

11/02/17/12

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université 8Mai 1945 – Guelma
Faculté des Sciences et de la Technologie
Département d'Electronique et Télécommunications



**Mémoire de Fin d'Etude
pour l'obtention du Diplôme de Master Académique**

Domaine : Sciences et Techniques
Filière : Electronique
Spécialité : Systèmes Electroniques

**APPLICATION DES RESEAUX DE NEURONES
POUR LA RECONNAISSANCE DES CHIFFRES
MANUSCRITS ARABE**

Présenté par :
- **Zeghdani Imen**

Sous la direction de :
MR : Menasria Azzeddine



JUIN 2012

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
مَا أَذْكُرُونِي أَذْكُرْكُمْ وَأَخْذُرُوا إِلَيَّ وَلَا تَخْذُرُونِي

صَدَقَ اللَّهُ الْعَطِيَّةَ

Remerciements



Louange à Allah, Seigneur des mondes. Que la paix et le salut soient sur le seigneur des prophètes et le Sceau des messagers, notre prophète Mohamed, sa famille et ses compagnons.

Dans les lignes suivantes je vais tenter d'exprimer toute mon gratitude aux personnes qui m'ont soutenus et permis de mener à terme ce travail.

Je remercie avant tout mon encadreur M^K MENASRIA AZZEDDINE pour son aide précieuse aussi que ces conseils et de vous avoir toujours fait confiance et d'avoir guidé la conduite de mon travail.

Je le remercie pour sa patience et la disponibilité qui m'a été accordée pendant la réalisation de mon travail.

Je tiens à exprimer ma reconnaissance au prêtre de tous les membres du département d'électronique et télécommunication et son chef de département le docteur HAMDI RACHID.

Ainsi que tous nos enseignants qui m'ont soutenu tout le long de mes études.

Dédicace

Louange à Allah, Seigneur des mondes, Que la paix
et le salut

Soient sur le seigneur des prophètes et le
Sceau des messagers

Notre prophète Mohamed, sa famille et ses
compagnons.

Je dédie ce modeste travail à ceux que
l'expression de

Reconnaissance était difficile et qui ont
dépensés leur

Sacrifices pour que j'arriverais à ce jour là, The Flower
by Vasilisa
ceux qui m'ont

Appris c'est quoi la vie, **chers parents:**

A mes frères « **Khayre Eddine, Billel** ».

A mes sœurs « **Khawla, myasser** ».

A mon marie « **Merabet Amar** »

A mes amies « **Brahmia M^{ed} amine,
Fernene mahdi, Wissem, Aida** »

Imen



SOMMAIRE

INTRODUCTION GENERAL

CHAPITRE I LA RECONNAISSANCE DES FORMES

I.1-Introduction	4
I.2-Définition	4
I.3-Processus De Reconnaissance Des Formes	5
3.1 Etat réel et le codage	6
3.2 prétraitement	6
3.3 Apprentissage.....	7
3.4 l'analyse	7
3.5 Décision	8
I.4-Différentes Approches De Classification	8
4.1 La méthode statistique	8
4.2 La méthode de bayes	9
4.3 Classificateur paramétrique	9
• Classificateur euclidien	9
• Classificateur quadratique	9
4.4 Classificateur non paramétrique	10
• La méthode du plus proche voisin	10
• Réseaux de neurones artificiels	10
• La méthode de Markoviennes	11
• Les méthodes mixtes	11
I.5- Domaines D'applications	11
5.1 Reconnaissances des formes sur signaux temporels	11
5.2 Reconnaissances des formes dans l'image numérique	13
I.6 -Tableau d'Applications	15
CONCLUSION	16

CHAPITRE II RECONNAISSANCES DE L'ECRITURE

I. Introduction	17
II.2. les Différents aspects de reconnaissance de l'écriture	17
II.2.1 Reconnaissance en-ligne ou hors-ligne.....	18
II.3. Méthodes de reconnaissance	19
II.3.1 Reconnaissance globale ou analytique	19
II.4. Stratégies de reconnaissance	19
II.5. Organisation générale d'un système de reconnaissance	20

II.5.1 Prétraitements	21
II.5.1.1 Binarisation	21
II.5.1.2 Filtrage	22
II.5.1.3 Normalisation de la taille	24
II.5.1.4 Extraction des contours	24
II.5.1.5 Extraction des composantes connexes	26
II.5.2 Segmentation	28
II.5.2.1 La segmentation des mots	29
II.5.3 Extraction des primitives (caractéristiques)	32
• Les caractéristiques adaptées à l'écriture	33
II.5.3.1 Détection des caractéristiques structurelles	34
II.5.3.2 Détection des caractéristiques statistiques	35
II.5.3.3 Détection des caractéristiques géométriques	36
II.5.4 Classification	45
CONCLUSION	46

CHAPITRE III LES RESEAUX DE NEURONE

1. Introduction	47
III.2. Modélisation du neurone	48
III.2.1 Réseaux de neurones formels	48
III.2.1.1 Model simplifié (premier modèle de neurone)	49
III.2.1.2 Modélisation générale	51
III.2.2 Différentes structures des réseaux de neurones	52
III.2.2.1 Les Réseaux à couches	52
III.2.2.2 Les Réseaux entièrement connectés	53
III.2.3 Les réseaux les plus célèbres	54
III.2.3.1 Le perceptron	54
III.2.3.2 Le modèle ADALINE	54
III.2.3.3 Le perceptron multicouche (PMC)	35
III.2.4 L'apprentissage.	55
III.2.4.1 Apprentissage supervisé	55
III.2.4.2 Apprentissage semi-supervisée	56
III.2.4.3 Apprentissage non-supervisée (auto-organisation)	56
III.2.5 Règles d'apprentissages	56

III.2.5.1 La règle de Hebb.....	56
III.2.5.2 La règle de Widrow-Hoff.....	57
III.2.5.3 Apprentissage du perceptron multicouche.....	58
III.3 Le Perceptron multicouches comme un classifieur.....	58
III.4 Evolution du perceptron multicouche.....	62
CONCLUSION.....	63
CHAPITRE IV APPLICATION SUR LES CHIFFRES MANUSCRITE	
Introduction.....	64
IV.1. Application à la reconnaissance des chiffres manuscrits.....	64
IV.1.1. Base de données.....	64
IV.1.2. prétraitement.....	65
IV.1.3.1 Le choix des primitives.....	65
IV.1.3.2. les paramètres de distribution (PD).....	65
IV.1.3.3. Les sept moments invariants (MI).....	66
IV.2. Constitution du vecteur des caractéristiques.....	67
IV.3. Classification (apprentissage et reconnaissance).....	68
IV.3.1. Apprentissage.....	68
IV.3.2. Reconnaissance.....	69
CONCLUSION.....	69
CONCLUSION GENERAL.....	70

INTRODUCTION GENERALE

Dans le domaine de l'intelligence artificielle, l'un des objectifs les plus recherchés est de doter les ordinateurs de capacités qui sont naturelles à l'être humain. Un premier pas dans ce sens consiste à faciliter la communication homme-machine en la rendant la moins contraignante possible.

Aujourd'hui les recherches intenses portant sur le traitement de l'information, la théorie de la décision et les méthodes de reconnaissance de formes d'une part, et sur la disponibilité des bases de données et le développement continu des moyens de stockage, de calcul et de traitement d'une autre part. Ont permis la conception de la machine et de systèmes qui facilitent le travail de l'homme et le remplacent souvent dans les tâches répétitives.

La reconnaissance des formes est une science de définition d'algorithmes permettant de classer des images ou bien des signaux qui caractérisent des objets dont l'aspect varie par rapport à un objet type appelé, en d'autres termes elle permet le passage de la conception des objets à leur reconnaissance.

L'écriture manuscrite est un moyen naturel de communication qui présente l'avantage d'être familier à la majorité des gens. De ce fait, elle constitue un moyen d'interaction facile avec l'ordinateur. Un système capable de lire le texte manuscrit devrait permettre aux ordinateurs de traiter une grande variété d'applications non encore accessibles. Cette tâche n'est pas triviale car les mots possèdent une infinité de représentations dues au fait que chaque personne possède une écriture qui lui est propre.

En outre depuis quelques années, les recherches dans le domaine de la reconnaissance de l'écriture se sont multipliées de manière spectaculaire et 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 manuscrite s'avère un problème extrêmement complexe qui n'a pas de solution satisfaisante jusqu'à ce jour.

On distingue dans l'écriture deux types d'écriture : l'imprimé et le manuscrit. Dans le premier type les principales difficultés ont été surmontées ; la situation diffère complètement en ce qui concerne la reconnaissance des textes manuscrits. Les problèmes rencontrés sont liés d'une part au manque de consistance des caractères, et aux lettres liées à

l'intérieur des mots et qui ont des formes différentes. Ils sont liés aussi aux nombres de scripteurs.

En revanche pour les systèmes de la reconnaissance de l'écriture, on distingue deux grandes catégories d'applications à vocabulaire limité où le nombre des mots à reconnaître constitue un lexique de taille restreinte (*approche globale*); et les applications à vocabulaire étendu nécessitent obligatoirement une phase de segmentation (*approche analytique*).

Des nombreuses études présentées dans la littérature, la plupart portent sur la reconnaissance de sous-unités isolées d'écriture, tels que les caractères, les mots ou les chaînes de chiffres, parmi les travaux récents réalisés dans ce domaine on trouve :

Leroux et al [6] proposent une stratégie de contrôle pour la reconnaissance des montants littéraux des chèques postaux, c'est un système omni-scripteur, fondé sur la stratégie de coopération entre une méthode analytique et méthode globale, dans ce travail aucun résultat n'a été présenté, c'est uniquement une stratégie proposée qui peut être adoptée pour la réalisation des système de reconnaissance de montants littéraux de chèques,

Lemaire et al [7] effectuent une coopération neuro-Markovienne pour la reconnaissance des montants littéraux des chèques, le bilan de cette technique été positif, un taux de reconnaissance de l'ordre de 80.28% est atteint sur une base de teste de 2879 mots.

Miled [8] propose deux approches Markoviennes pour la reconnaissance hors-ligne omni-scripteur de l'écriture arabe, dans un vocabulaire relativement étendu (232 classes de mots différents correspondants à des noms de villes tunisiennes), Les modèles ont été entraînés sur une base de 4720 mots et les tests ont été effectués sur une autre base de 5900 mots. Des résultats modestes de 60% ont été enregistrés.

Morita [10], présente un système hybride HMM-MLP pour la segmentation et la reconnaissance de l'écriture manuscrite et plus particulièrement des dates et les montants littéraux présentent sur les chèques brésiliens. Le bilan de cette technique a été positif, un taux de reconnaissance de 82.3% pour les dates et 96.8% pour les mots.

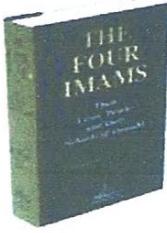
Dans le premier chapitre nous donnons un aperçu général sur la reconnaissance des formes, et nous décrivons le schéma général d'un système de reconnaissance ainsi que les méthodes de classification et leurs domaines d'utilisation.

Le second chapitre traite des systèmes de reconnaissance de l'écriture 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 d'écriture.

Le chapitre 3 est consacré à l'état de l'art des réseaux de neurones, entre autre le MLP adopté dans notre projet. Une description de la méthode d'apprentissage utilisée est aussi présentée.

Le quatrième chapitre est consacré à la conception et à l'implémentation du modèle appliqué à la reconnaissance des chiffres manuscrits. Dans ce chapitre, nous abordons les différentes étapes de reconnaissance et les aspects pratiques liés aux choix des différents paramètres. Nous terminons ce chapitre par une présentation des résultats de notre application.

Nous terminons notre travail par une conclusion général.



CHAPITRE I

la reconnaissance des formes

I.1 INTRODUCTION

I.2 DEFINITION

I.3 PROCESSUS DE RECONNAISSANCE DES FORMES

I.3.1 Etat réel et le codage

I.3.2 La prétraitement

I.3.3 l'Apprentissage

- Des Classes Connues
- Des Classes Non Connues

I.3.4 l'analyse

I.3.5 La Décision

I.4 DIFFERENTES APPROCHES DE CLASSIFICATION

I.4.1 La méthode statistique

I.4.2 La méthode de bayes

I.4.3 Classificateur paramétrique

- Classificateur euclidien
- Classificateur quadratique

I.4.4 Classificateur non paramétrique

- Réseaux de neurones artificiels
- La méthode de Markoviennes
- Les méthodes mixtes
- La méthode du plus proche voisin

I.5 DOMAINES D'APPLICATIONS

I.5.1 Reconnaissances des formes sur signaux temporels

- Signal de parole
- Signaux biomédicaux

I.5.2 Reconnaissances des formes dans l'image numérique

I.6 TABLEAU D'APPLICATIONS

CONCLUSION



1-Introduction

La reconnaissance des formes (abrégé RdF) est historiquement un chapitre de l'intelligence artificiel ; elle vise à automatiser le discernement de situation typiques au niveau de la perception.

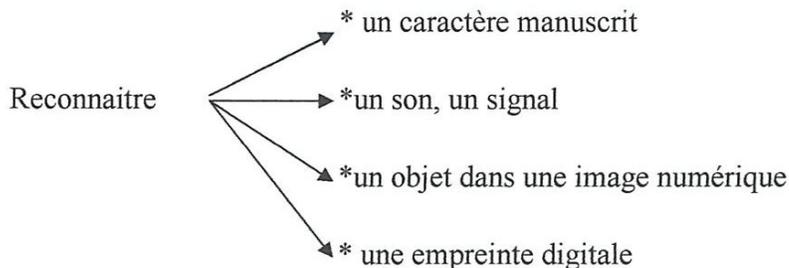
La RdF est avant tout la réduction méthodique d'information, a partir d'une donnée très riche on considère donc souvent la reconnaissance des formes comme un problème de classification, c'est-à-dire un problème de synthèse d'une fonction qui affecte chaque donnée prévisible à la catégorie pertinente.

Dans certains cas les images données dans lesquelles on veut discerner des objets de forme connue sont composés de plusieurs formes élémentaires dont les positions relatives sont significatives ; la RdF se généralise à ce cas pour une description sommaire mais appropriée de la forme composée permettant ainsi son interprétation ultérieure.

I.2-Définition

La reconnaissance des formes consiste en une automatisation de tâches de perception artificielle réalisée usuellement par le cerveau et le système sensoriel humains.

Exemple :



Une forme est une représentation simplifiée du monde extérieur définie sous une forme acceptable par l'ordinateur (par exemple un vecteur de réels, ou bien un mot d'un langage donné,.....).



I.3-Processus De Reconnaissance Des Formes

Un système de reconnaissance des formes peut comporter une phase d'apprentissage qui va consister à apprendre à reconnaître des formes sur la base d'échantillons.

Lorsque cette base d'apprentissage sera achevée le système sera alors prêt à fonctionner pour reconnaître des formes inconnues qui lui seront soumises.

Mais un système de reconnaissance des formes peut être aussi un système qui trie (fait des paquets homogènes suivant certains critères), un ensemble de formes inconnues ; il n'y a alors pas d'apprentissage à proprement parlé.

On imagine donc qu'il y a de nombreuses approches possibles pour ce problème.

Le schéma général d'un système de reconnaissance des formes avec une phase d'apprentissage est donné par la figure suivante :

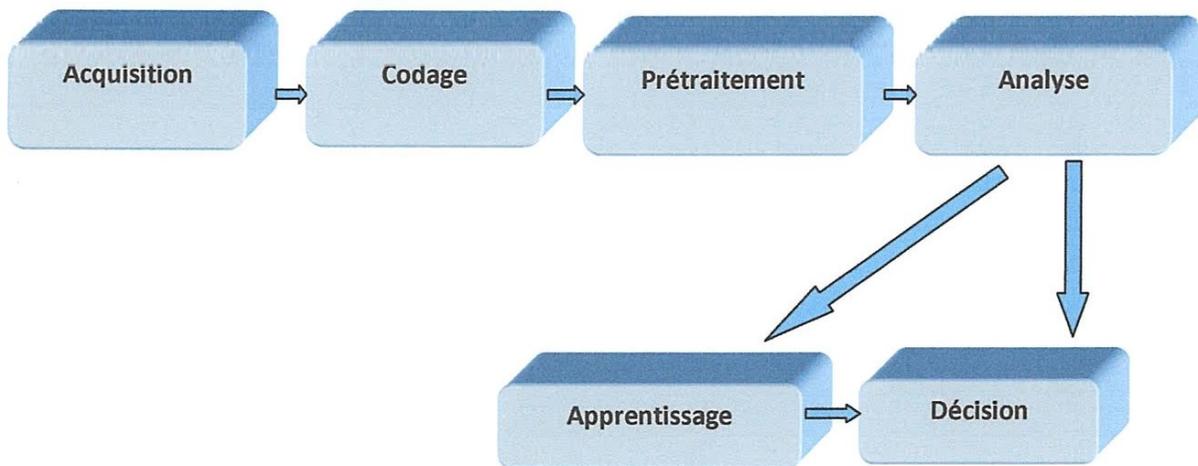


Figure -1- schéma général d'un système de reconnaissance des formes

Nous allons d'abord rappeler le rôle de chaque étape dans l'élaboration du processus complet de reconnaissance.



I.3.1-Etat réel et le codage

La chaîne part du monde physique qui est un espace analogique de dimension infini appelé espace de forme(f), dans ce vaste espace il est difficile de tenir compte de chaque objet seul lors du traitement d'un problème de reconnaissance de formes. La loi de passage au monde discret nécessite forcément une sélection et par conséquent une certaine simplification.

