du réseau à retropropagation est formée de plusieurs couches de neurone (3 au minimum) sans communication à l'intérieur d'une même couche [44] :

- Une couche d'entrée.
- Une ou plusieurs couches cachées.
- Une couche de sortie.

Chaque neurone sur une couche ( $i$ ) est relié à chaque et tout neurone de la couche ( $i + 1$ ).

L'apprentissage  $y$  est supervisé, c'est à dire que l'on présente au réseau une forme et son modèle. Le nombre de neurones au niveau de chaque couche cachée n'est défini qu'expérimentalement [49].

Le nombre de neurones dans la couche d'entrée (la couche de sortie respectivement) est égale à la dimension du vecteur d'entrée (respectivement de sortie) [50].

Chaque élément de la couche d'entrée est présenté à tout neurone de la première couche cachée. Il est important de préciser qu'aucun calcul ne s'effectue dans les neurones de la couche d'entrée, ils ne servent qu'à mémoriser le vecteur d'entrée [47]. Donnons par exemple la structure générale d'un réseau multiPerceptron à une seule couche cachée (2 neurones), 2 neurones en entrée et un neurone en sortie.

La retropropagation est capable d'approximer aussi étroitement que l'on désire n'importe quelle association entre un vecteur de dimension  $n$  et un vecteur de dimension  $m$  (n'importe quelle application de  $R^n \rightarrow R^m$ ) c'est à dire qu'à partie d'un nombre limités d'exemples, le réseau évalue la sortie à associer à une entrée différente de ceux-ci.

C'est une interpolation ou approximation, c'est à dire qu'un vecteur d'entrée proche d'un des exemples entraînera l'émission en sortie d'un vecteur proche de la sortie associée à cet exemple [44].

Cet algorithme est destinés à plusieurs applications de classification automatique, de prédiction, de reconnaissance (parole, image,...etc) et au diagnostic médical,... etc, et s'est révèle assez performant [39].

### Algorithme de la retropropagation

L'algorithme de retropropagation du gradient :

N : neurones en entrée, et M en sortie

1. Un ensemble d'entrées  $(x_1, x_2, \dots, x_n)$  et de sortie souhaitées  $(s_1, s_2, \dots, s_n)$  est présente.
2. Initialiser poids et seuil (aléatoires).
3. Choisir une paire  $(x, s)$  du domaine d'enchaînement.
4. Propager l'activation de la couche d'entrée à la couche de sortie.
5. Calcul de l'erreur dans la couche de sortie, et l'erreur dans la dernière couche cachée.
6. Ajuster les poids (couche de sortie, dernière couche cachée).
7. Continuer calcul (erreur) et ajustement (poids) pour toutes les couches cachées.
8. Revenir à 3 jusqu'a épuiser l'ensemble des données présente ntées.

$$E_j = (s_j - y_j) F_j' \left( \sum W_{ij} X_i \right)$$

$S_j$  : Sortie désirée.

$Y_j$  : Sortie calculée.

$F'$  : Dérivée de la fonction d'activation sigmoïdal.

$$\text{Couche de sortie : } \Delta W = \frac{\beta E_j X}{\|X^2\|} = \frac{\beta (S_j - Y_j) F_j' \left( \sum W_{ij} X_{ij} \right) X}{\|X^2\|}$$

$$\text{Couche cachée : } \Delta W = \frac{\beta (S_i - Y_i) \left[ \sum W_{ij} X_{ij} F_j' \left( \sum W_{ij} X_{ij} \right) \right] X}{\|X^2\|}$$

#### VI.3.1.3.3. Limite de l'algorithme de la retropropagation

- 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.

## **VII.Apprentissage**

La règle de Hebb est le premier mécanisme d'évolution proposé pour les synapses. Son interprétation pour les réseaux de neurones formels est la suivante :

La conception d'un réseau de neurone passe par une étape d'apprentissage : une fois l'architecture choisie, il reste à déterminer les bonnes valeurs des poids synoptiques de façon à obtenir le comportement voulu.L'apprentissage nécessite en général deux jeux de donnée:

- Un jeu de données d'apprentissage pour le calcul des poids synaptiques.
- Un jeu de données de test, différentes des précédentes, pour évaluer la qualité de l'apprentissage, et la capacité du réseau à traiter des données inconnues.

Après l'apprentissage, le réseau est d'abord testé : sur les exemples (jeu de données d'apprentissage). Il est inutile d'examiner le comportement du réseau sur les données de test, si les performances ne sont pas bonnes sur ceux-ci [44].

Généralement, il existe trois sortes d'apprentissages :

- Apprentissage supervisé ou apprentissage avec maître.
- Apprentissage semi-supervisé.
- Apprentissage non supervisé ou apprentissage sans maître.

### **VII.1. Apprentissage supervisé ou apprentissage avec maître**

L'algorithme d'apprentissage utilise cette mesure pour ajustement est répéter sur chaque exemple. Le rôle du maître est de mesurer l'erreur en sortie du comportement du réseau appliqué à un exemple. En général, un bon apprentissage nécessite plusieurs passages sur le jeu complet.

### **VII.2. Apprentissage semi-supervisé**

Cette forme d'apprentissage s'apparente à la précédente par la présence du maître.Mais la valeur de l'erreur en sortie n'est plus disponible. La seule information disponible est un signal d'échec ou simplement le signe de l'erreur. C'est une information très pauvre. L'apprentissage par renforcement appartient à cette famille d'algorithmes.

### **VII.3. Apprentissage non supervisé ou apprentissage sans maître**

L'absence de jugement extérieur (mesure de l'erreur en sortie) entraîne un traitement particulier des données. Suivant l'algorithme, les exemples sont utilisés comme modèles vers lesquels le réseau converge au cours du fonctionnement. Autre possibilité, les exemples forment des nuages de points dans l'espace d'entrée et l'apprentissage permet de délimiter les différentes classes de données.

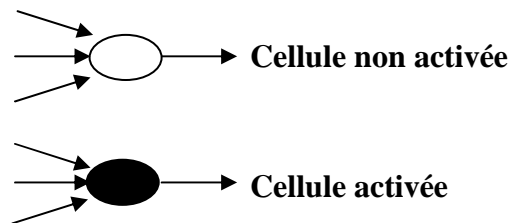
## VIII. Les règles d'apprentissage

### VIII.1. Règle de Hebb [39]

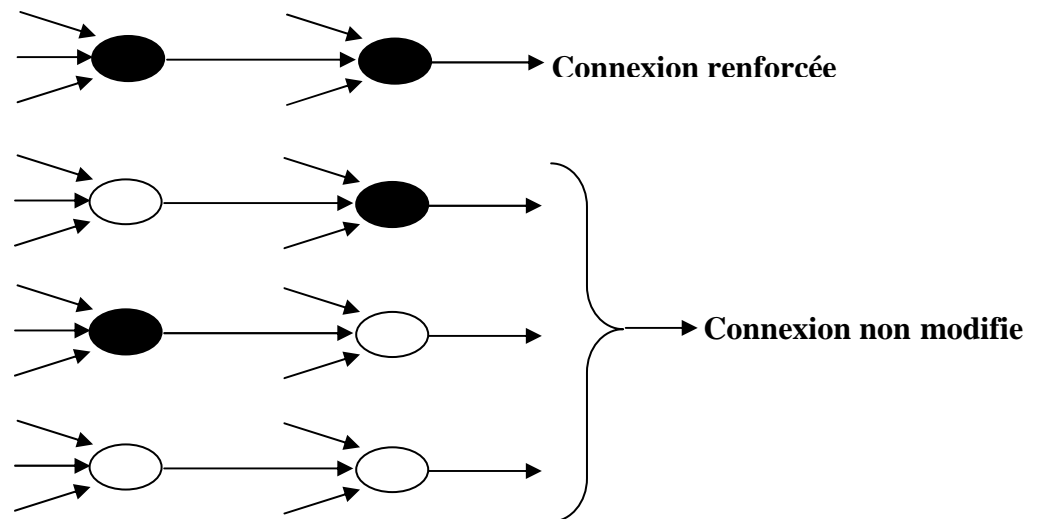
La règle de Hebb est le premier mécanisme d'évolution proposé pour les synapses. Son interprétation pour les réseaux de neurones formels est la suivante :

On considère que si deux neurones connectés entre eux sont activés au même moment, la connexion qui les relie doit être renforcée. Dans les cas contraires, elle n'est pas modifiée. D'une autre façon :

« *Quand une cellule A excite par son axone, une cellule B et que de manière répétée et persistante, elle participe à la genèse d'une impulsion dans B, un processus de croissance ou un changement métabolique a lieu dans l'une ou dans les deux cellules, de telle sorte que l'efficacité de A à déclencher une impulsion dans B est, parmi les autres cellules qui ont cet effet, accrue* ».



La règle de Hebb e alors :



Ainsi, ceci traduit que l'évolution d'une synapse est sensible aux échanges d'informations entre les deux neurones que relie cette synapse, ce qui cadre bien avec la théorie selon laquelle l'évolution du réseau nerveux est due à l'interaction entre l'environnement et le programme génétique.

L'évolution du réseau se fait par un mécanisme sélectif qui imprime sur le réseau l'image de l'environnement.



L'apprentissage, comme la mémorisation, ne peut se caractériser au niveau biochimique que par l'évolution des connexions entre les neurones.

L'apprentissage est l'acquisition de propriétés associatives stables. La mémorisation est l'enregistrement et le rappel de ces propriétés ou modifications.

### VIII.2. la règle de Widrow -Hoff [39]

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 :

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

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 :

$$E_k = e_k^T e_k = (t_k - y_k)_k^T (t_k - y_k) = \frac{1}{2}(t_k^T t_k + y_k^T y_k - 2y_k^T t_k)$$

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

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

Le gradient se calcule comme suit :

$$\frac{\nabla E_k}{w} = \frac{\partial E_k}{\partial w} = \frac{\partial E_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial w}$$

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

$$\frac{\partial E_k}{\partial y_k} = y_k - t_k.$$

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

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

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

Avec

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

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

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

## IX. Applications des réseaux de neurones

L'intérêt porté aujourd'hui aux réseaux de neurones tient sa justification dans les propriétés fascinantes qu'ils possèdent et qui devraient permettre de dépasser les limitations de l'informatique traditionnelle, à savoir : Parallélisme, capacité d'adaptation, capacité de généralisation, facilite de construction [39]. Cependant, ces réseaux présentent des limites parmi lesquelles on peut citer :

- Difficultés liées à leur implémentation hardware : la plupart des réseaux de neurones sont simulés sur des machines séquentielles.