Les outils les plus utilisés pour la digitalisation des images (passage du domaine continu vers le domaine discret) sont les caméras et les scanners qui captent l'intensité de la lumière émise par un certain nombre de points de l'image, et la codent pour former un fichier selon différents formats.

C'est une opération de conversion numérique du monde physique continu vers un monde numérique discret de manière à pouvoir être traité, à transformer un ensemble de données analogiques en un ensemble de données numériques de manière à pouvoir être traité par ordinateurs. Cette transformation doit se faire sans perte d'information pertinente et en conservant les propriétés essentielles de l'objet physique, cette espèce est appelée aussi « espace de représentation » à une dimension encore trop importante même si elle est finie. La dimension de cet espace est choisie volontairement grande de manière à pouvoir disposer d'un maximum d'information sur la forme.

I.3.2- prétraitement

Les étapes de prétraitement consistent à sélectionner dans l'espace de représentation de l'information à l'application, cette sélection passe souvent par l'élimination du bruit des conditions d'acquisition, par la normalisation des données aussi bien que par la suppression de la redondance.

Le rôle du prétraitement est de préparer les données reçues du capteur à la phase suivante d'analyse consacrée à l'extraction des paramètres.

Les principales opérations de prétraitement sont :



I.3.3- Apprentissage

Dans le cas d'apprentissage il s'agit en fait de fournir au système un ensemble de formes qui sont déjà connues (on connaît la classe de chacune d'elles) c'est cet ensemble (apprentissage qui va permettre de « régler » le système de reconnaissances de façon à ce qu'il soit capable de reconnaître ultérieurement des formes de classe inconnue.

Il y a plusieurs cas possible :

- Des Classes Connues

On considère la classe de chaque forme de l'ensemble d'échantillons. On parle alors d'apprentissage supervisé (ou avec professeur).

- Des Classes Non Connues

On ne connaît la classe d'échantillons (ce qu'en sait c'est seulement que l'ensemble des échantillons représente ce qu'on doit connaître). On parle alors d'apprentissage non supervisé (ou sans professeur).

En fait le rôle du module d'apprentissage consiste à caractériser chaque classe par exemple par des relations entre les paramètres définissant la forme.

I.3.4- Analyse

Lors de l'étape d'analyse, les techniques de RdF calculent un certain nombre de caractéristiques ou paramètres. Ces paramètres correspondent à des mesures de natures géométrique, topologique ou statistique et servent comme seules données représentant la forme. Ils sont généralement limités en nombre.

Donc, le but de l'analyse est d'extraire les propriétés caractéristiques de l'objet et de les exprimer sous une forme numérique ultérieure d'apprentissage et de reconnaissance.

I.3.5- Décision

La décision ou le classement est l'étape de reconnaissances proprement dite, son rôle est d'identifier la formes teste à partir de l'apprentissage réalisé.

La méthode de décision est souvent celle utilisé par l'apprentissage, ce qui veut dire que les critères utilisé pour la comparaison sont les même que ceux utilisé pour l'apprentissage,



en effet il est évident que le choix de critère différent pour la décision, n'assure pas l'uniformité de la description ne peut pas conduire avec certitude à un résultat cohérent. Parmi la technique utilisée certaines sont fondées sur la notion de proximité et nécessitent de calcul une distance ou une probabilité de ressemblance avec les modèles définis. D'autres sont fondés sur l'analyse de la structure de la forme et essayant plutôt de vérifier certains 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 cas, la réponse peut être accompagnée d'un taux ou score de confiance.

I.4. Différentes Approches De Classification

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.

I.4.1 La méthode statistique

Se fonde sur une caractérisation des paramètres étudiés permettant de prendre la décision de classification d'une forme inconnue selon le critère «probabilité maximale d'appartenance à une classe».

Les exemplaires des classes correspondant aux observations d'une variable aléatoire X , chaque réalisation X est représentée par un vecteur de R^n . Chaque composante du vecteur correspondant 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 distributions de probabilité des exemplaires des classes.

I.4.2 La méthode de Bayes

L'approche bayésienne consiste à choisir parmi un ensemble des formes, celle pour laquelle la suite de primitives (observations) extraite a 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

$$P\{C|X\} = \frac{P(C)P(X|C)}{P(X)} \quad \text{I.1}$$

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 $P(X, C)$

$$P(X/C) = \text{MAX} P(X/C) * P(C). \quad \text{I.2}$$

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

$$P(X/C) = \prod_{i=1}^n P(x_i / c_i) \quad \text{I.3}$$

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 .



I.4.3 Classificateur paramétrique

- **Classificateur euclidien**

Il s'agit de l'un des plus simple classificateur qui puissent être conçus, la classe dont le vecteur de caractéristique moyen est le plus proche au sens de la distance euclidienne du les fonctions discriminante utilisé sont donc de la suivante :

$$\varphi_i(x) = \frac{1}{2} [(x - M_i)]^T (x - M_i)$$

- **Classificateur quadratique**

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

$$\varphi_i(x) = \frac{1}{2} [(x - M_i)]^T (x - M_i)$$

Où $M_i = E(XC_j)$ est le vecteur de caractéristique moyen des élément qui appartiennent a la classe C_i , $E\{\cdot\}$ désignant l'opérateur d'espérance mathématique ,et $\{\cdot\}$ celui de transposition le terme XX^T EST indépendante de la classe de la objet et les fonctions discriminantes peuvent également s'écrire :

$$\varphi_i(x) = M_i^T X - \frac{1}{2} M_i^T M_i$$

Les frontières qui séparent les classes dans l'espace R^d sont ici linéaires

I.4.4 Classificateur non paramétrique

- **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_i 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 a l'objet est alors du prototype le plus proche de celui-ci. Les fonctions discriminantes sont donc du forme :

$$\varphi_i(x) = -\min_k \frac{1}{2} [(X - X_k)]^T (X - X_k)$$



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

$$\varphi_i(x) = \min \left(X_K^T X - \frac{1}{2} X_K^T X_K \right) \quad X_K \in \varphi_i$$

Les frontières de décision entre classes linéaire par morceaux, c'est-à-dire constituées de nombreux petit polygone convexe chacune contenant un seul prototype d'une seule classe chaque classe est alors délimité 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 complexe lorsque suffisamment d'exemplaires de chaque classe sont disponibles.

- **Réseaux de neurones artificiels**

Un réseau de neurone est un outil statistique permet 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'exemple de chaque classe qu'il a appris un réseau de neurones est normalement capable de déterminer à quelle appartient un nouvel élément qui lui est soumis.

Pour pouvoir utiliser les capacités de classification d'un réseau de neurone; il faut suivre un certain nombre d'étapes :

- ❖ La construction de la structure du réseau.
- ❖ La construction d'une base de données de vecteur représentant au mieux le domaine à modéliser.
- ❖ Le paramétrage du réseau par apprentissage.
- ❖ La phase de reconnaissances.

- **La méthode de Markoviennes**

Les MMC ou modèle de Markov cachés (HMM : Hidden Markov Models), sont devenus une technique largement utilisée dans le domaine de la reconnaissance de l'écriture en-ligne et hors-lignes, ils ont été créés pour modéliser la production par une source cachée de séquence de signaux variable au cours du temps. Ces derniers sont engendrés par un processus aléatoire. Le premier processus peut ce trouver un instant donné dans un état appartenant à un ensemble de n .



- **Les méthodes mixtes**

Elles consistent à mettre en œuvre des systèmes à structure mixte comme par exemple un système combinant une règle floue et modèle de Markov ou encore carte Kohonen et règle floues.

I.5- Domaines D'applications

Nous citons tout d'abord quelques applications actuelles de la reconnaissance des formes et quelques-unes des performances atteintes.

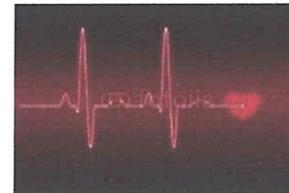
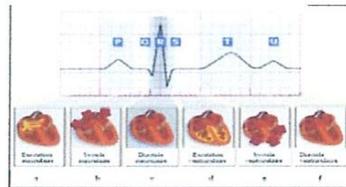
I.5.1 Reconnaissances des formes sur signaux temporels

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 étendu (quelques milliers de mots). On estime que d'ici une dizaine d'années, les machines seront capables de comprendre la parole humaine et auront de performances suffisantes pour être utilisables dans des tâches complexes (intégration de bases des données, acquisition de connaissance spécialisées,...etc.).

Un autre domaine extrêmement utile est celui des signaux biomédicaux ; la reconnaissance des formes permettrait d'automatiser ou de simplifier des tâches à la fois très complexes et très répétitives, comme un dépouillement électroencéphalogramme, ou encore assurer une surveillance automatique sur des mesures en temps réel. Un grand nombre d'études sont en cours dans ce vaste domaine, et déjà certains systèmes sont implantés en laboratoires hospitaliers.

Si l'on quitte les signaux physiologiques, la reconnaissance des formes s'intéresse aussi aux mesures des signaux d'origines artificielles : surveillance des machines, interprétation des échos,... etc. Un bon exemple est la détection d'objets sur signal radar.

- **Signal de parole :** a partir de ce signal la machine peut reconnaître le (locuteur qui a parlé), et aussi reconnaître la parole (qu'a-t-on dite). Cette dernière est liée à plusieurs disciplines comme le décodage acoustico-phonétique, la reconnaissance des mots, la reconnaissance de phrase, la reconnaissance du locuteur et compréhension du dialogue oral homme machine, parmi les applications les plus courantes on peut citer :
 - ✓ La commande vocale.
 - ✓ La dictée automatique.
 - ✓ La traduction en temps des langues étrangères.
- **Signaux biomédicaux**



1.5.2 Reconnaissances des formes dans l'image numérique

L'autre domaine prépondérant est celui de l'analyse et l'interprétation des images. Depuis les dessins les plus simples, comme les chiffres ; jusqu'aux images multi spectrales complexes issues de satellites, le champ des applications est immense.

Dans le domaine des images médicales, on trouve les problèmes de comptage de cellules ou de chromosomes, de sélection de radiographies, d'interprétation des résultats de tous les systèmes d'imagerie.

Un grand nombre d'images proviennent du domaine de la robotique, en particulier industrielle: reconnaissance de pièces pour saisie, par exemple. L'analyse des paysages est également très utile (photos aériennes, guidage en temps réel d'engins, etc.).

Classification d'empreintes digitales : les empreintes digitales sont intéressantes pour identifier les personnes.

La lecture automatique de texte manuscrit, un des vieux et difficiles problèmes de la RdF. Seuls des cas relativement simples sont résolus avec un taux d'erreur tolérable, Par

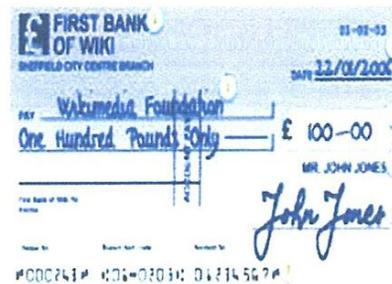


exemple la lecture des codes postaux pendant le tri du courrier et les montants littéraux de chèques.

- lecture automatique de caractères : parmi les applications de cette reconnaissance
- ✓ Reconnaissances des adresses



- ✓ Authentification des chèques bancaires



- la reconnaissance des empreintes digitales.



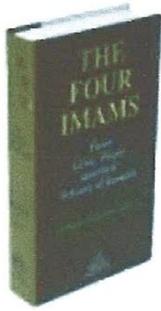
**I.6-Tableau d'Applications**

	Domaines D'études	Applications
Traitement Du	Reconnaissances de la parole	Bureautique, commande vocale, ordinateur sans clavier ni souris
	Reconnaissance du locuteur	Banques, commerces
Signal	Electrocardiogramme, électro-encéphalogramme	Médical
Traitement Des Images	Reconnaissances des caractères, du manuscrit	Saisies de texte, bureautique, tri postal, compression télécopie, chèque
	Reconnaissances des signatures	Banque, commerces
	Reconnaissance des empreintes digitales, des visages	Banque, commerces, police
	Analyse de radiographies, échographies, reconnaissances chromosomes, comptage globules	Contrôles systématiques de santé
	Détection de défaut circuits intégrés, pièces métalliques, manufacturées	Contrôle de qualité industriel.
	Identification d'objets	Tri d'objets industriels, surveillances militaire
	Localisation d'objets	Guidage de robots industriel, guidage missiles
	Analyse d'images de satellite	Météorologie, agriculture, ressources terrestres, surveillance militaire
Analyse de photos aériennes	Agriculture, surveillance militaire	
Analyse d'échos radar	Poursuite de cibles, pilotage missiles	



Conclusion

La tâche de reconnaissance du forme est de développer un système qui rapproche le plus de être humaine dans sa capacité. Ce domaine de reconnaissances de formes reste toujours à explore étant donnée sa complexité et sa diversité ; du point de vue de sa réalisation informatique. Dans ce chapitre j'ai présenté un aperçu général sur le processus de la reconnaissance de formes et les méthodes de classification ainsi que leurs domaines d'utilisation.



CHAPITRE II

la Reconnaissance De L'écriture Manuscrite

1. INTRODUCTION

II.2. LES DIFFERENTS ASPECTS DE RECONNAISSANCE DE L'ECRITURE

II.2.1 Reconnaissance en-ligne ou hors-ligne

II. 3. METHODES DE RECONNAISSANCE

II.3.1 Reconnaissance globale ou analytique

II.4. STRATEGIES DE RECONNAISSANCE

II.5. ORGANISATION GENERALE D'UN SYSTEME DE RECONNAISSANCE

II.5.1 Prétraitements

II.5.1.1 Binarisation

Seuillage

II.5.1.2 Filtrage

II.5.1.4 Extraction des contours

II.5.1.5 Extraction des composantes connexes

II.5.2 Segmentation

II.5.2.1 La segmentation des mots

II.5.2.2 Extraction des caractères

II.5.3 EXTRACTION DES PRIMITIVES (CARACTERISTIQUES)

a) Caractéristiques géométriques

b) Caractéristiques structurelles

c) Caractéristiques Statistiques

- Les caractéristiques adaptées à l'écriture

II.5.3.1 Détection des caractéristiques structurelles

- Localisation des hampes et jambages
- Les boucles

II.5.3.2 Détection des caractéristiques statistiques

- Caractéristiques issues de la technique de zonage « ZONING »

II.5.3.3 Détection des caractéristiques géométriques

- Codage des contours
- Extraction des profils
- Description de formes au moyen de moments invariants
- Les moments invariants
- Description de formes par descripteur de Fourier
- Aspects théoriques de la TF
 - ✚ TF discrète (TFD)
 - ✚ Définition complexe de descripteurs de Fourier

II.5.4 CLASSIFICATION

CONCLUSION



1. Introduction

De nos jours, l'écriture est toujours le moyen de communication visuelle le plus utilisé par l'homme. Il n'est donc pas surprenant de voir que de nombreux travaux scientifiques portent sur sa reconnaissance automatique. L'écriture est en fait la réalisation d'un message à transmettre, c'est-à-dire la représentation physique d'un contenu sémantique. Le média ou support généralement utilisé est le papier. Le but de la reconnaissance de l'écriture est de prendre une décision quant au contenu sémantique du message transmis à partir de sa représentation physique. Les applications de systèmes capables de remplir cette tâche sont nombreuses ; nous pouvons citer entre autres la lecture automatique de bons de commande, le traitement automatique des chèques, la vérification de signatures ou encore le tri automatique du courrier.

La reconnaissance de l'écriture est rattachée au vaste domaine de la reconnaissance de formes. Sa spécificité vient bien sûr des données à analyser et de leurs diverses sources de variation. En effet, un même mot écrit par plusieurs personnes peut avoir des formes assez différentes. Pour cette raison, les caractéristiques extraites de l'écriture sont très importantes pour la suite du processus de reconnaissance. Ce dernier quant à lui, peut être mis en œuvre à l'aide de presque toutes les techniques développées en reconnaissance de formes.

II.2. les Différents aspects de reconnaissance de l'écriture

La reconnaissance optique des caractères (OCR) est un processus qui permet de convertir un texte écrit sur papier en texte sous forme d'un fichier informatique en représentation symbolique. Pour l'écriture latines, le codage typiques set opéré par le code ASCII (American Standard Code for Informaticien Intercharacter), tandis que pour l'arabe on utilise généralement l'ASMO (arabic Standard Metrology Organisation).

Par ailleurs, il n'existe pas de système universel d'OCR (Optical character Recognition) qui peut traiter tous les cas d'écriture, mais plutôt des voies d'approche dépendant du types de données traitées et bien évidemment de l'application visée [2]. Il existe différents modes de classification des systèmes OCR, parmi les principaux, nous citons :

Les systèmes qualifiés de « *en-ligne* » ou « *hors-ligne* » relatifs au mode *d'acquisition* pour cela il existe un grand nombre de techniques classées généralement en deux grandes catégories :

- Les approches *globales* et *analytiques* selon que l'analyse s'opère sur la totalité du mot, ou par segmentation en caractères (ou en parties de caractères) et analyse individuelle de ceux-ci.
- Les approches *statistiques*, *structurelles* et *stochastiques* relation que traits caractéristiques extraits des formes considérées relatifs au mode de *classification*.

II.2.1 Reconnaissance en-ligne ou hors-ligne

Ce sont deux modes différents d'OCR ayant chacun ses outils propres d'acquisition et ses algorithmes correspondants de reconnaissance.