Nous rappelons les principaux domaines d'application pour lesquels les réseaux de neurones sont les mieux adaptés :

1- La reconnaissance de formes : relative à la perception automatique, c'est à dire l'identification des formes fournies par divers senseurs, afin de parvenir à leur classification.

- Traitement d'image ⇒ Reconnaissance aérienne (Type d'avion à partir d'image, avec invariance en rotation et facteur d'échelle) ⇒ Imagerie biomédicale.
- Reconnaissance de caractères manuscrits ou dactylographies (Reconnaissance des codes postaux).

2- Reconnaissance de la parole (dictée vocale : cabine téléphonique).

3- Traitement du signal (reconnaissance de signature radar ou sonar).

4- Reconnaissance de la vision (Reconnaissance en temps réel d'objets en trois dimensions).

5- Contrôle adaptatif (contrôle d'un bras de robot, contrôle de qualité).

6- Prévion (marche financier).

7- Optimisation.

8- Gestion de l'information : (Interrogation d'une base d'images).

## X. Conclusion

Les réseaux de neurones présentés dans ce chapitre sont des systèmes paramétriques non linéaires. Ils sont déterminés lors d'une phase d'apprentissage à partir d'une base de données d'exemple, appelée base d'apprentissage.

La principale difficulté posée par l'utilisation de la méthode neuronale est l'optimisation de la phase d'apprentissage. Le choix de l'architecture adéquate ou la détermination du *pas d'apprentissage* se fait par essais successifs. Nous avons appliqué dans le chapitre suivant les réseaux de neurones comme méthode de classification des caractères arabes manuscrits.

# **CHAPITRE V**

## *Résultats Et Discussion*

## **Introduction**

Notre contribution s'articule autour du développement d'un système de reconnaissance analytique de l'écriture arabe manuscrite par la combinaison des caractéristiques statistiques et structurelles du caractère, qui peut être exploitée pour effectuer une classification en utilisant les réseaux de neurones donc une identification du caractère.

Le système proposé est composé de quatre sous-systèmes : le prétraitement, la segmentation, la paramétrisation et la reconnaissance. Il reçoit en entrée l'image d'un mot arabe manuscrit écrit sans contrainte représentant une des wilayas algériennes stockée dans un fichier BMP. Après une série de prétraitements dont le but est de simplifier les étapes ultérieures du processus de reconnaissance, l'image du mot à reconnaître sera segmentée en parties connexes (car certains caractères arabes ne se rattachent pas à leurs prédécesseurs ou à leurs successeurs et de ce fait divisent le mot en deux ou plusieurs sous-mots), enfin les parties connexes segmentés en caractères.

## **I.Architecture du système réalisé**

L'organigramme de la figure récapitule le travail présenté dans le cadre de cette recherche (FigureV.1).

### **I.1.Acquisition**

Dans tout système de reconnaissance des formes, il est nécessaire d'avoir une base de données afin d'effectuer tous les traitements. La base de données que nous avons utilisée contient des caractères que nous avons obtenus de la phase de segmentations stockés sous forme d'image binaires (FigureV.2).

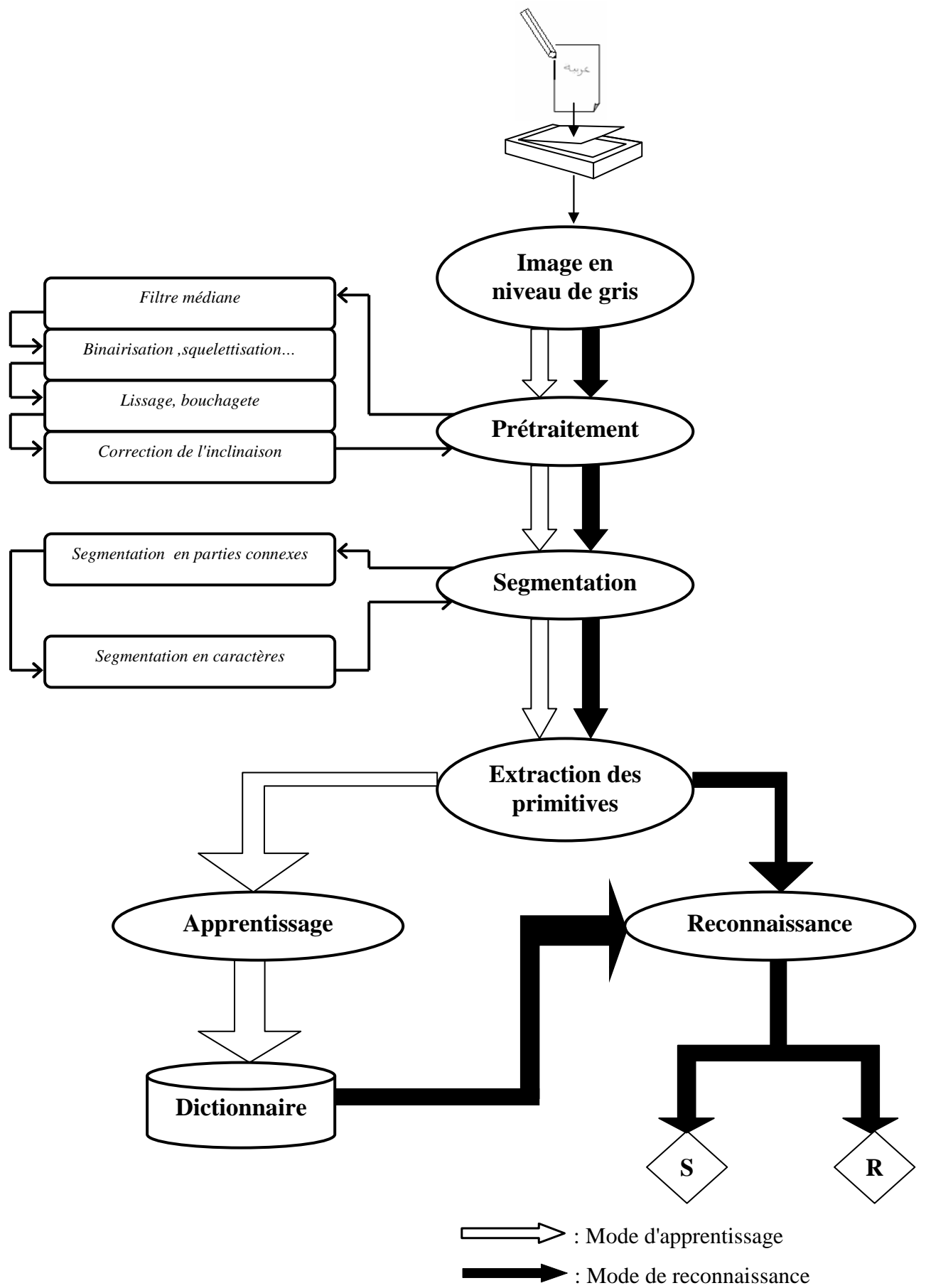
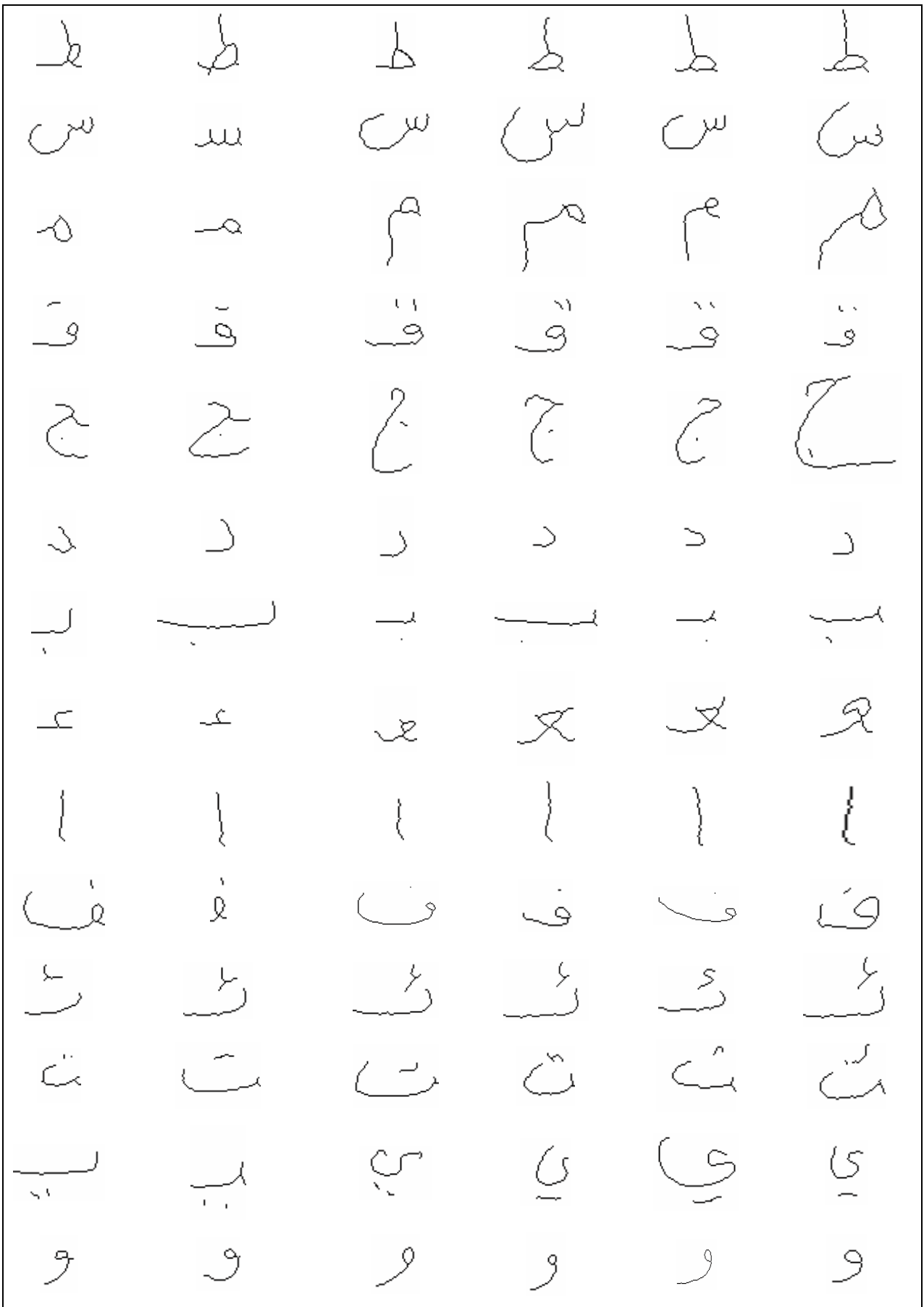