- **La reconnaissance en-ligne (on line)**, dite encore dynamique s'opère en temps réel pendant l'écriture : les symboles sont reconnus au fur et à mesure qu'ils sont écrits à la main.

Ce mode, réservé généralement à la reconnaissance de textes manuscrits, présente plusieurs avantages. C'est une approche « signal » où la reconnaissance effectuée sur des données ou une seule dimension, l'écriture présente par une séquence de points dont les coordonnées sont fonction du temps ; ainsi l'ordre de l'écriture, l'information temporelle, les levers du stylo sont exploités reconnaissance. De plus, la réponse en contenu du système permet à l'utilisateur de corriger et de modifier son écriture de manière interactive. L'acquisition du tracé est assurée généralement par une tablette graphique munie par d'un stylo électronique, les coordonnées du point cursif .une résolution typique de 200 dpi (*dot per inche*) et une fréquence d'échantillonnage de 100 points par seconde, format le plus caractéristique les plus usitées.

- **La reconnaissance hors-ligne (off line)** ou en différé dite encore *statique*, démarre par l'acquisition. Elle convient pour les documents imprimés et les manuscrits déjà rédigés. Le mode hors-ligne peut être considéré comme le cas le plus général de la reconnaissance d'une écriture. Il se rapproche du mode de reconnaissance visuelle : l'interprétation de l'information est indépendante de source de génération, ce qui laisse les champs ouverts à différents types d'applications

La reconnaissance hors-ligne relève d'une approche de traitement généralement « image-2D », ce qui constitue une différence fondamentale avec la reconnaissance en-ligne classique. La quantité d'informations manipulées dans les deux cas est nettement différente. A titre d'exemple, un mot cursif écrit en mode en-ligne (échantillonné à 100 point par seconde) nécessite environ 230 octets, alors que le mode hors-ligne, il occupe 80 K octets (numérisation à une résolution 300dpi) [5], les taux réalisées en mode en ligne sont



généralement plus élevés que dans le cas hors ligne pour des conditions de teste similaires en se basent sur les divers degrés de complexité des problèmes à traiter.

II. 3. Méthodes de reconnaissance

II.3.1 Reconnaissance globale ou analytique

- *L'approche globale*, dit encore holistique considère le mot comme une seule entité et le décrit indépendamment des caractères qui le constituent. Cette approche présente l'avantage notamment de garder le caractère dans son contexte avoisinant qui permet une modélisation plus efficace des variations de l'écriture et des dégradations qui peuvent l'entacher. Cependant cette méthode est pénalisée par la taille mémoire, le temps de calcul et la complexité du traitement croient linéairement par la taille du lexique considéré, d'où une limitation du vocabulaire.
- *L'approche analytique* contrairement à l'approche globale, le mot est segmenté en caractère ou en fragments morphologiques significatifs inférieurs aux caractères appelés *graphèmes*. la reconnaissance consiste alors à reconnaître les entités segmentées puis tendre vers une reconnaissance du mot, ce qui consiste une tâche assez délicate pouvant générer différents types d'erreurs.

II.4. Stratégies de reconnaissance

Les principales applications de la reconnaissance de l'écriture manuscrite sont la lecture des montants littéraux de chèques et la lecture des enveloppes postales. Dans les deux cas le système doit reconnaître un ensemble de mots. À ce moment deux stratégies peuvent être envisagées : reconnaître les mots séparément ou alors essayer de reconnaître le groupe de mots.

Dans la plupart des systèmes existants la reconnaissance des mots est tout d'abord mise en œuvre, puis la combinaison des différents mots permet d'obtenir un résultat pour le groupe de mots. Dans le cas de la reconnaissance des montants littéraux de chèques cette dernière étape permet d'effectuer une vérification grammaticale et ainsi de proposer seulement des montants grammaticalement corrects. Dans le cas de l'application à la reconnaissance des adresses le fonctionnement est similaire : la reconnaissance se fait au niveau du mot et non pas de la phrase entière.

II.5. Organisation générale d'un système de reconnaissance

L'objectif de la reconnaissance de l'écriture manuscrite est de développer un système qui se rapproche le plus de l'être humain dans sa capacité de lire. Cependant, cette reconnaissance de l'écriture consiste à extraire d'une forme inconnue (mot, lettres, chiffres) une description plus simple et à établir sur celle-ci une décision. Cette décision est effectuée généralement en mesurant la ressemblance d'une forme inconnue avec un ensemble de références stockées en mémoire et décrites dans une représentation analogue. Les références sont obtenues lors d'une phase intérieure qualifiée d'apprentissage. Cette phase est très importante dans tout système de reconnaissance de l'écriture. Autrement dit c'est un passage de l'espace observable vers un espace de décision d'appartenance à une classe.

La construction d'un système de reconnaissance de l'écriture comprend plusieurs étapes distinctes représentées par la *figure II.1*.

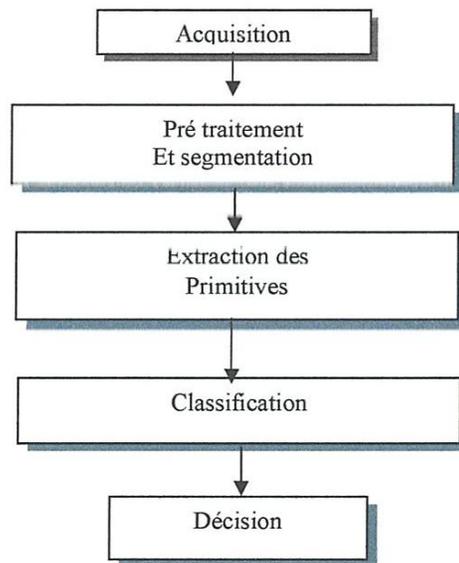


Figure II.1 : les étapes de reconnaissance d'écriture

Par la suite plusieurs traitements sont opérés sur ces images et ces fichiers. Le but de ces prétraitements est d'éliminer les phénomènes qui provoquent une dégradation des performances du système, de réduire les bruits de quantification (binarisation) et de préserver la connexité des composantes connexes dans l'image. Le résultat de cette phase permettra d'extraire ou de mettre en évidence des particularités locales ou globales. Cette étape permet d'engendrer pour chaque image, un vecteur de primitives qui sert comme entrée au module responsable de la classification.



Dans la section suivante nous abordons l'état de l'art de chaque phase de la reconnaissance à savoir : les prétraitements, l'extraction des primitives et la classification.

II.5.1 Prétraitements

L'objectif des prétraitements est de faciliter la caractérisation de la forme (caractère, chiffre, mot) ou de l'entité à reconnaître soit en nettoyant l'image représentant la forme ou en réduisant la quantité d'information à traiter pour ne garder que les informations les plus significatives. Le nettoyage de l'image consiste essentiellement à éliminer les bruits résiduels issus de la binarisation. La réduction de la quantité d'information à traiter peut être obtenue à partir des opérations visant à ramener l'épaisseur du trait à un seul pixel (soit par squelettisation, soit par suivi de trait), ou à partir d'extracteurs de contours supérieurs, inférieurs et/ou intérieurs.

Les différentes étapes des prétraitements sont :

II.5.1.1 Binarisation

Pour les images acquises en niveaux de gris, la binarisation devient nécessaire avant d'attaquer la phase du traitement. La binarisation permet à 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 texte), 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 tout les pixels ayant un niveau de gris inférieur à cette valeur sont représenté par un zéros "0" le noir, et tout les pixels de niveaux de gris supérieur auront la valeur un "1" (le blanc).

La valeur du seuil est déterminée à partir de l'histogramme de niveaux de gris de tous les pixels de l'image, elle sera égale à la valeur de niveaux de gris se trouvant dans la vallée entre les deux pics de l'histogramme *figure II.2*.

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

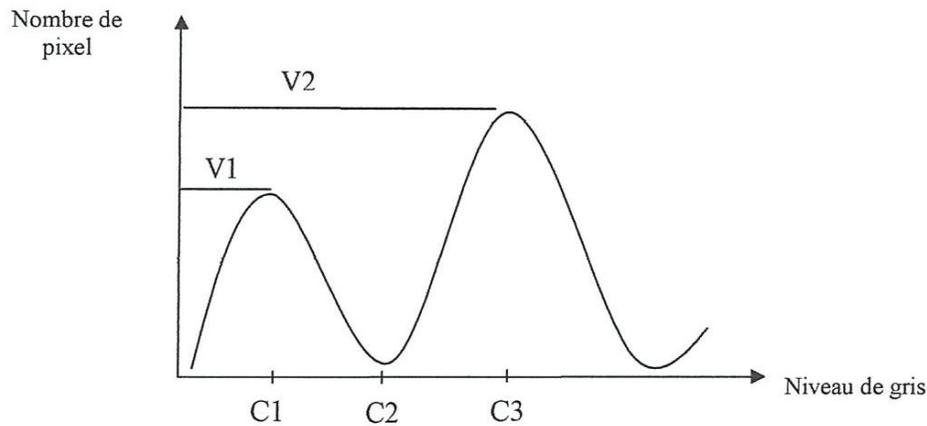


Figure II.2 : histogramme de niveaux de gris

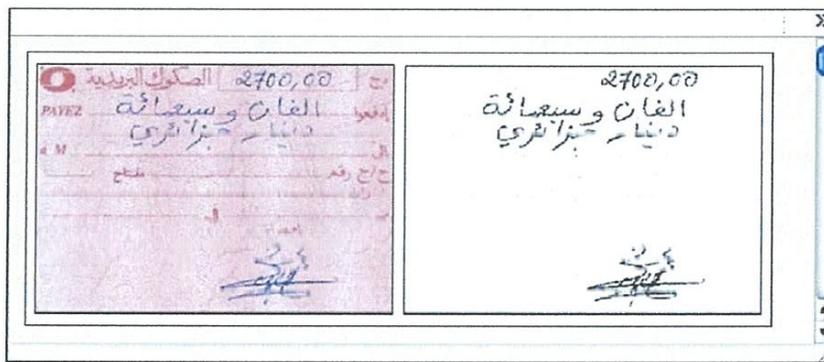


Figure II.3 : Exemple de binarisation

II.5.1.2 Filtrage

L'image de caractères peut être entachée de bruit dont l'origine est la qualité du document ou le mode d'acquisition utilisé, ce qui conduit soit à des absences de points (trous), soit des points en plus (des empâtements). Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales appelées *Nettoyage* et *bouchage*.

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 dont les huit directions

0	0	0
X	1	0
0	0	0

- Par élimination des points formant des angles droits ou des ex croisement
Du contour, en appliquant le masque suivant dans le huit directions :

X	0	0
1	1	0
0	0	0

Pour le *bouchage*, il s'agit d'égaliser les contours et de boucher les trous internes à la forme des caractères en lui ajoutant des points noirs on distingue :

- Le bouchage 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

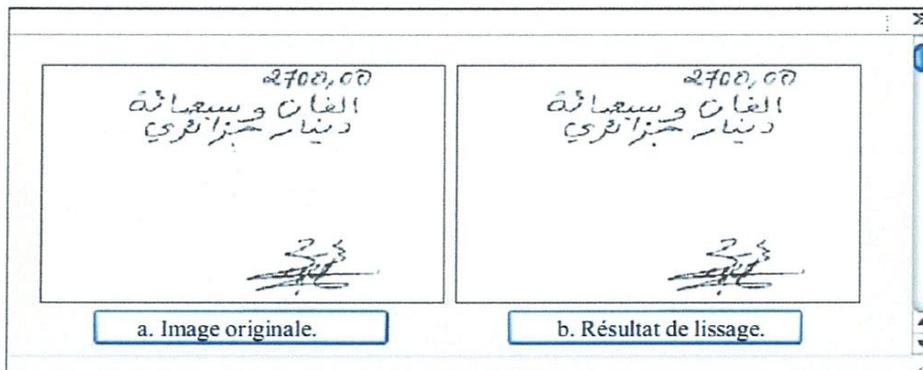


Figure II.4. Exemple de lissage par les filtres présentés.

II.5.1.3 Normalisation de la taille

La taille des caractères peut varier d'une fonte à l'autre, et même au sein d'une même fonte après agrandissement ou réduction, ce qui peut causer une instabilité des paramètres.

Une technique de prétraitement consiste à ramener les caractères à la même taille, tel que l'algorithme de *Sihari*.



Cet algorithme opère en deux étapes. La première normalise le caractère en hauteur et la seconde, en largeur. L'ordre de normalisation ainsi choisi évite que les caractères fins ne se déforment pas.

Il s'agit de transformer l'image du caractère de dimensions h_i, l_i en une image de dimensions h, l . Une étape intermédiaire consiste à produire une image de dimension h', l' , avec :

$$P = \frac{h}{h_i} \text{ Et } l' = p.l \quad \text{II.1}$$

La normalisation en hauteur est exécutée en transformant chaque pixel (x, y) noir de l'image du caractère en $(p.x, p.y)$.

La normalisation en largeur de l'image ainsi obtenue se fait par l'examen de deux cas.

Si $l' < l$, alors l'image normalisée en hauteur est centrée dans une surface de dimension h, l .

Si $l' > l$, chaque pixel noir de l'image (x, y) subit la transformation en $((x/l'), l, y)$.

II.5.1.4 Extraction des contours

On appelle point de contour, un pixel qui se situe sur la frontière séparant deux régions voisines.

Dans les images à niveaux de gris, il est intéressant d'extraire le contour à l'aide d'un calcul de *gradient*. Ce contour est alors d'autant plus marqué que le niveau des pixels résultant du gradient est élevé. En revanche, dans les images binaires, il est plus avantageux d'utiliser un algorithme de *suivi de contour* car il fournit directement une liste ordonnée de points.

a) Le gradient :

Le contour se manifeste dans l'image par des variations locales importantes des valeurs de niveaux de gris mis en évidence par d'élévations de la dérivée première de la fonction image

Le gradient d'une fonction $f(x, y)$ est défini par $\nabla f(x, y) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j}$ II.2

\vec{i} et \vec{j} sont les vecteurs unitaires sur x et y .

Dans le domaine discret les dérivées suivant x et y peuvent être exprimées par les approximations suivantes :

$$A_x = f(x+1, y) - f(x, y) \text{ suivant } x$$

$$A_y = f(x, y+1) - f(x, y) \text{ suivant } y$$

L'amplitude du gradient est donc $A = \sqrt{A_x^2 + A_y^2}$ II.3

La **figure II.5** montre l'exemple d'un mot avant et après l'extraction de ses contours.

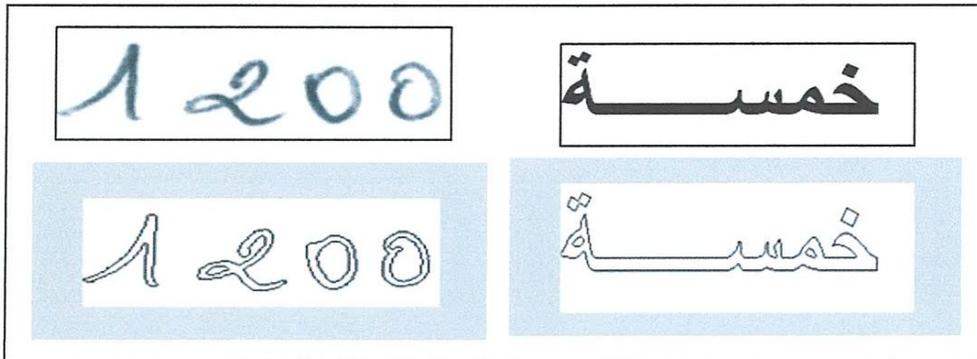


Figure II.5 Exemple d'extraction de contour

b) Procédure de suivi de contour

L'algorithme de traçage de contour utilise le concept de 8-voisinage, ou 4-voisinage. Il peut être décrit comme étant un observateur qui traverse les pixels de l'image du texte (la composante connexe) d'un pixel à un autre, et sélectionne le plus à droite. Le point de départ peut être détecté en parcourant la composante de droite à gauche haut en bas. L'algorithme s'achève quand le pixel courant se coïncide avec celui de départ. Ce type d'algorithme retrouve seulement le contour extérieur de la composante connexe *figure II.6*.

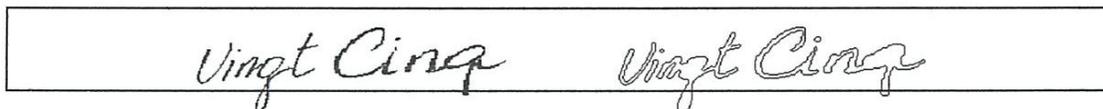


figure II.6 : Exemple du mot « vingt cinq » avant et après l'extraction de contours extérieur

Le contour est également utilisé comme étape préalable à un changement de représentation de l'information, en tant qu'empreinte des formes contenant une quantité réduite de données.

II.5.1.5 Extraction des composantes connexes

L'extraction des composantes connexes, procédure également appelée capture des connexités ou étiquetage des pixels, est largement utilisée en Reconnaissance des Formes (RdF) pour segmenter les images binaires. La technique consiste à regrouper les pixels voisins dans un ensemble appelé composante connexe. Chaque ensemble est disjoint des autres et peut ensuite être aisément isolé. La 4-connexité est distinguée de la 8-connexité suivant que le critère de voisinage comprend les 4 ou les 8 voisins d'un pixel.