Figure V.1. Architecture du système



FigureV.2. Aperçu sur la base de donnée

## **I.2.Prè-traitements**

Cette étape a pour but de préparer l'image aux étapes suivantes pour l'extraction des primitives et la reconnaissance. Le bruit sera supprimé par les différents techniques de prétraitement (binarisation, seuillage, squelettisation,.....).

## **I.3.Segmentation**

Le module de segmentation que nous avons implémenté est réalisé en deux étapes: la segmentation du mot en parties connexes et les parties connexes en caractères.

## **I.4.Extraction des primitives**

Nous avons combiné entre les primitives statistiques, structurelles et topologique car nous avons essayé avec chacun d'eux nous avons eu de mauvaises taux de reconnaissance.

Le vecteur de primitives est composé de 62 caractéristiques:

- 7 moments invariants.
- 9 paramètres de distribution (zonage).
- 8 la densité de profils.
- 12 caractéristiques extraites à partir des histogrammes de transitions et de projections.
- 16 descripteurs de FOURIER.
- 4 pour remplissage des quatre coins d'un caractère.
- 5 pour les points diacritiques et le "hamza".
- 1 pour l'existence d'une boucle.

## **II. Classification et Reconnaissance**

Le classifieur que nous avons utilisé est un PMC à rétropropagation de gradients d'erreur à une couche cachée. Les 62 primitives qui représentent les caractéristiques possibles pour chaque caractère, sont les entrées du réseau. Les classes à discriminer sont 58 classes qui représente le nombre de caractères des différentes situations (début, fins, milieu ou isolée) d'où le choix de 58 neurones pour la couche de sortie. La fonction d'activation des neurones est la fonction sigmoïde unipolaire. Pour déterminer les valeurs du poids de toutes les connexions du réseau, nous avons choisi les valeurs initiales entre -0,5 et +0,5. Nous effectuons un apprentissage sur un nombre total de 5000 itérations. A la fin de l'apprentissage, nous obtenons les matrices de poids  $W$ ,  $Z$  et de biais  $W_0$ ,  $Z_0$ . Comme les poids et les biais sont initialisés de façon aléatoire on n'obtient jamais les mêmes valeurs à chaque itération de l'apprentissage.

Nous avons effectué un grand nombre d'expériences pour définir le nombre de neurones pour la couche cachée. Pour cela, on fait varier le nombre de neurones de la couche cachée ainsi que le

pas d'apprentissage du réseau, puis on fait un apprentissage complet pour chacun et enfin on choisit la structure qui conduit aux bons résultats.

La base d'apprentissage représente 58 caractères (classes) différents et pour chaque caractère nous avons pris 30 images de prise de vues différentes. Donc pour l'apprentissage nous avons utilisées une base de donnée de 1740 images.

Pour la phase de test nous avons utilisée une base de donnée comprend 580 (10 prototype pour chaque classe) indépendante de la base d'apprentissage.

### III. Validation des résultats

Pour étudier l'effet de la variation du nombre de neurones de la couche cachée sur le taux de reconnaissance, Nous avons fixé le nombre d'itérations à 5000 et le pas d'apprentissage à 0,5, puis Nous avons effectué les tests d'apprentissage et de reconnaissance.

Le tableau V.1 récapitule les résultats obtenus.

Taille couche entrée	Taille couche cachée	Taille couche sortie	Taux de reconnaissance
62	20	58	87,98
62	25	58	86,47
62	30	58	89,14
62	<b>40</b>	58	<b>92,56</b>
62	45	58	88,31
62	50	58	90,06

**Tableau V.1. Taux de reconnaissance pour différent Nombre de neurone de la couche cachée**

Nous remarquons que le taux de reconnaissance augmente en fonction du nombre de neurone de la couche cachée, puis à partir d'un certain nombre, ce taux diminue. Le taux de reconnaissance maximum est atteint à un nombre de 40 neurones de la couche cachée.

Pour une meilleure visualisation des choses, on a essayer de faire l'apprentissage en étudiant la variation de le taux de reconnaissance en fonction du pas d'apprentissage d'où les résultats suivants :



Taille couche entrée	Taille couche cachée	Taille couche sortie	Pas d'apprentissage	Taux de reconnaissance
62	40	58	0,2	85,32
62	40	58	0,3	90,74
62	40	58	<b>0,5</b>	<b>92,56</b>
62	40	58	0,6	87,43
62	40	58	0,8	84,66
62	40	58	0,9	86,97

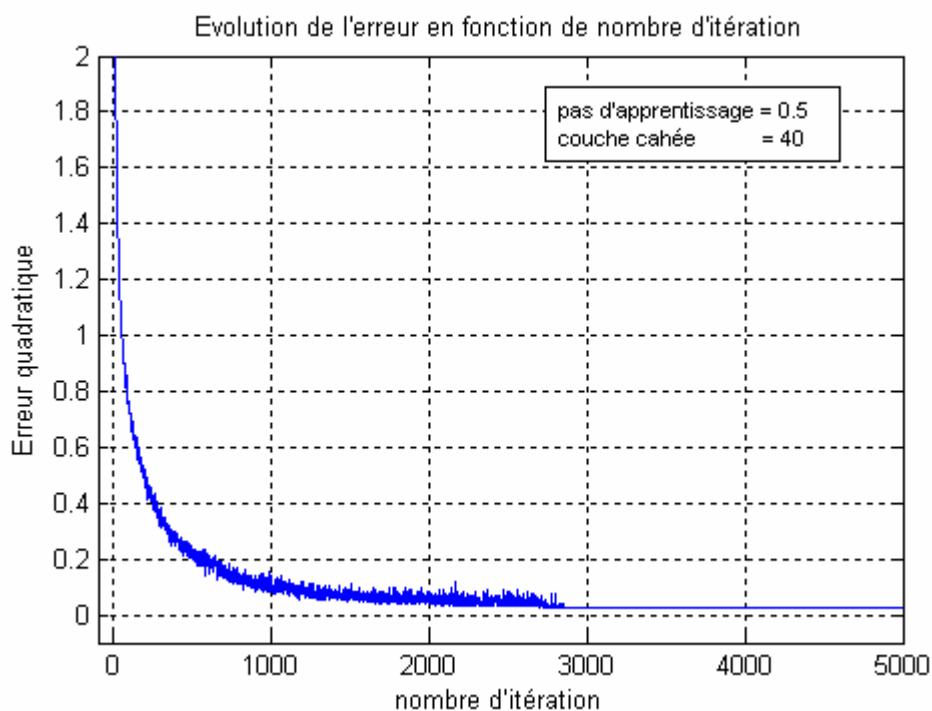
**Tableau V.2. Taux de reconnaissance pour différent pas d'apprentissage**

Nous constatons que le meilleur taux de reconnaissance est acquis pour un pas d'apprentissage égal à 0,5.

L'étude précédente nous a permis de fixer les paramètres du classifieur de notre application :

- Tailles des couches : 62 x 40 x 58.
- Pas d'apprentissage : 0,5.

Dans la phase d'apprentissage, nous avons utilisé un grand nombre d'itérations pour minimiser l'erreur à 0.0216. La figureV.3 montre l'évolution de l'erreur en fonction de nombre d'itérations.



**FigureV.3. Evolution de l'erreur en fonction du nombre d'itérations**

Nous avons effectué un test sur la base d'apprentissage elle-même nous avons obtenues un taux de 99,12%, puis nous avons effectuées les tests sur une base de test de 580 caractère (10 prototypes pour chaque classe).

Les résultats sont illustrés dans le tableau suivant :

	<b>Taux de Rec. (%)</b>	<b>Taux de Rejet (%)</b>
<b>Base d'apprentissage</b>	99,12	0,88
<b>Base de test</b>	92,56	7,44

**Tableau V.3. Taux de reconnaissance sur les bases d'apprentissage et de test**

#### **IV. Conclusion**

Dans ce chapitre, nous avons présenté un aperçu sur la reconnaissance de l'écriture arabe manuscrite qui représente de nos jours un axe de recherche important dans la reconnaissance des formes en générale. On a présenté aussi les différentes étapes de notre application, ainsi le système de reconnaissance de l'écriture arabe manuscrite qui est composé principalement d'un réseau de neurones à couches entraîné par un algorithme de retropropagation de gradient.

Nous avons fixé l'architecture du réseau approprié qui est l'étape la plus importante pour la reconnaissance.

En suite nous avons fait des tests sur le réseau choisi où nous avons étudié la variation du taux de reconnaissance en fonction du nombre de neurone de la couche cachée tout en fixant le pas d'apprentissage puis nous avons fixé le nombre de neurone de la couche cachée qui nous a donné le meilleur taux de reconnaissance précédemment toute en étudions la variation taux de reconnaissance en fonction du pas d'apprentissage, le réglage optimal du réseau neuronal nous a permis d'atteindre un taux de reconnaissance égale à 92,56% .

Le taux de reconnaissance pour l'écriture arabe manuscrite dépend de la segmentation donc on base pour l'approche analytique sur le développement des techniques de la segmentation parce que une bonne segmentation conserve le caractère complet d'où la facilité des opérations qui suivent (extraction des paramètres, classification et la reconnaissance).