Il existe deux principaux algorithmes pour accomplir cette tâche :

- le premier est basé sur une procédure de suivi de contour: en parcourant le contour d'un objet et en revenant au point de départ, une composante connexe est délimitée, à l'exclusion cependant des contours intérieurs correspondant aux éventuels trous.

- le second algorithme procède par une propagation d'un étiquetage des pixels lorsque l'on effectue un balayage des lignes et des colonnes de l'image.

Nous avons élaboré un algorithme de ce type fonctionnant en une seule passe, suivant le critère de 4-connexité

La propagation de l'étiquette des pixels suivant les colonnes (verticalement) est prioritaire sur la propagation suivant les lignes (horizontalement). On procède de la manière suivante :

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 II.6a*.

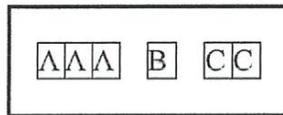


Figure II.6a : Propagation de l'étiquetage des pixels de gauche à droite sur une ligne horizontale

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 II.6b*.

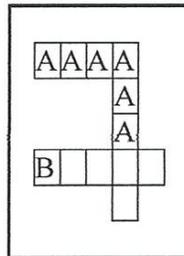


Figure II.6b : Propagation de l'étiquetage des pixels de haut en bas sur une colonne verticale

Lorsqu'un conflit se présente entre la propagation horizontale et la propagation verticale des étiquettes, deux cas se présentent alors *Figure II.6c* :

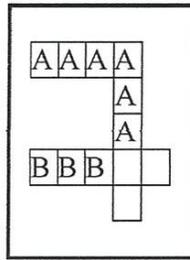


Figure II.6c : Conflit entre la propagation horizontale et verticale

1^{er} 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 II.6d*. La nouvelle étiquette est alors annulée ;

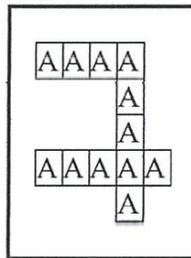


Figure II.6d : Résolution du conflit correspondant au 1^{er} cas

2^e 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 II.6e*.

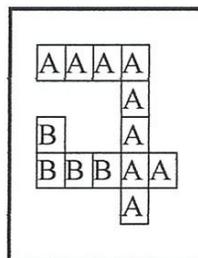


Figure II.6e : Résolution du conflit correspondant au 2^e cas

II.5.2 Segmentation

Après la phase de prétraitement, plusieurs systèmes de reconnaissance divisent la page du texte en lignes, chaque ligne en mot, et certains systèmes divisent le mot en caractères ou en segments de caractères dans le but d'aborder les phases ultérieures, ce découpage est appelé segmentation.

Elle est définie comme étant un processus qui consiste à diviser l'image en régions à nuance rapprochées, l'idée de base est de regrouper dans la même région les pixels possédants la même valeur de niveau de gris.

Pour la reconnaissance de l'écriture la segmentation se fait séquentiellement en deux phases : décomposition des pages en lignes pour *extraire les mots* et *l'extraction des caractères*. Pour cela nous avons appliqué des méthodes de segmentation basée sur

- analyse de l'histogramme horizontal et vertical.
- Les minimaux des contours extérieurs
- L'extraction des composantes connexes

a)- segmentation des mots elle regroupe deux étapes :

- *Séparation des lignes* : localisation des lignes de texte.
- *Séparation des mots* : séparation des lignes en éléments constitutifs.

b)- segmentation n des caractères

- *La segmentation caractères* : séparation des parties connexes du mot en caractères.

II.5.2.1 La segmentation des mots

Après l'étape d'acquisition et prétraitement, l'image obtenue est une matrice binaire. L'extraction des mots se fait en deux étapes distinctes : une séparation de lignes, suivie à une séparation de mots, la séparation de lignes se fait par projection horizontale qui ne rien d'autre qu'une simple somme des nombres de points allumés par ligne, La séparation de mots se fait par projection verticale, une simple somme des nombres de points allumés par colonne.

Les projections horizontale et verticale sont définies par les équations suivantes :

$$h(i) = \sum_j g(i, j) \quad : \text{Pour la projection horizontale.}$$

$$v(i) = \sum_i g(i, j) \quad : \text{Pour la projection verticale.}$$

Ou $g(i,j)$ vaut 0 ou 1

a) Séparation des lignes

La détection des lignes de texte commence une fois que la projection horizontale de toutes les lignes d'image est faite.

On peut dire le début ou la fin d'une ligne de texte est détectée, si la valeur de projection horizontale (P.H) *figure II.7*, est inférieure à un seuil.

Le seuil est obtenue au minimum de l'histogramme (la fonction h) La confirmation de cette détection se fait si le nombre de ligne d'image détectées est adjacentes est lui même supérieur à un autre seuil. C'est de cette façon qu'on confirme l'existence d'une ligne de texte.



(a) (b)
Figure II.7. a-Exemple de localisation des lignes d'écriture.

b-Correction de l'inclinaison

b) La séparation des mots

La majorité des études consacrées à la segmentation des mots dans la phrase reposent sur l'hypothèse d'existence d'un espace inter-mots toujours supérieur à l'espace inter-lettres ou à la largeur moyenne d'une lettre.

La segmentation des mots permet de localiser les bornes de chaque mot de la ligne isolée, les bornes des mots peuvent être déterminées en effectuant une projection vertical et en cherchant l'espace noire qui est représentés par les pixels de valeur supérieur a zéros, le début et la fin de la suite de ces espaces noire entre deux espace blanc représente les bornes, *Figure II.8*.

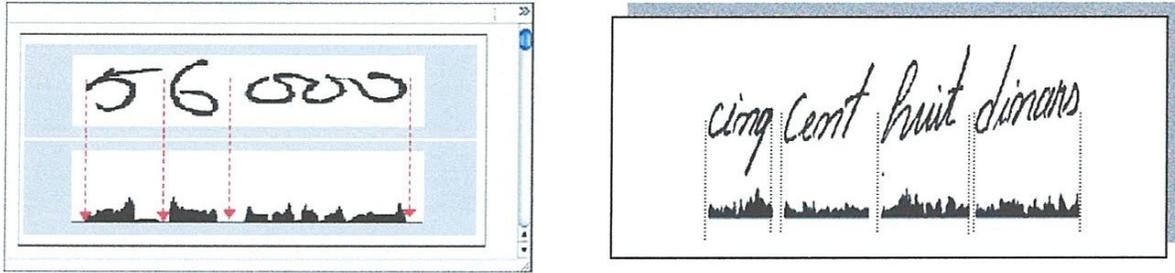


Figure II.8 : histogramme de la segmentation de mots

II.5.2.2 Extraction des caractères

Les études dans le domaine de la segmentation du mot en caractère sont principalement basées sur l'analyse du contour. Les critères de segmentation que nous avons répertoriés sont liés à la détection des minimums du contour supérieur du mot *figure II.9 (c)*. Les hypothèses de base sont les suivantes :

- les caractères composant le mot ne sont reliés que par un seul trait ;
- la segmentation peut être effectuée par une coupure verticale *figure II.9 (d)* située au minimum local du trait de connexion suivi par une procédure de capture des composantes connexes.

La détection de ce minimum local est obtenue par le calcul de la dérivée du contour, c'est-à-dire, dans l'espace discrétisé, de la différence des ordonnées de deux points d'abscisses consécutives. Puis le minimum est validé en fonction des critères d'unicité du trait de connexion et de son épaisseur, afin de ne pas couper un caractère comportant un minimum local à l'intérieur de son tracé.

Afin de renforcer les hypothèses de base, plusieurs autres critères ont été proposés dans différentes études incluant la segmentation des mots manuscrits :

On peut aussi appliquée la technique des composantes connexes, procédure également appelée capture des connexités ou étiquetage des pixels, La technique consiste à regrouper les pixels voisins dans un ensemble appelé composante connexe. Chaque ensemble est disjoint des autres et peut ensuite être aisément isolé *figure II.9 (e)*.

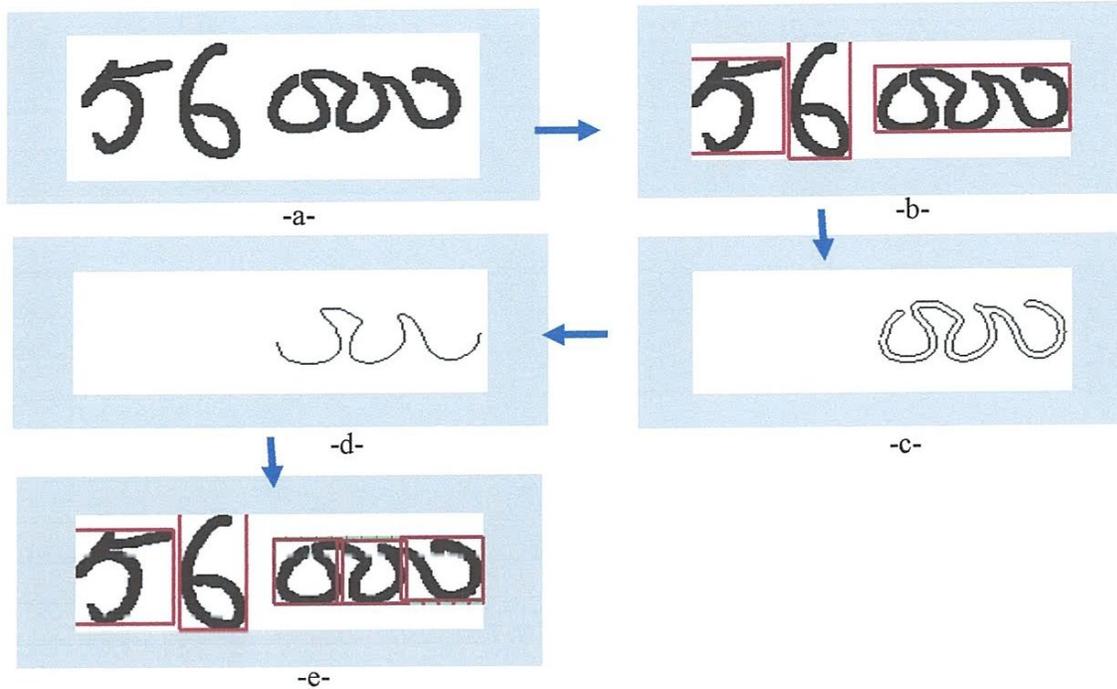


Figure II.9e: Extraction des chiffres

- a) image du chiffre '56000'
- b) extraction des composantes connexes
- c) contour inférieur
- d) détection des points de coupure
- e) localisation des chiffres

II.5.3 Extraction des primitives (caractéristiques)

Un des problèmes fondamentaux de la reconnaissance d'écritures est de déterminer quelles caractéristiques employer pour obtenir un bon résultat de classification.

L'objectif de l'extraction et de la sélection des caractéristiques est d'identifier les caractéristiques qui sont importantes dans la discrimination de classes de formes. Cela signifie de trouver des primitives les plus robustes possibles, les plus pertinentes (afin de diminuer la taille du vecteur de primitives) et les moins coûteuse en calculs.

Le résultat de cette phase d'extraction de primitives est une séquence de symboles représente l'image binaire. Cette séquence est représentée sous forme d'un vecteur qui constituera les données d'entrée de la phase de reconnaissance.

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.

Au-delà de cette classification un peu formelle, la différence entre les approches systématique et heuristique comme entre le caractère a priori ou a posteriori, s'avère plus nuancée dans la pratique.

Les primitives sont classées ont trois catégories sont *géométriques, structurelles, et statistiques*.

a) Caractéristiques géométriques

Les mesures géométriques ou physiques sont généralement des mesures de distance entre points appartenant ou non au mot et qui permettent de caractériser ses propriétés métriques.

Parmi les caractéristiques géométriques les plus utilisées on retrouve

- Les moments géométriques, invariants en changement d'échelles, rotation et translation.
- Descripteur de Fourier ces caractéristiques sont basées sur des informations extraites à partir de l'image du contour
- Description locale des contours (Codage des contours) est une technique exprimant les frontières discrètes de l'image par une séquence de code comme la chaîne de Freeman spécifiant la direction de déplacement d'un vertex à un autre.
- Description des profils des caractères.



b) Caractéristiques structurelles

Les caractéristiques structurelles décrivent le mot dans sa globalité, parmi les plus utilisées dans la reconnaissance globale de mots on trouve :

- Les hampes et les jambages qui sont généralement des primitives recherchées dans la perception humaine
- Les boucles et les cavités qui permettent aussi de distinguer entre les mots qui possèdent des primitives presque similaires.
- Les points diacritiques, car ils permettent de distinguer les caractères ayant le même corps principal.

c) Caractéristiques Statistiques

Les mesures de densité de pixels noirs ou blancs sont généralement utilisées pour décrire l'influence de certaines parties du mot et donc la présence éventuelle de traits caractéristiques. La plus part des méthodes utilisant ce type de mesures procèdent à un découpage préalable de l'image en zones et ce sont les densités de pixels dans chaque zone qui forment les composantes du vecteur de caractéristiques.

• Les caractéristiques adaptées à l'écriture

La détection des caractéristiques est l'une des étapes les plus délicates dans la construction d'un système de reconnaissance de caractères manuscrits.

Nous présentons dans cette partie les principales caractéristiques et leurs détections.

II.5.3.1 Détection des caractéristiques structurelles

• Localisation des hampes et jambages

Grâce à la connaissance de la zone médiane du mot, chaque caractère est identifié à une classe générique parmi cinq.

Les cinq classes génériques sont notées 'H', 'J', 'f', 'M', et '?' respectivement Hampes, Jambages, hampes et jambages, Médiants et ambigus *figure II.10*.

Une hampe est détectée si l'ordonnée supérieure du caractère est située dans la zone supérieure à la zone médiane du mot à une distance supérieure à la marge (Position absolue).

Un jambage est détecté si l'ordonnée inférieure du caractère est située dans la zone inférieure à la zone médiane du mot à une distance supérieure à la marge (Position absolue).

Le caractère est classé Médian si sa hauteur est inférieure à celle de la zone médiane du mot (hauteur relative), à condition toutefois que le graphème ne dépasse pas les zones définies par les marges (Position absolue).

Sinon, si le caractère ne vérifie aucune de ces conditions, il est classé ambigu.

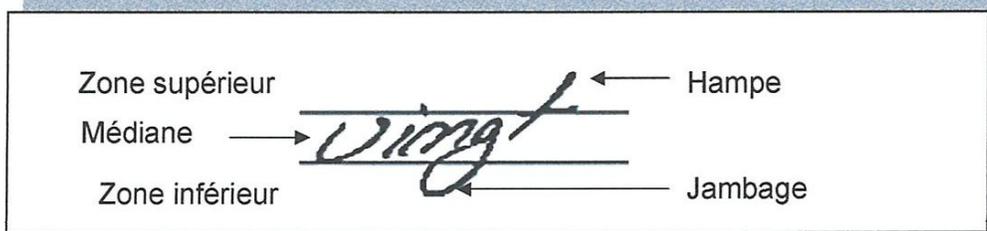


Figure II.10 Exemples de Localisation des hampes, jambages et médiane du mot 'vingt'

- Les boucles

Il y a deux méthodes permettant d'extraire les boucles : celles qui procèdent par l'étiquetage des zones blanches de l'image et celle qui opèrent sur le contour de la forme. L'analyse des boucles par étiquetage (la connexité) des zones blanches de l'image consiste à associer à chaque zone blanche une étiquette, ensuite une procédure de remplissage par des zéros (noir) de tous les zones a une contacte avec les limites du cadre de l'image *Figure II.11*.



Figure II.11 : Extraction des boucles méthode d'étiquetage

La méthode de détection qui opère sur les contours du caractère nécessite une procédure de suivi de contour, en parcourant le contour d'un objet et en revenant au point de départ, une composante connexe est délimitée, à l'exclusion cependant des contours intérieurs correspondant aux éventuels trous *figure II.12*.



Figure II.12 : Extraction des boucles méthode basée sur les contours

II.5.3.2 Détection des caractéristiques statistiques

- **Caractéristiques issues de la technique de zonage « ZONING »**

Une information globale caractérisant la répartition des points constituant le caractère. Dans la fenêtre minimale est également exploitée. La fenêtre englobant le caractère ou le mot étudié est divisée en régions de façons différentes *figure II.13* Pour chaque découpage, on détermine la répartition des points de l'image entre les régions constituant ce découpage.

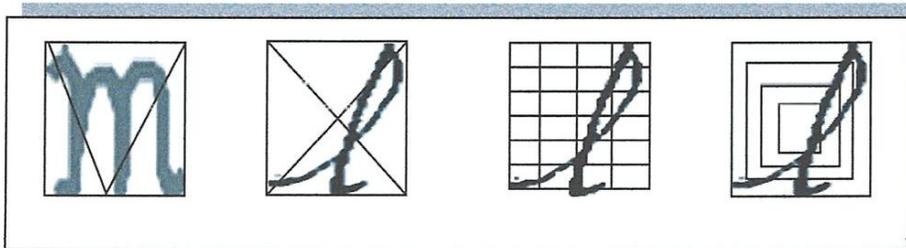


Figure II.13 : Division de l'image en des zones de quatre manières

Ce type de caractéristiques a été proposé pour problème mono scripteur.

II.5.3.3 Détection des caractéristiques géométriques

- **Codage des contours**

Le codage des contours est une technique exprimant les frontières discrètes de l'image du caractère par une séquence de code ou une suite des symboles. Un premier type de symboles donne la position absolue d'un point appartenant au contour appelé *point de départ*. Un deuxième type décrit la position relative de tous les autres points des contours *figure II.14* pris séquentiellement dans un sens défini arbitrairement.