# Conclusion Générale

Le problème de la reconnaissance de l'écriture arabe manuscrite non-contrainte est complexe. Dans un premier lieu, il faut trouver une modélisation qui tire profit de ses propriétés bidimensionnelles intrinsèques. De plus, cette modélisation doit prendre en compte des erreurs qui peuvent affecter les données qui arrivent souvent entachées de bruit dû à l'acquisition ou aux prétraitements (ex. binarisation, squelettisation...) qui se trouvent en amont du système de reconnaissance. La modélisation de l'écriture manuscrite peut se faire à travers deux approches totalement différentes : une approche globale ou bien une approche analytique.

Notre étude est basée sur l'approche analytique en utilisant les réseaux de neurones sur des caractères arabes manuscrits.

Dans un premier temps l'image de mot est acquise et prétraitée, nous effectuons ensuite l'étape la plus importante et la plus difficile du processus de reconnaissance c'est la segmentation. Le mot est segmenté en parties connexes, puis les parties connexes en caractère.

La phase suivante est l'extraction des paramètres pertinents de chaque caractère.

A la phase de reconnaissance chaque vecteur de primitive obtenue de la phase précédente est une entrée du réseau de neurone de type MLP.

La segmentation de l'écriture arabe manuscrite que nous avons utilisée dans ce travail nous a donné des résultats encourageantes basée sur le squelette de l'image du mot et nous avons résolu le problème lié surtout à la segmentation des lettres (shin "س") et (chin "ش") et nous espérons dans la future d'avoir des résultats mieux que celles-ci et d'augmenter le taux de reconnaissance.

# Bibliographie

- [1] Olivier Monga et Radu Horaud, «*Vision par ordinateur*», Editions Hermes, 2eme édition, 1995.
- [2] C.Remy, «*L'intelligence artificielle*», Dunod, 1994.
- [3] L. Hamami, «*Système de classification des chromosomes par utilisation des Réseaux de Neurones*», first Polytechnic Symposium on Signal Processing, 25 Avril 2002, ENP, Algiers-Algeria.
- [4] L. Hamami et N. Lassouaoui, «*An Approach Fractal and Analysis of Variogram for Edge Detection of Biomedical Images*», 6<sup>th</sup> International Work Conference on Artificial and Natural Neural Network, IWANN 2001, Granada, Spain, June 13.15.2001, Proceedings, PartII, pp.336344, published for Springer LNCS 2085 Bio-Inspired Applications of Connectionism.
- [5] Olivier Pavie, «*Le starter photos numériques, Scans et retouches* », 1999.
- [6] A. Belaid, «reconnaissance des forme : Méthode et application»,1995.
- [7] Catherine Achard, «*Cours de Traitement d'images*», 2002/2003].
- [8] J-P. Cocquerez, «*Analyse d'image : filtrage et segmentation* », 1995.
- [9] J.J. Toumazet, «*traitement de l'image sur micro-ordinateur*», Edition Sybex, France,
- [10] M. Kunt, «*Traitement d'images*», Edition 1993.
- [11] Oivid Due Trier, Anil K. Jain, & Torfinn Taxt.. «*Feature extraction methods for character recognition*». *Pattern recognition*, 29(4), 641-662. (1996)
- [12] Richard O. Duda, Peter E. Hart, & David G. Stork. « *Pattern Classification* ». (Second edition), New York: Wiley-Interscience. (2001).
- [13] Gosselin.B « *traitement de l'information, classification et Reconnaissance de formes* », Copyright © 2000 Faculté polytechniques de Mons.
- [14]: Matougui Farida, « *Système interactif pour la reconnaissance des caractères arabes manuscrits SIRCAM* » thèse de magistère C.D.T.A (1990).
- [15] M.B Kurdy and A.Joukhadar, « *Multi\_font Recognition System for Arabic Characters*», Proceedings of the 3<sup>rd</sup> International Conference and Exhibition on Multilingual Computing, Durham U.K. Dec. 1992, pp. 73 1-73 9.
- [16] L.Hamami and D.Bcrkani, «*Réognition System for Printed multi-font and multi-size Arabie characters writing*», The Arabian Journal for Science and Engineering (AJSE), Vol.27, n°!B, April 2002.
- [17] R.H. Cheng, C.W. Lee and Z. Chen, «*Recognition of radicals in Handwritten Chinese characters by means of problem reduction and Knowledge Guidance*», International Journal of Pattern Recognition and Artificial Intelligence, 10(6), pp.657-678, 1996.

- [18] H.J. Kim and S.K. Kim, «*On-line recognition of cursive Korean characters using Art-Based Stroke Classification*», International Journal of Pattern Recognition and Artificial Intelligence, 10(7), pp.791-812, 1996.
- [19] Belaïd, «*Analyse et Reconnaissance de Documents, Le Traitement Electronique du Document*», Cours INRIA, 3-7 Oct, 1994, Aix-en-Provence, pp.49-92.
- [20] H Almuallim and S. Yamaguchi, «*A method of recognition of Arabic cursive handwriting*», IEEE Transaction on Pattern Analysis and Machine Intelligence, n°5, pp.715-722. 1987.
- [21] T.S. El-Sheikh and R.M. Guindi, «*Computer Recognition of Arabic Cursive Script*», Pattern Recognition, Vol.21, No.4, 1988, pp 293-302.
- [22] A.Amin and J.F.Mari, «*Machine Recognition and correction of printed Arabic text*», IEEE Trans. Man Cybern., 9(1), 1989, 1300-1306.
- [23] Shoukry, «*A sequential algorithm for the segmentation of typewritten Arabic digitized text*», Arabian J. Se. and Eng. 16(4), 1991, 543-556.
- [24] V. Margner, SARAT-«*A System for the recognition of Arabic printed text*», 11 Int. Conf. on Pattern Recognition, 1992, 561-564.
- [25] A.Amin and H.B.Al-Sadoun, «*A new segmentation technique of Arabic text*», 10<sup>1</sup> Int. Conf. on Pattern Recognition, 1992, 441-445.
- [26] Al-Badr and R. Haralick, «*Segmentation free word recognition with application to Arabic*», 3<sup>rd</sup> Int. Conf. On Document Analysis and Recognition, Montréal, 1995, 355-359.
- [27] Amin, «*Arabic Character Recognition, Handbook of character recognition and Document Image Analysis*», edited by H. Bunke and P.S.P. Wang, 1996.
- [28] Motawa, A. Amin and R Sabourin, Segmentation of Arabie Cursive Script, 4<sup>h</sup> Int. Conf. On Document Analysis and Recognition, ICDAR97, Germany, 1997,
- [29] Abdel.Belaïd and Y.Belaïd, «*Reconnaissance de formes* », Inter-Edition, Paris, 1992, 412p.
- [30] L.Hamami-Mitiche, Segmentation d'un paragraphe de texte imprimé en caractères, COMAET'98, Bedjaia,-Algérie.
- [31] Olivier Monga et Radu Horaud, Vision par ordinateur, Editions Hermès, 2<sup>eme</sup> édition, 1995.
- [32] A.Belaïd, K.Tombre, «*Analyse de documents: de l'image à la sémantique*», colloque national sur récrit, et le document, Nancy, juillet 1992.
- [33] T.L. Abegnoli, «*Quatre phases de reconnaissance de caractères*». Revue Electronique n°29. France, juin 1993.
- [34] N.Ragot: "Reconnaissance de formes par modélisation mixte intrinsèque /discriminante à base de système d'inférence floue hiérarchisés". Thèse de docteur de l'université de Rennes i, octobre 2003.
- [35] M.Pechwitz and V.Maergner, «*HMM-based approach for handwritten Arabic word*

*recognition using the IFN/ENIT*».Database.in ICDAR,pages 890-894, IEEE computer society,20003.

- [36] HEBBACHE Nadia et KIOUS Houria, « Réalisation d'une interface en langue arabe pour un tuteur intelligent », Thèse d'ingénieur d'état, Institut d'informatique. USTHB, 1997.
- [37] DARGENTON Patrice, « Contribution à la Segmentation et à la Reconnaissance de l'écriture Manuscrite », Thèse de doctorat, Institut National des Sciences Appliqués de Lyon, 14 Décembre 1994.
- [38] M.GUIVARCH « *Réseaux de neurones : Techniques de base et apprentissage* » 3ème université d'été, Réseaux connexionnistes en informatique, Méthodes et applications, pp II.2-II31, Lyon 1989.
- [39] E.DAVALO, « *Des Réseaux de Neurones* » Ed. EYROLLES, PARIS 1989.
- [40] J.HOPFIELD, D.TANK «*Les réseaux de neurones formels* » Pour la science, pp 80-89,1988.
- [41] P.Verschure. « *Formal Minds And Biological Brains* » IEEE Expert, PP 66-75, Octobre 1993.
- [43] H.LEMBERG «*Les réseaux neuronaux artificiels* »INFOPC N°71,pp 20-28, 1989.
- [44] P.J.MELSE« *Neural Networks: a conceptuel overview* » Technical report, TCR-89-08,PP 1-41, Aout 1989.
- [45] F.FOGELMAN SOULIE, « *Applications des méthodes connexionnistes en IA*»,3<sup>ème</sup> Université d'été, Réseaux connexionnistes en informatique, Méthodes et Applications, PPI.1-113, Lyon 1989.
- [46] P.LIPPMAN «*An introduction to computing with neural nets*» IEEE ASSP Magazine, pp4-22, Avril 1987.
- [47] B.WIDROW « *30 Years Of Adaptive Neural Networks: Perceptron, Madaline, And Backpropagation* ».
- [48] « *Matlab:Neural Networks Tool Box User's Guide* », June 1992.
- [49] F.BLAYO, « *Tour d'horizon sur les implémentations de réseaux de neurones* »,3<sup>ème</sup> Université d'été. Réseaux connexionnistes en Informatique. Méthodes et Applications. PP VII.1-VII.68.LYON 1989.
- [50] T.FUKUDA, T.SHIBATA, « *Theory for industrial controlsystem* » IEEE Transactions on Industrial Electronics Vol 39 N°6, Décembre 1992.
- [51] HU, M.K, «Usual pattern recognition by moment invariants», IRE transaction in information theory, 82, 1991, pp 409-509.