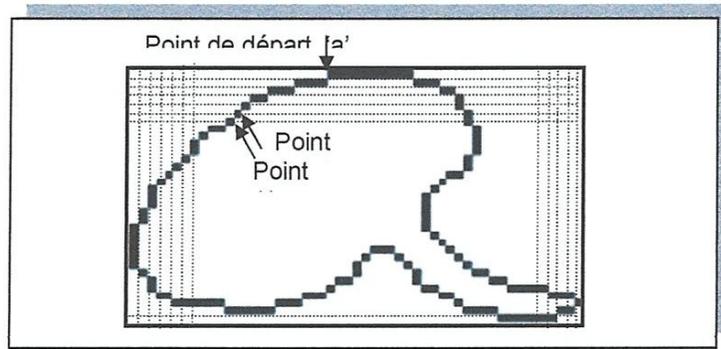


Figure II.14 : Représentation d'un contour de la lettre « a » à la moyenne
D'une grille

Il existe une technique qui permet de définir la position relative d'un ensemble des points à savoir le code de chaîne ou codage de Freeman.

Dans ce code la position d'un nouveau point voisin $i+1$ par rapport au point i , connexe de connectivité 4 ou 8, s'exprime au moyen d'un mot binaire à 2 ou à 3 bits.

Le nombre spécifique indiqué par le mot binaire correspond à l'une des directions qui relie les deux points comme indiqué à la *figure II.15*

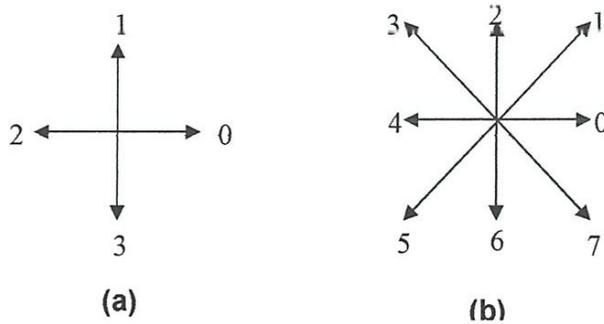


Figure II.15 Définition du code de chaîne : (a) Code de chaîne de connectivité 4;
(b) code de chaîne de connectivité 8 ;

15 de connectivité 8 en prenant le point 'a' comme point de départ et le sens des aiguilles d'une Montre comme sens de parcours est :

```
C8=0000007007776766665554566655454454453444444344333222222212266666777707
07000700007654445443444344333344550000001211111111101010001000000700777
67666655545660000001211111111101010001.
```

- **Extraction des profils**

L'utilisation des profils pour caractériser les caractères manuscrits a été suggérée par divers auteurs ([7] par exemple). Le caractère est encadré. Dans une fenêtre minimale. Le profil compte le nombre de Pixel (distance) entre la boîte de bondissement d'image de caractère et le bord du caractère. Le profil d'un caractère peut être pris à n'importe quelle position, gauche, droite, supérieur, inférieur et orienter *Figure II.16*.

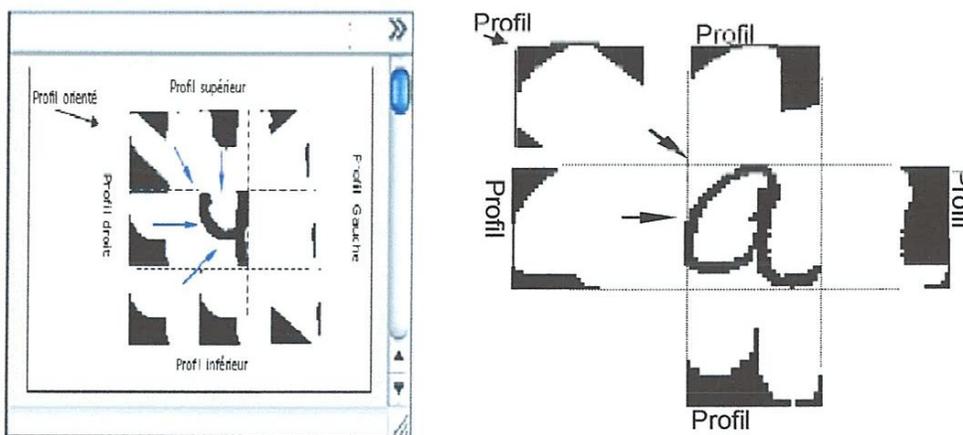


Figure II.16. Exemple des différentes projections des profils pour « le chiffre 4 et la lettre a »

Les profils décrivent bien les formes externes des caractères et laissent distinguer entre un grand nombre de lettres.

- **Description de formes au moyen de 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 ont le mérite de répondre à trois critères qui sont.

- ✓ invariance par rapport à la translation,
- ✓ invariance par rapport au changement d'échelle,
- ✓ invariance par rapporte à la rotation.

• **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 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:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) . dx dy \quad \text{Avec} \quad p \text{ et } q \in \mathbb{N} \quad \text{II.4}$$

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

$$\mu_{pq} = \iint_{-\infty-\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy. \quad \text{II.5}$$

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

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Et μ_{pq} est le moment centralisé d'ordre (p, q) .

Pour une image discrète les relations (II. 4) et (II. 5) deviennent

$$m_{pq} = \sum_x \sum_y (x)^p (y)^q f(x, y) \quad \text{II.6}$$

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

II.7

Ce qui donne pour p et q = 0, 1 :

$$\mu_{00} = m_{00}$$

$$\mu_{10} = \mu_{01} = 0$$

$$\mu_{11} = m_{11} - (m_{00} \cdot \bar{x} \cdot \bar{y})$$

Les moments centrés et normalisés sont définis par :

$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \text{ avec : } \gamma = \frac{p+q}{2} + 1$$

Afin de rendre les moments φ_i invariants par rotation et changement d'échelle, les moments suivants ont été introduits :

$$\varphi_1 = n_{20} - n_{02}$$

$$\varphi_2 = (n_{20} - n_{02})^2 - 4n_{11}^2$$

$$\varphi_3 = (n_{30} - 3n_{12})^2 + (3n_{11} - n_{03})^2$$

$$\varphi_4 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2$$

$$\varphi_5 = (n_{30} - 3n_{12})^2 (n_{30} - n_{12}) \left[(n_{30} + 3n_{12})^2 - (3n_{21} - n_{03})^2 \right] \\ + (n_{21} - n_{03})(3n_{11} + n_{03}) \left[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2 \right]$$

$$\varphi_6 = (n_{20} - n_{02})^2 \left[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2 \right] + 4n_{11}(n_{30} + n_{12})(n_{21} - n_{03})$$

$$\varphi_7 = (3n_{21} - n_{03})(n_{30} + n_{12}) \left[(n_{30} - 3n_{12})^2 - 3(n_{21} - n_{03})^2 \right] \\ + 3(n_{12} - n_{30})(n_{21} + n_{03}) \left[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2 \right]$$

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

La figure suivante montre une image de composant connexe (5) déformés par rotation, translation et par changement d'échelle ainsi les moments φ_i .

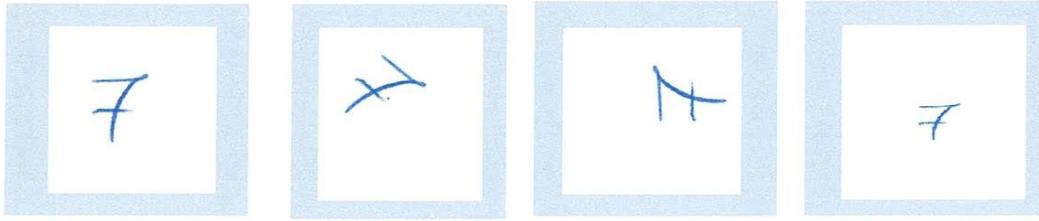


Figure. II.17. Opération géométrique

- a) Image originale.
- b) Image tournée de 45°.
- c) Image tournée de 90°.
- d) Image dont la taille est réduite par 50%.

Moments φ_i , $i=1, \dots, 7$, calculés sur les quatre images décrites au tableau (II. 1).

	Figure .a	Figure .b	Figure .c	Figure .d
φ_1	0.9061	0.8828	0.9061	0.8021
φ_2	0.3480	0.3200	0.3480	0.1454
φ_3	0.2132	0.2054	0.2132	0.1758
φ_4	0.0046	0.0043	0.0046	0.0007
φ_5	0.0001	0.0001	0.0001	0.0000
φ_6	0.0001	0.0001	0.0001	0.0000
φ_7	0.0004	0.0005	0.0004	0.0003
	-0.0001	-0.0001	-0.0001	-0.0000

Tableau .II. 1. Résultat de calcul des moments invariants.

En observant les moments décrits dans le Tableau (II. 1), on constate une assez bonne propriété d'invariance.



- **Description de formes par descripteur de Fourier**

L'analyse harmonique est l'instrument majeur de la théorie des signaux. La transformation de Fourier, généralisée par l'emploi des distributions, permet d'obtenir une représentation spectrale des signaux déterministes. Celle-ci exprime la répartition fréquentielle de l'amplitude, de la phase, de l'énergie ou de la puissance des signaux considérés.

L'analyse de Fourier consiste à décomposer un signal en une somme de signaux élémentaires particuliers, qui ont la propriété d'être faciles à mettre en œuvre et à observer. Le descripteur de Fourier correspond à une suite de nombres dont le but est de caractériser au mieux la forme à décrire. Il existe principalement deux méthodes (angulaire et complexe) qui permettent de les définir.

- **Aspects théoriques de la TF**

- ✚ **TF d'un signal réel**

Considérons un signal monodimensionnel de type temporel. La TF de ce signal est définie de la façon suivante :

$$X(f) = \text{TF}[x(t)] = \int_{-\infty}^{+\infty} x(t) \exp(-i2\pi ft) dt \quad \text{II.8}$$

Les signaux traités dans cette étude sont codés sur leurs parties réelles seulement, leurs parties imaginaires étant fixées à zéro.

$X(f) = \text{TF}[x(t)]$, avec $x(t)$ réel

$$\begin{aligned} &= \int_{-\infty}^{+\infty} x(t) \cos(2\pi ft) dt - i \int_{-\infty}^{+\infty} x(t) \sin(2\pi ft) dt \\ &= \text{Re}\{X(f)\} + i \text{Im}\{X(f)\} \end{aligned}$$

$$= |X(f)| \exp(i V(f)). \quad \text{II.9}$$

Le module, $|X(f)| = \sqrt{(\text{Re}\{X(f)\})^2 + (\text{Im}\{X(f)\})^2}$ appelé spectre d'amplitude, est une fonction paire. Le spectre de phase $V_X(f) = \arg(X(f))$, qui est une fonction impaire, sensible aux translations, ne sera pas pris en considération et nous nous limiterons donc à l'étude de $|X(f)|$ pour les seules fréquences positives.



✚ TF discrète (TFD)

Soit T_e la cadence à laquelle on prélève N échantillons pendant la durée T d'un signal x ; en introduisant les simplifications de notation suivantes :

$$x(k) = x(kT_e/T) ;$$

$$X(n) = X(n/T * f_e) ;$$

$$W_N = \exp(2ip/N) \text{ avec } N \text{ pair}$$

La Transformation de Fourier Discrète (TFD) est définie par :

$$X(n) = \sum_{k=N/2}^{N/2-1} x(k) W_N^{-nk} \quad \text{II.10}$$

Et la transformation inverse par :

$$x(k) = \sum_{n=N/2}^{N/2-1} X(n) W_N^{nk} \quad \text{II.11}$$

Où x est le vecteur du signal discret monodimensionnel et X sa TFD.

✚ Définition complexe de descripteurs de Fourier

Après une opération de *suivi du contour* extérieur d'un objet, on obtient la suite ordonnée de ses N points représentée sous la forme complexe **figure II.18**.

Un point s 'y déplaçant génère un signal monodimensionnel complexe $z(n)$ où n représente l'abscisse curviligne du contour.

$$z(n) = x(n) + jy(n) \quad \text{II.12}$$

x : représente la ligne,

y : représente le colonne.

Pour n variant de 0 à $N-1$.

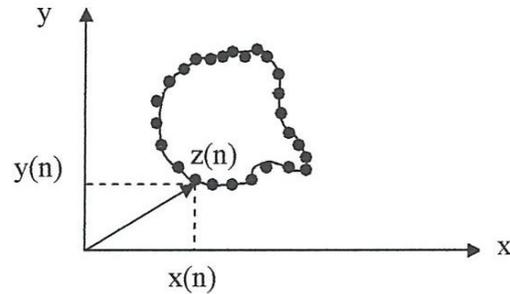


Figure II.18 : Représentation complexe d'un contour

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

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

Compte tenu de la nature périodique (de période N) de cette suite, on peut la représenter en utilisant la transformation de Fourier discrète (DFT) de l'équation II.12.

$$z(n) = (1/N) \sum_{k=0}^{N-1} Z(k) e^{2j\pi kn/N}, \text{ pour } 0 \leq n \leq N-1. \quad \text{II.13}$$

$$Z(k) = \sum_{n=0}^{N-1} z(n) e^{-2j\pi kn/N}, \text{ pour } 0 \leq k \leq N-1. \quad \text{II.14}$$

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) équation (II.14) aux coefficients.

On peut également effectuer une opération de "**filtrage**", par exemple en supprimant certains coefficients. Après transformée inverse, on obtient un contour fermé qui approxime plus ou moins bien le contour initial *Figure II.19*.

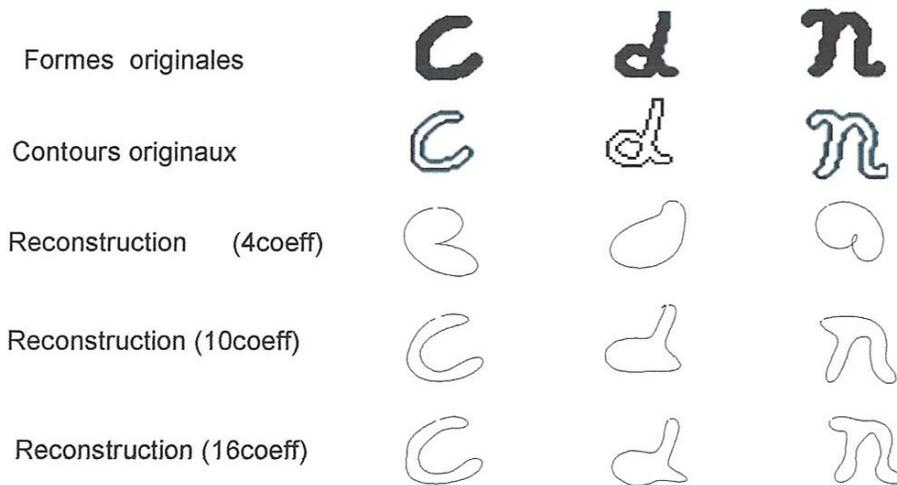


Figure II.19 : Reconstruction de la forme originale du contour extérieur

II.5.4 Classification

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). 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*.

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*.

Dans le cas de l'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.

Dans le cas de l'apprentissage non supervisé, on fournit au système de reconnaissance un grand nombre de formes non étiquetées. L'étape de la classification va se charger d'identifier automatiquement les formes appartenant à une même classe.

Dans le domaine de la reconnaissance de l'écriture, ce sont les méthodes basées sur un apprentissage supervisé qui sont le plus souvent utilisées ; et plus particulièrement pour les caractères manuscrits isolés car les classes naturelles sont connues et en nombre limité.

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. En effet, il est difficile de trouver *a priori* les primitives pertinentes et la méthode d'apprentissage la plus adaptée au problème posé d'où l'utilité de procéder par itérations successives. Ces itérations consistent à extraire des primitives jugées utiles au problème de reconnaissance à résoudre et de tester la performance du système avec cet ensemble de primitives. Au fur et à mesure que les performances du système souhaitées ne sont pas atteintes alors il suffit de trouver à nouveau une nouvelle famille de primitives ou de combiner les primitives extraites avec de nouvelles primitives.

Ces phases d'apprentissage et de test sont réalisées préalablement à l'intégration du module de décision dans le système de reconnaissance. Dans tous les cas, on peut permettre au système de reconnaissance d'itérer les phases d'apprentissage et de test tant que l'on n'a pas atteint les performances désirées. Le calcul de cette performance est le résultat du classifieur utilisé. Pour construire un classifieur, il existe trois approches : structurelle, statistique et hybrides.

L'approche réseau de neurones est une implémentation dérivée de l'approche statistique et de l'approche structurelle.

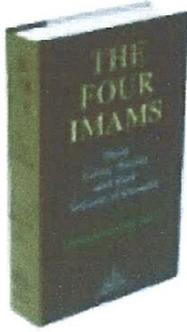
Dans notre cas, nous abordons les deux approches : structurelle et statistique en considérant que les réseaux de neurones font partie des approches statistiques.



Conclusion

Dans ce chapitre nous présentons les différentes étapes nécessaires à la construction d'un système de reconnaissance hors-ligne de l'écriture manuscrite. Pour chacune de ces étapes à savoir : les prétraitements, segmentation, l'extraction des primitives et la classification.

Notre effort s'est principalement dirigé vers l'extraction des meilleures primitives pertinentes et les différents modèles des classifieurs utilisant les réseaux de neurones.



CHAPITRE III

Réseaux de neurones

III 1. INTRODUCTION

III.2. MODELISATION DU NEURONE

III.2.1 Réseaux de neurones formels

III.2.1.1 Model simplifié (premier modèle de neurone)

III.2.1.2 Modélisation générale

III.2.2 Différentes structures des réseaux de neurones

III.2.2.1 Les Réseaux à couches

III.2.2.2 Les Réseaux entièrement connectés

III.2.3 Les réseaux les plus célèbres

III.2.3.1 Le perceptron

III.2.3.2 Le modèle ADALINE

III.2.3.3 Le perceptron multicouche (PMC)

III.2.4 L'apprentissage

III.2.4.1 Apprentissage supervisé

III.2.4.2 Apprentissage semi-supervisée

III.2.4.3 Apprentissage non-supervisée (auto-organisation)

III.2.5 Règles d'apprentissages

III.2.5.1 La règle de Hebb

III.2.5.2 La règle de Widrow-Hoff

III.2.5.3 Apprentissage du perceptron multicouches

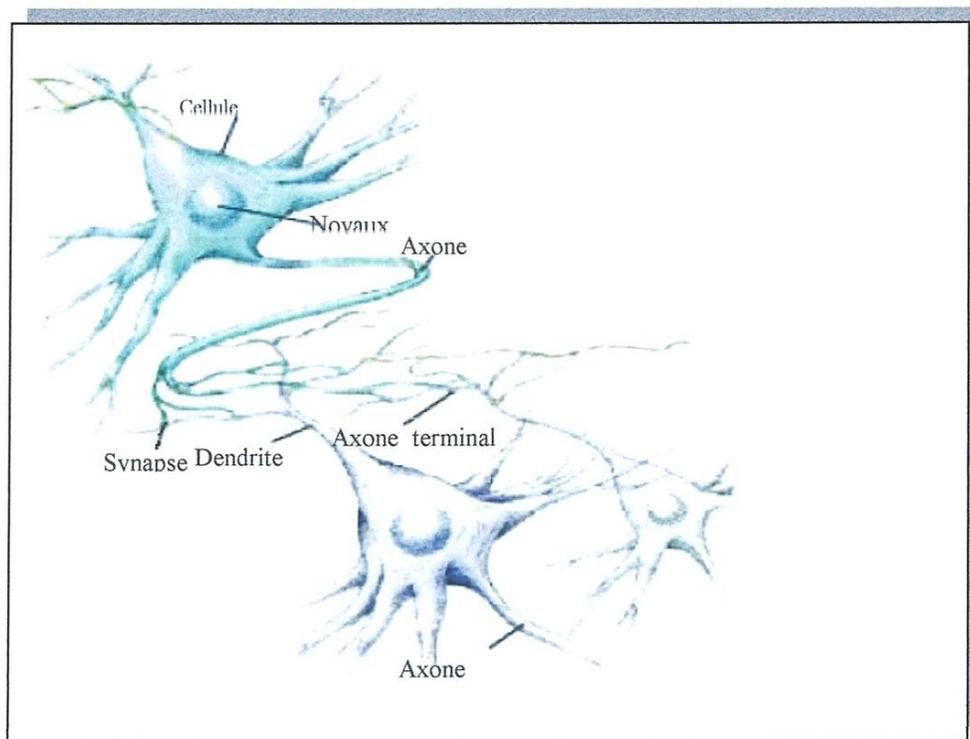
III.3 LE PERCEPTRON MULTICOUCHES COMME UN CLASSIFIEUR

III.4 EVOLUTION DU PERCEPTRON MULTICOUCHES

CONCLUSION

1. Introduction

Nous abordons les réseaux de neurones dans la seule perspective de faire de la reconnaissance des formes et de la classification. Notre objectif n'est donc pas celui des « neurosciences » consistant à chercher à utiliser la puissance des ordinateurs pour simuler l'intelligence humaine ou animale. Rappelons quelques généralités sur ces cellules particulières que sont les neurones biologiques. Il s'agit donc de cellules comportant un noyau comme toutes les cellules vivantes mais comportant des branches appelées axons. Ces axons peuvent être reliés pour « Connecter » des neurones.



**Figure III.1 : Vue d'artiste des connexions neuronales dans le cerveau.
Cellule émettrice en haut et cellules réceptrices en bas.**

Ces connexions se comptant par milliards et produisant ce qu'on appelle un réseau de neurones. Chacun de ces neurones peut être activé ou non par ce qu'on appelle la « force synaptique » et transmettre son « activation » par l'intermédiaire des synapses aux autres neurones. Chaque neurone pouvant être considéré comme comportant des entrées et des sorties. Ce réseau de neurones peut recevoir des activations par l'intermédiaire des sens notamment et l'état d'excitation des neurones définit alors un état de la perception.



III.2. Modélisation du neurone

III.2.1 Réseaux de neurones formels

Les réseaux de neurones formels sont à l'origine une tentative de modélisation mathématique du cerveau humain. Les premiers travaux datent de 1943 et sont l'œuvre de McCulloch et Pitts. Ils présentent un modèle assez simple pour les neurones et explorent les possibilités de ce modèle. L'idée principale des réseaux de neurones "modernes" est la suivante : On se donne une unité simple, un neurone, qui est capable de réaliser quelques calculs élémentaires. On relie ensuite entre elles un nombre important de ces unités et on essaye de déterminer la puissance de calcul du réseau ainsi obtenu. Il est important de noter que ces neurones manipulent des données numériques et non pas symboliques.

Citons quelques étapes dans la formalisation des réseaux de neurones :

- Première définition d'un neurone formel par McCulloch et Pitts en 1943
- Les percepts ou concepts sont physiquement représentés dans le cerveau par l'entrée en activité (simultanée) d'une assemblée de neurones (Hebb, 1949).
- Deux neurones entrant en activité simultanément vont être associés (c'est-à-dire que leurs contacts synaptiques vont être renforcés. On parle de loi de Hebb et d'associationnisme
- Le perceptron de Frank Rosenblatt (1958) : le premier modèle pour lequel un processus d'apprentissage a pu être défini. De cette période, date également les travaux de Widrow et Hoff.
- Le livre de Minski et Papert "Perceptrons" (1969). Cet ouvrage contient une étude critique très complète des perceptrons. On lui reproche parfois violemment d'avoir sonné le glas des recherches sur les réseaux neuronaux dans les années 70, ce que nient leurs auteurs. Ce livre a été réédité en 1980, avec des ajouts et corrections manuscrites dans les marges, sans doute pour qu'on ne puisse pas les accuser de camoufler la première version du texte.
- L'algorithme de rétro propagation du gradient dans les réseaux multicouches découvert au début des années 80 par Rumelhart et McClelland, Parker, Hinton, Le Cun. Les « inventeurs » sont nombreux car l'idée de descente de gradient est naturelle. La plupart de ces travaux étaient associés à des études empiriques montrant la puissance du modèle.

- Le modèle de Hopfield (1982) qui utilise des réseaux totalement connectés basés sur la règle de Hebb qui ont permis de définir la notion d'attracteurs et de mémoire associative.

III.2.1.1 Model simplifié (premier modèle de neurone)

Le modèle de neurone formel présenté ici, dû à McCulloch et Pitts, *Figure III.2* est un modèle mathématique très simple dérivé d'une analyse (elle aussi assez simple) de la réalité biologique. On ne constate tout d'abord que le modèle biologique fait intervenir une notion temporelle qui est difficile à intégrer dans un modèle simple. On oublie donc cette notion et de ce fait on remplace l'intégration temporelle par une simple sommation des signaux arrivant au neurone (ces signaux sont communément appelés les entrées du neurone). On compare ensuite la somme obtenue à un seuil et on déduit de la comparaison la sortie du neurone. Cette sortie sera par exemple *Figure III.3* égale à 1 si la somme est supérieure au seuil et à 0 dans le cas contraire. Plus formellement encore, il suffit pour obtenir ce comportement de soustraire le seuil considéré à la somme des entrées, et de faire passer le résultat par la fonction *de* transfert du neurone qui est ici la fonction de Heaviside *figure III.4*. Le résultat après transfert est alors la sortie du neurone.

Cet enchaînement "sommation" puis "non-linéarité" représente finalement les propriétés "physiques" du neurone.

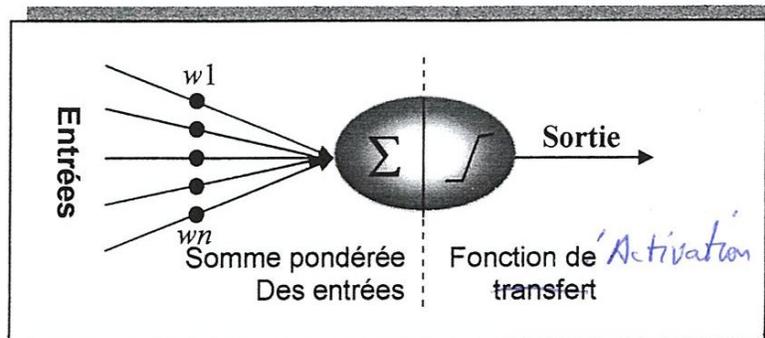


Figure III.2 : Le neurone formel de McCulloch et Pitts

La modélisation des synapses est assez simpliste en première approche. On se contente en effet d'introduire la notion de synapse excitatrice et de synapse inhibitrice en multipliant la sortie d'un neurone par 1 ou -1 avant de la transmettre aux autres neurones. Afin de donner plus de puissance au modèle, on généralise ce modèle en introduisant ensuite une connexion synaptique qui est une valeur réelle. La sortie d'un neurone est alors

multipliée par la valeur de la connexion synaptique avant d'être transmise à un autre neurone. On modélise ainsi la connexion entre deux neurones.

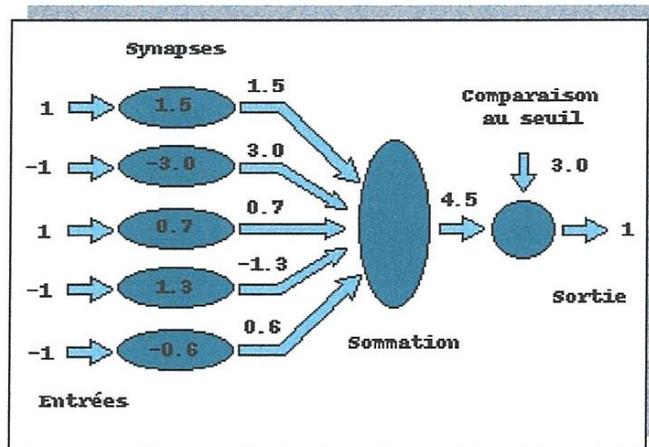


Figure III.3 : Exemple d'un neurone formel.

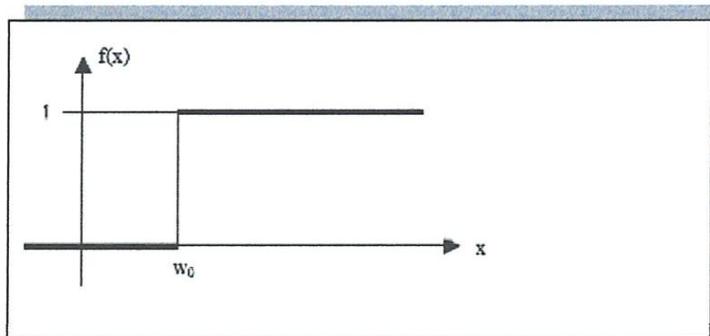


Figure III.4 : Fonction de Heaviside, utilisée par MacCulloch et Pitts.

En résumé, un neurone formel réalise simplement une somme pondérée de ses entrées, ajoute un seuil à cette somme et fait passer le résultat par une fonction de transfert pour obtenir sa sortie.

Le fonctionnement du neurone formel est donné par l'équation suivante :

$$S = \left(\sum_{i=1}^n w_i \times e_i \right) \quad \text{III.1}$$

Tel que :

e_i : $i=1 \dots n$: les entrées du neurone formel.

S : sa sortie.

w_i : les paramètres de pondération.

III.2.1.2 Modélisation générale

D'une façon plus générale, on peut définir un neurone formel par les quatre éléments suivants :

- La nature des entrées.
- La fonction de sortie qui calcule la sortie du neurone.
- La nature de la sortie du neurone.

Nous adapterons par la suite les notations suivantes :

- $(e_i)_{i=1 \dots n}$: seront les entrées.
- h sera la fonction d'entrée totale.
- F sera la fonction de sortie.
- D'autre part, nous utiliserons aussi :

$E=h(e_1 \dots e_n)$ comme entrée totale.

$S=F(E)$ comme sortie.

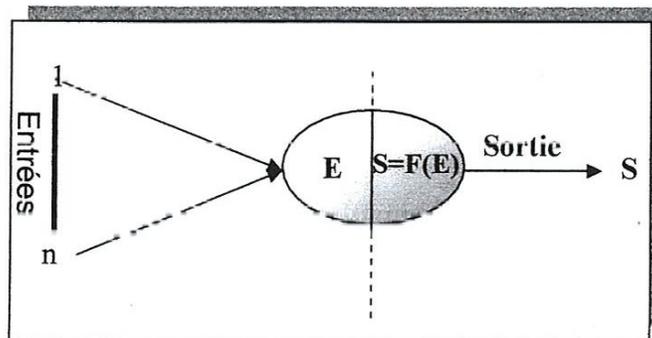


Figure III.5 : Modélisation d'un neurone.

La nature des entrées et de la sortie peuvent être :

- Binaire : $(-1, +1)$ ou $(0, 1)$.
- Réelles.

La fonction d'entrée totale h peut être :

- Booléenne
- Linéaire : $h(e_1, \dots, e_n) = \sum w_i \cdot e_i$

La fonction de sortie 'F' peut être :

- Une fonction binaire à seuil ; dans ce cas, on s'arrange pour que la forme de h soit telle qu'on puisse utiliser la fonction de Heaviside ou la fonction Signe.
- Une fonction as seuil où multi-seuils
- Une fonction non-linéaire (la fonction sigmoïde)
- Une fonction stochastique

III.2.2 Différentes structures des réseaux de neurones

Les structures qui peuvent être utilisées sont très variées. Si l'on se réfère aux études biologiques du cerveau, on constate que le nombre de connexions est énorme. Par exemple, des chercheurs ont montré que le cortex était divisé en différentes couches, à l'intérieur d'une même couche les interactions entre neurones sont très grandes mais les neurones d'une couche sont reliés aux neurones des autres couches, le tout formant un système gigantesque.

D'une manière générale, l'architecture des réseaux de neurones formels peut aller d'une connectivité totale à une connectivité locale où les neurones ne sont reliés qu'à leurs plus proches voisins. Il est courant d'utiliser des réseaux à structure régulière pour faciliter leur utilisation.

Les réseaux de neurones peuvent se répartir en deux grandes catégories, selon qu'ils interdisent ou qu'ils permettent les liens récurrents entre leurs neurones. Les premiers dites *réseaux à couches*, voient l'activation de leurs neurones les traverser en une vague unique n'ont donc pas véritablement de comportement dans le temps.

En revanche, les réseaux appartenant à la seconde catégorie, nommée *réseaux entièrement connectés*, sont caractérisés par des comportements dynamiques qui peuvent être plus ou moins-complexes selon les contraintes d'architecture qu'ils comportent.

III.2.2.1 Les Réseaux à couches

Il est possible, à l'étude du système visuel, de concevoir une organisation par couches successives du système nerveux, les signaux arrivant sur les cellules sensibles, sont transmis à un système de neurone hyper connectés, avec entre autres des mécanismes de bouclage puis aboutissent aux cellules motrices qui déclenchent un comportement.

Ce système portant simplifié est déjà très complexe à modéliser si l'on tient à conserver la structure de connectivité quasi-totale dans les couches intermédiaires des traitements du signal. C'est pourquoi on utilise plutôt une structure de réseaux à couches telle que les neurones appartenant à une même couches ne soient pas connectées entre eux, *figure III.6*, chacune des couches recevant des signaux de la couche précédente, transmet le résultat des ces traitements à la couche suivante.

Les deux couches extrêmes correspondent à la couche qui reçoit ses entrées du milieu extérieur à la couche qui fournit le résultat des traitements effectués. Les couches intermédiaires sont appelées couches cachées, leur nombre est variable.

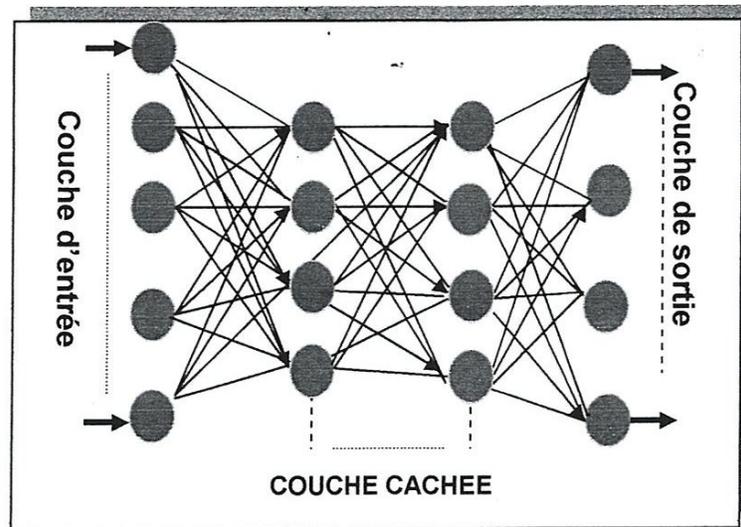


Figure III.6 : Réseaux multicouches.

III.2.2.2 Les Réseaux entièrement connectés

Dans ces réseaux, chaque cellule est reliée à toutes les autres et possède un retour sur elle-même, il n'y a en général qu'une seule couche vue et au mieux deux; une d'entrée et une de sortie. Un neurone reçoit des impulsions de tous ses compagnons et même de lui-même.

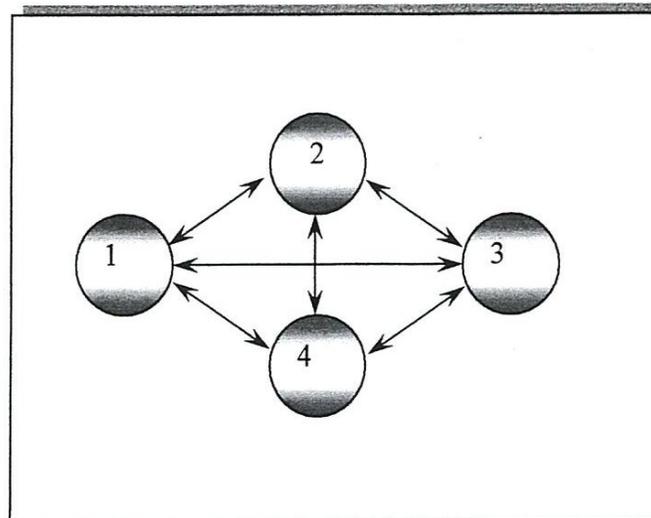


Figure III.7. Réseaux entièrement connectés

III.2.3 Les réseaux les plus célèbres

Il y a de très nombreuses sortes de réseaux de neurones actuellement. Personne ne sait exactement combien. De nouveaux réseaux (ou du moins des variations de réseaux plus anciens) sont inventés chaque semaine. On en présente ici de très classiques.

III.2.3.1 Le perceptron

C'est un des premiers réseaux de neurones, conçu en 1958 par **Roseblatt**. Il est linéaire et monocouche. Il est inspiré du système visuel.

Essentiellement, le perceptron est composé de deux couches *figure III.8*. La première (d'entrée) appelée à l'origine la rétine du perceptron, quant à la deuxième couche, elle donne la réponse du perceptron correspondant à la stimulation présentée en entrée.

Les sorties des neurones ne peuvent prendre que deux états (-1 et 1 ou 0 et 1) seuls les poids des liaisons entre la couche d'association et la couche finale peuvent être modifiés.

III.2.3.2 Le modèle ADALINE

L'ADALINE (Adaptatif Linéaire Élément) conçu par B.Widrow dans les années 1960, est un perceptron sans couche cachée donc à un seul neurone *figure III.9*, qui reçoit le stimulus arrivant de la couche d'entrée et donne la réponse correspondante.

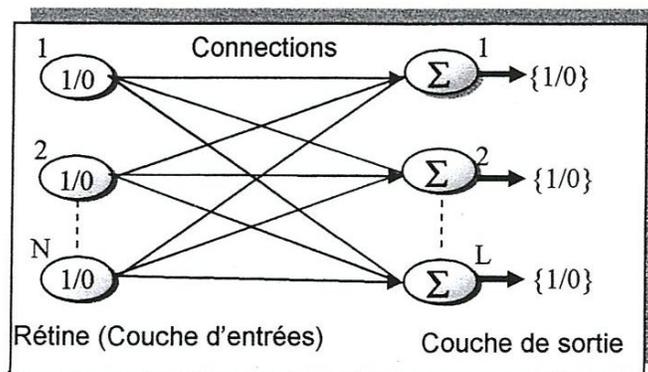


Figure III.8 : Structure d'un perceptron.

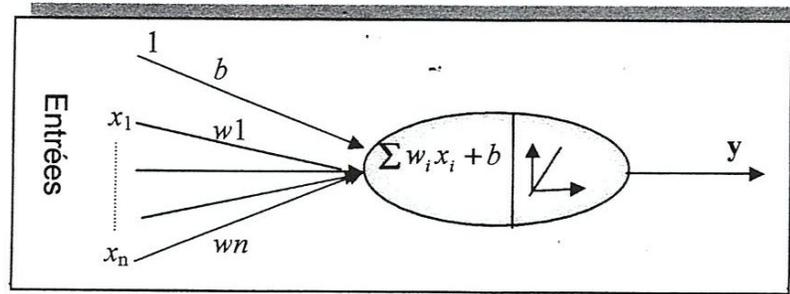


Figure III.9 : Architecture de L'ADALINE

III.2.3.3 Le perceptron multicouche (PMC)

C'est une amélioration du perceptron comprenant une ou plusieurs couches intermédiaires dites couche cachée *figure III.6*. Ils utilisent, pour modifier leur poids, un algorithme de rétropropagation du gradient, qui est une généralisation de la règle de WIDROW-HOFF. Les PMC agissent comme un séparateur non linéaire et peuvent être utilisés pour la classification, le traitement d'image, l'aide de la décision ou la commande d'un processus.

Vu l'importance de ce modèle dans notre étude, il fera l'objet d'une étude détaillée dans ce qui suit.

III.2.4 L'apprentissage

Le principal problème pour les réseaux de neurones est d'arriver à trouver un ensemble de valeurs des synapses (poids), qui sont les porteurs de l'information, tel que les configurations d'entrée se traduisent par les réponses voulues, et cela en partant d'une valeur particulière des poids des connexions, les réseaux améliorent leurs réponses en ajustant leurs coefficients selon un algorithme ou une règle d'apprentissage.

Il existe trois classes d'apprentissage : l'apprentissage supervisé, semi supervisé et non-supervisé.

III.2.4.1 Apprentissage supervisé

Les apprentissages supervisés demandent que l'on donne au réseau neuronal des exemples résolus, c'est à dire des couples de vecteurs entrée/sortie. Cet apprentissage exploite le plus souvent plusieurs idées simples. L'idée principale est la minimisation itérative d'un critère de l'erreur en sortie du réseau. On initialise les matrices de connexion

au hasard, puis l'on fait évoluer ces matrices de manière à ce qu'elles autorisent l'association souhaitée, c'est à dire jusqu'à ce qu'un critère de l'erreur (entre les sorties réellement obtenues et les sorties souhaitées) soit quasi nul.

III.2.4.2 Apprentissage semi-supervisé

Dans ce type d'apprentissage l'utilisateur possède seulement des indications imprécises sur le comportement final du réseau, mais en revanche, il est possible d'obtenir des indications qualitatives (correctes/ incorrectes).

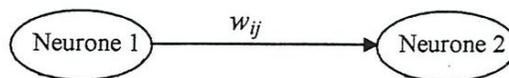
III.2.4.3 Apprentissage non-supervisée (auto-organisation)

Dans ce type d'apprentissage (sans professeur) les poids synaptiques du réseau sont modifiés selon des critères internes comme la co-activation des neurones. Le comportement de ce type d'apprentissage est comparable à des techniques d'analyse de données. Un exemple des réseaux à apprentissage non-supervisée : les cartes topologiques de Kohonen. Enfin, certains réseaux associent les deux types d'apprentissage supervisé et non-supervisée, c'est le cas par exemple du réseau de Boltzmann.

III.2.5 Règles d'apprentissages

III.2.5.1 La règle de Hebb

Dans le domaine de la recherche sur le fonctionnement des neurones biologiques et sur les mécanismes d'apprentissage de l'intelligence humaine, Hebb a proposé un type de réseau de neurones totalement interconnecté (c'est à dire ou les neurones sont reliés par des connexions de type synapse fonctionnant à la fois en « entrée » et en « sortie ». Ces connexions sont affectées de poids qui évoluent au cours du temps et en fonction de l'activation de chacun des deux neurones extrémités de cette connexion. Il a ainsi défini une règle d'apprentissage dite de Hebb :



Cette règle considère alors que toute connexion entre deux neurones se renforce si ces deux neurones sont actifs au même moment. Si on note A_1 et A_2 l'activation des neurones 1 et 2 et si on suppose qu'un neurone actif à son activation qui vaut 1 et qu'un neurone inactif a son activation qui vaut 0 alors l'expression de la règle de Hebb est la suivante :

$$w_{ij}(t + \delta t) = w_{ij}(t) + \mu \times A_i \times A_j \text{ Avec } \mu > 0 \quad \text{III.2}$$

Au départ on a $w_{ij}(0) = 0$ pour tout i, j .

Ce type de réseaux a surtout une vocation pour la modélisation des neurones biologiques et reste peu utilisé en reconnaissance des formes. Un des inconvénients majeurs de ce modèle vient du fait que les w_{ij} ne peuvent qu'augmenter au cours du temps.

III.2.5.2 La règle de Widrow-Hoff

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

$$w(k+1) = w(k) + \eta(t_k - y_k)x_k^T \quad \text{III.3}$$

A chaque pas d'apprentissage k , l'erreur en sortie est calculée comme la différence entre la cible recherchée t et la sortie y des réseaux

$$E_k = e_k^T e_k = (t_k - y_k)^T (t - y) = \frac{1}{2}(t_k^T t_k + y_k^T y_k - 2y_k^T t_k) \quad \text{III.4}$$

Le gradient de se calcule comme suite :

$$\nabla E_{kw} = \frac{1}{2} \nabla [y_k^T y_k - 2y_k^T t_k]_w \quad \text{III.5}$$

le gradient se calcule comme suite :

$$\nabla E_{kw} = \frac{\partial E_k}{\partial w} = \frac{\partial E_k}{\partial y_k} \frac{\partial y_k}{\partial w} \quad \text{III.6}$$

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

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

$$\frac{\partial (wx_k + b)}{\partial w} = x_k$$

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

$$w(k+1) = w(k) + \eta \nabla E_{kw} \quad \text{III.7}$$

Avec η : le gain d'apprentissage, $0 < \eta < 1$

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

$$b(k+1) = b(k) + \eta(t_k - y_k) = b(k) + \eta \nabla E_{kb} \quad \text{III.8}$$

III.2.5.3 Apprentissage du perceptron multicouches

L'apprentissage du perceptron multicouches est supervisé et consiste à adapter les poids des neurones de manière à ce que le réseau soit capable de réaliser une transformation données, représentée par un ensemble d'exemples constitué d'une suite de N vecteurs d'entrées.

Donc l'apprentissage fonctionne sur le même principe que la règle de Widrow-hoff ; on dispose d'un ensemble d'exemples qui sont des couples (entrées-sorties désirées). A chaque étape un exemple est présenté en entrée du réseau. Une sortie réelle est calculée. Ce calcul est effectué de proche en proche de la couche d'entrée à la couche de sortie. Cette phase est appelée propagation avant ou relaxation du réseau. Ensuite l'erreur (somme quadratique des erreurs sur chaque sortie) est calculée. Cette erreur est ensuite rétro propagée dans le réseau, donnant lieu à une modification de chaque poids.

L'apprentissage du perceptron multi-couches consiste à minimiser l'erreur quadratique commise sur l'ensemble des exemples, Ce problème de minimisation de l'erreur a été résolu par des méthodes de *rétropropagation du gradient d'erreur*.

III.3 Le Perceptron multicouches comme un classifieur

L'introduction de couches intermédiaires dans le réseau MLP permet de résoudre des problèmes plus complexes que la simple séparation linéaire. Lorsqu'il existe au moins une couche cachée, les états internes du réseau ne peuvent plus être donnés directement par les exemples et les sorties désirées puisque les sorties des neurones appartenant aux couches intermédiaires sont inconnues. La *figure III.10* représente un réseau de neurones multicouches avec comme entrée le vecteur de primitives et en sortie les classes où seront classées les formes. Pour déterminer le nombre de couches cachées dans un réseau cela dépendra du problème à résoudre (comment les classes des formes sont-elles séparées?).

Plusieurs types de réseaux de neurones multicouches ont été développés. Nous focalisons notre étude sur le type MLP. Dans le processus de la construction de ce type de classifieur, l'apprentissage a comme objectif d'adapter les poids des connexions entre les neurones de sorte que le réseau donne en sortie la classe d'appartenance des formes qui lui sont proposées en entrée. Ce qui revient à minimiser l'erreur commise par le réseau sur l'ensemble de formes de la base d'apprentissage. Pour une forme de la classe C_i , l'état du neurone O_i de la couche

De sortie du MLP doit être proche de la valeur maximale et tous les autres proches de la valeur minimale. Ce problème de minimisation de l'erreur a été résolu par des méthodes de *rétropropagation du gradient d'erreur*. Toutefois, ces méthodes ne garantissent pas l'obtention du minimum global de la fonction à optimiser (blocage dans un minimum local). Le terme «*rétropropagation*» est utilisé pour décrire l'apprentissage du réseau de neurones de type MLP utilisant la descente du gradient appliquée à la fonction de la somme des erreurs quadratiques. Le réseau de type MLP à une couche cachée peut estimer n'importe quelle fonction dans R^n avec une précision arbitraire. Ainsi, le MLP est capable d'estimer des hyper-surfaces discriminantes très complexes. Les fonctions estimées sont soit linéaires ou non linéaires. Elles sont appelées fonctions d'activation. La fonction non linéaire que nous utiliserons est la fonction *sigmoïde*.

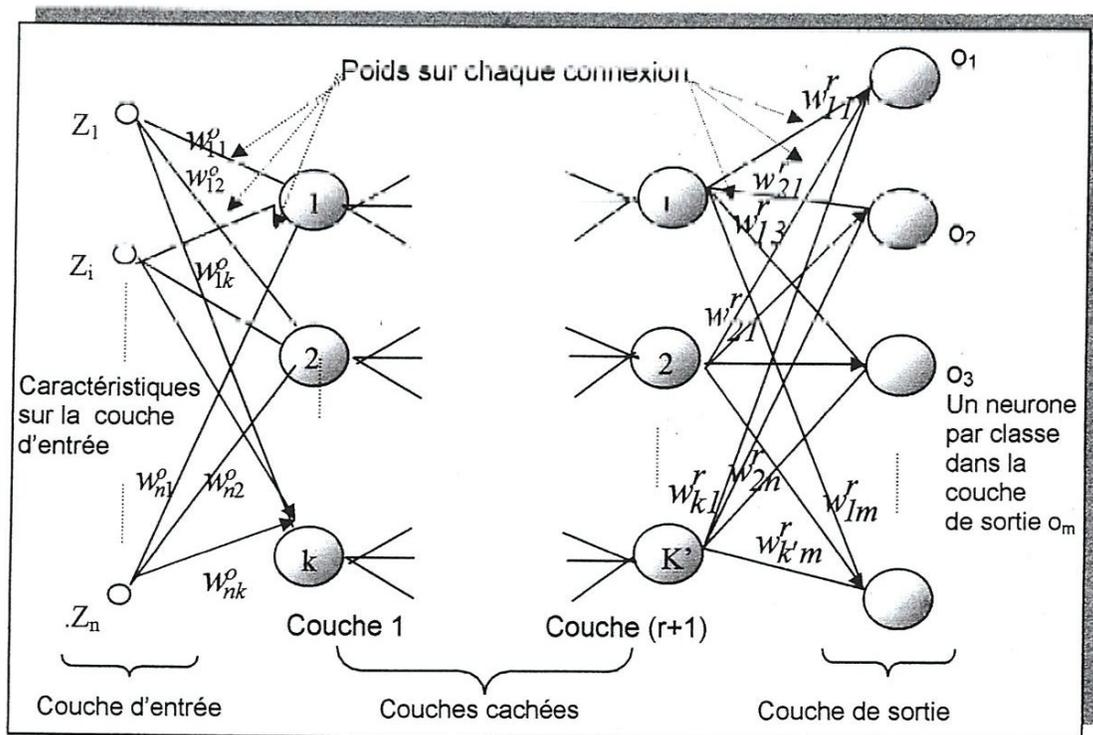


Figure III.10 : Réseau de neurones multicouches.

Par contre la difficulté majeure rencontrée lors de l'utilisation de ce type de classifieur consiste à déterminer le nombre de couches cachées, le nombre de neurones dans chacune des couches et les poids des connexions entre les différentes couches (taux d'apprentissage...). De ce fait, pour une application donnée, la construction du classifieur

de type MLP utilise des règles empiriques et nécessite un certain nombre d'essais afin d'obtenir des performances de généralisation intéressantes.

Pour déterminer les poids de toutes les connexions du réseau, l'utilisation des algorithmes d'apprentissage est obligatoire. L'objectif des algorithmes d'apprentissage est de minimiser l'erreur de décision effectuée par le RNA en ajustant les poids à chaque présentation d'un vecteur d'entraînement. Pour ce qui est de l'ajustement des poids à une étape donnée de la phase d'apprentissage, l'erreur à minimiser est habituellement celle produite lors de l'application d'un vecteur de l'ensemble d'apprentissage à l'entrée du réseau. Pour ce qui est de l'évaluation de la qualité d'apprentissage du réseau, l'erreur cumulée par tous les vecteurs de l'ensemble d'entraînement est évaluée. Cette erreur cumulée est calculée pour tous les cycles de la phase d'entraînement et est définie à partir de l'erreur quadratique. Cette mesure de l'erreur illustre la précision obtenue après P cycles d'apprentissage. En résumé, nous utilisons pour l'apprentissage du réseau l'algorithme de rétro-propagation avec minimisation du gradient d'erreur qui est défini par les étapes suivantes :

1. Initialiser les poids à des petites valeurs et les seuils du réseau.
2. Insérer à l'entrée du réseau une observation (exemple) de la base de données en forme de vecteur de caractéristiques, puis calculer sa valeur d'activation et sa fonction d'activation en utilisant les formules (III.9) et (III.10).
3. Évaluer le signal d'erreur des sorties du réseau en utilisant la formule (III.11).
4. Ajuster les poids en utilisant la formule (III.9).
5. Évaluer le signal d'erreur pour chaque couche cachée en utilisant la formule (III.12).
6. Ajuster les poids de la couche cachée en utilisant la formule (III.10).
7. Répéter les étapes 2 à 6 pour l'ensemble des observations de la base d'apprentissage tant que le critère d'arrêt n'a pas été atteint.

Il existe plusieurs critères d'arrêts. Ces critères peuvent être combinés entre eux.

Le premier critère est basé sur l'amplitude du gradient de la fonction d'activation, puisque par définition le gradient sera à zéro au minimum. L'apprentissage du réseau du type MLP utilise la technique de recherche du gradient pour déterminer les poids du réseau.

Le second critère d'arrêt est de fixer un seuil que l'erreur quadratique ne doit pas dépasser. Toutefois ceci exige une connaissance préalable de la valeur minimale de l'erreur qui n'est pas toujours disponible. Dans le domaine de la reconnaissance de formes, il suffit de s'arrêter lorsque tous les objets sont correctement classés

Les deux premiers critères sont sensibles aux choix des paramètres (par exemple : le nombre de noeuds dans la couche cachée, le seuil d'erreur, ...) et si le choix n'est pas bon alors les résultats obtenus seront mauvais ou le temps de calcul de la performance du système de reconnaissance sera plus lent (par exemple définir un grand nombre de noeuds dans la couche cachée). Cependant, la méthode du crossvalidation n'a pas ce genre de problème.

En général, les formules utilisées par cet algorithme sont :

- pour l'ajustement de poids entre le noeud j (sortie ou cachée) et le noeud i (cachée ou entrée)

$$\Delta w_{ij} = \eta \delta_j o_i$$

III 9

Où η est la valeur de la constante d'apprentissage.

En général, $0.1 < \eta < 0.9$

et o_i est la valeur d'activation du neurone i tel que

$$o_i = f(\text{net}_i) \quad \text{net}_i = \sum_j w_{ij} o_j \quad \text{III.10}$$

Avec $f(\text{net}_i)$ est la fonction d'activation. La fonction utilisée dans le cas de notre projet est la fonction sigmoïde définie

$$f(\text{net}_i) = \frac{1}{1 + \exp^{-\text{net}_i}} \quad \text{III.11}$$

- pour le calcul du signal d'erreur du neurone j de la couche de sortie avec d_j est la valeur désirée du neurone j .

$$\delta_j = (d_j - o_j) o_j (1 - o_j) \quad \text{III.12}$$

Avec d_j est la valeur désirée du neurone j .

- pour le calcul du signal d'erreur du neurone j de la couche cachée avec k nœuds

$$\delta_j = o_j (1 - o_j) \sum_k w_{jk} \delta_k \quad \text{III.13}$$

Ces formules ont été dérivées de la formule de calcul de l'erreur quadratique de l'ensemble de la base d'apprentissage définie comme suit

$$E_p = \frac{1}{2} \sum_p \left(\sum_k (d_{pk} - o_{pk})^2 \right) \quad \text{III.14}$$

Où p est l'indice d'un exemple de la base et k est l'indice du noeud de sortie.

L'objectif est de minimiser cette erreur.

III.4 Evolution du perceptron multicouches

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

- La mise en œuvre de cet algorithme exige souvent des temps de calcul très longs qui rendent son application incommode pour de nombreux problèmes de taille raisonnable.
- La rétro-propagation peut trouver piège dans les minima locaux, et ne peut pas donner la bonne réponse. Pour remédier à ce problème, on peut entraîner le réseau à partir de plusieurs choix initiaux de poids pour ne garder que le meilleur ou bien ajouter un bruit aléatoire puis relancer l'apprentissage.
- Le réseau apprend à partir des exemples de la base d'apprentissage. Si cette dernière concerne un fonctionnement dans un domaine réduit, le réseau ne saura répondre en dehors de ce domaine. Si le superviseur donne des informations erronées, le réseau apprendra un modèle erroné et si le processus que l'on cherche à modéliser est non stationnaire ou change brutalement, le réseau réapprendra un bon modèle.

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

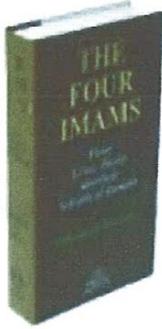


Conclusion

Nous avons essayé à travers ce chapitre de donner aperçu général sur les réseaux de neurones artificiels et les différentes règles d'apprentissage.

Les réseaux de neurones les plus connus présentés dans ce chapitre sont des systèmes paramétriques non linéaires qui feront l'objet des chapitres suivants avec son application au problème de reconnaissance. Les paramètres du réseau sont déterminés statistiquement lors d'une phase d'apprentissage à partir d'une base de données d'exemples, appelée base d'apprentissage.

En pratique, l'utilisation de la méthode neuronale pose certaines difficultés. La principale difficulté 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.



CHAPITRE IV

Application Sur Les Chiffres Manuscrits

IV.1. INTRODUCTION

IV.2. APPLICATION A LA RECONNAISSANCE DES CHIFFRES MANUSCRITS

IV.2.1 Base de données

IV.2.2 prétraitement

IV.2.2.1 Le choix des primitives

IV.2.2.2 les paramètres de distribution (PD)

IV.2.2.3. Les sept moments invariants (MI)

IV.3. CONSTITUTION DU VECTEUR DES CARACTERISTIQUES

IV.4. CLASSIFICATION (APPRENTISSAGE ET RECONNAISSANCE)

IV.4.1. Apprentissage

IV.4.2. Reconnaissance

CONCLUSION



1. Introduction

Dans ce travail, nous nous intéressons à l'élaboration d'un système de reconnaissance de l'écriture à vocabulaire limité car l'application visée derrière cette architecture est la classification des chiffres manuscrits. Le problème de la séparation des chiffres connectés ne sera pas étudié. Dans le cadre de notre étude, nous traiterons donc des chiffres préalablement isolés.

Dans ce chapitre, nous allons décrire brièvement dans la section 2 l'opération d'acquisition et les opérations de prétraitements. Dans la section 3, nous présenterons la procédure adoptée pour l'extraction des primitives qui sera plus au moins détaillée. Dans la section 4, nous abordons le classificateur à base de réseaux de neurones et nous concluons notre article en présentant les résultats obtenus dans notre travail.

IV.2. Application à la reconnaissance des chiffres manuscrits

IV.2.1 Base de données

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 dans notre étude contient des chiffres manuscrits isolés et stockée sous forme d'images binaires. Le nombre des classes est égale à dix (0, 1, 2, 3..., 9) (fig. IV.1)

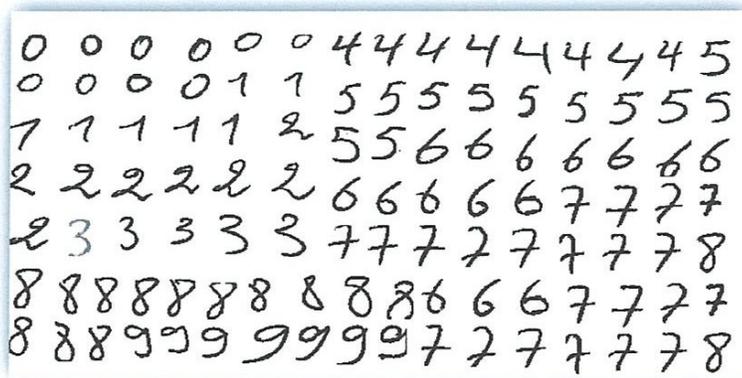


Figure IV.1 Aperçu sur la base des données.

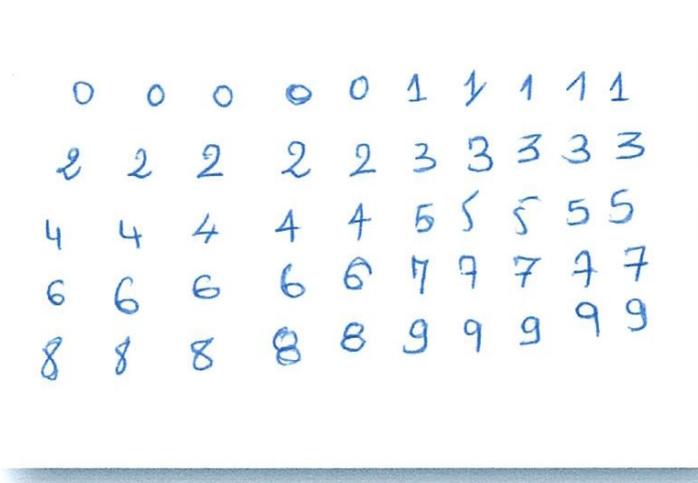


Figure IV. 2 Aperçu sur la base de test.

IV.2.2 prétraitement

Les prétraitements effectués sur cette base de données ont pour objectif le nettoyage des images. Autrement dit le bruit a été supprimé par les différentes techniques de prétraitement (binarisation, filtrage... etc.)

IV.2.2.1 Le choix des primitives

Nous avons effectué plusieurs simulations pour déterminer les primitives qui semblent être appropriées au système de reconnaissance de chiffre isolés. Les primitives extraites sont :

- ◆ Paramètre de distribution (Zonage).
- ◆ Les sept moments invariants
- ◆ Constitution du vecteur des caractéristiques
- ◆ Classification (apprentissage et reconnaissance)

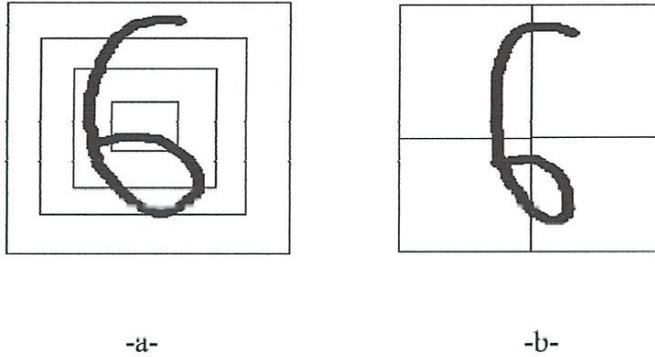


IV.2.2.2 les paramètres de distribution (PD)

Ces paramètres caractérisent la distribution des pixels, dans des zones obtenues en divisant l'image en 4 zones de deux manières différentes. Les paramètres de distribution qui sont désignés aussi par caractéristiques zonales ou locale sont définis par : $R_j = N_j/N$

N_j : nombre de pixels objets dans la région j.

N : nombre total des pixels dans le caractère.



0	0.0271	0.2990	0.6739	0.2129	0.2286	0.3830	0.1755
---	--------	--------	--------	--------	--------	--------	--------

-c-

Figure. IV.3 Extraction des paramètres de distribution normalisée

- a- Image de caractère est divisée en 4 zones
- b- Image de caractère est divisée en 4 zones
- c- Vecteur caractéristique associe à l'image a&b.

IV.2.2.3. Les sept moments invariants (MI)

Les moments géométriques permettent de trouver les caractéristiques propres aux formes à reconnaître telles que les axes principaux d'inertie, le centre de gravité, l'orientation de ses formes ainsi que d'autre l'exploitation de ces informations pourrait permettre de distinguer les différentes formes entre elles.

L'algorithme utilisé pour le calcul des sept moments invariants est :



POUR chaque caractère FAIRE :

- Lecture du modèle de caractère à partir du fichier.
- Calcul des moments m_{pq} (équation)
- Calcul des moments centrées μ_{pq} (équation)
- Calcul des moments invariants φ_i

FIN POUR

Avec $I=1 \rightarrow 7$.

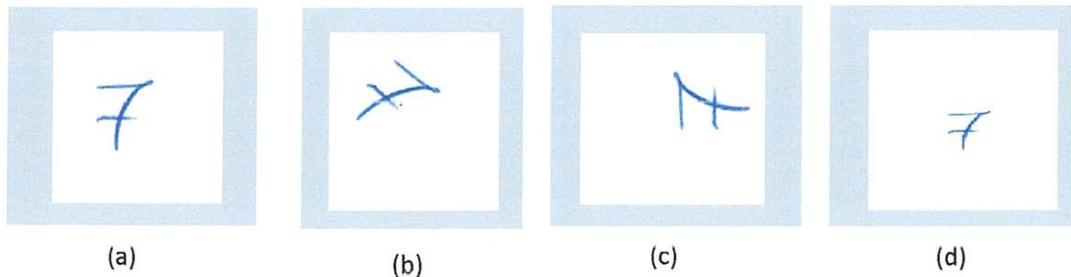


Figure IV.4. Extraction des moments invariants le chiffre « 7 ».

- a) Image originale.
- b) Image tournée de 45° .
- c) Image tournée de 90° .
- d) Image dont la taille est réduite par 50%.

Moments φ_i , $i=1, \dots, 7$, calculés sur les quatre images décrites au tableau (IV. 1).

	Figure .a	Figure .b	Figure .c	Figure .d
φ_1	0.9061	0.8828	0.9061	0.8021
φ_2	0.3480	0.3200	0.3480	0.1454
φ_3	0.2132	0.2054	0.2132	0.1758
φ_4	0.0046	0.0043	0.0046	0.0007
φ_5	0.0001	0.0001	0.0001	0.0000
φ_6	0.0004	0.0005	0.0004	0.0003
φ_7	-0.0001	-0.0001	-0.0001	-0.0000

Tableau. IV. 1. Résultat de calcul des moments invariants.



En observant les moments décrits dans le Tableau (IV. 1), on constate une assez bonne propriété d'invariance.

IV.3. Constitution du vecteur des caractéristiques

Notre application consiste à classifier un ensemble des chiffres manuscrits arabe de 0 à 9 dans le but de les reconnaître. Les paramètres extraits de chaque image de chiffre sont représentés par l'ensemble des primitives que nous avons développée précédemment (25 caractéristiques) réparties de la manière suivante :

- 07 primitives de type moments invariants appliquées à l'image binaire.
- 18 primitives de type de partition de distribution.

Donc la base d'apprentissage représente 10 chiffres (classes) différents et pour chaque chiffre nous avons pris 10 images de prises de vues différentes écrit sous contrainte par dix scripteurs. Donc, pour l'apprentissage nous avons utilisé une base constituée de 100 images.

IV.4. Classification (apprentissage et reconnaissance)

IV.4.1. Apprentissage

Le classificateur que nous avons utilisé est un perceptron multicouche à rétropropagation du gradient de l'erreur à une seule couche cachée. Pour l'apprentissage de réseaux on a fixée 25 neurones en entrée, La couche cachée est composée de 15 neurones. Les classes à discriminer sont 10 classes qui représentent tous les chiffres de 0 à 9, d'où le choix de 10 neurones pour la couche de sortie. La fonction d'activation des neurones est la fonction sigmoïde unipolaire.

Après avoir effectué plusieurs tests pour fixer les caractéristiques adéquates du réseau, nous avons procédé à la phase d'apprentissage en utilisant un nombre d'itérations considérable pour minimiser l'erreur, ce qui nous avons menés au graphe de la Fig. IV.5.

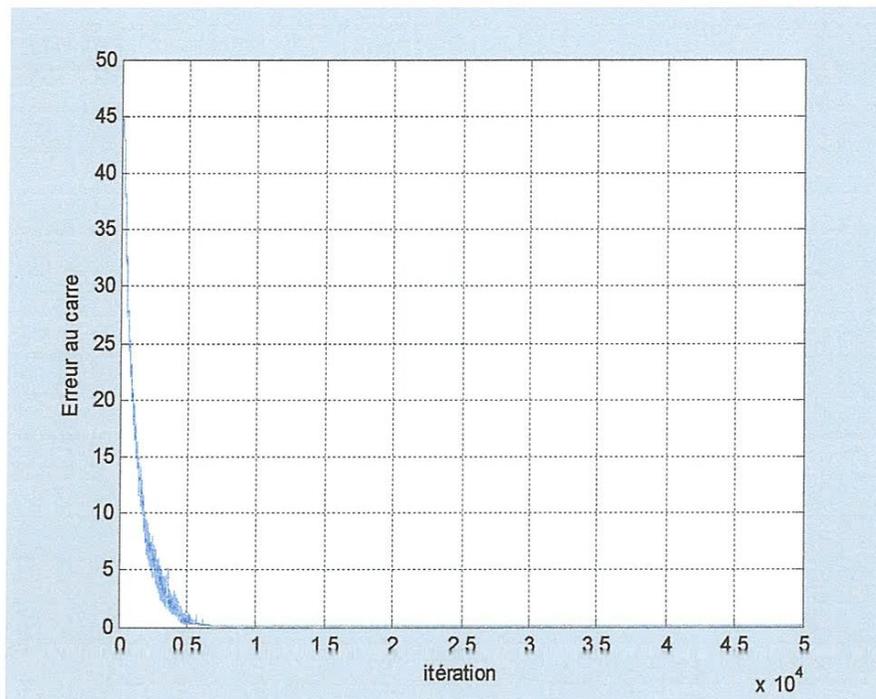


Figure. IV.6 Evolution de l'erreur en fonction du nombre d'iterations

IV.4.2. Reconnaissance

Pour tester la fiabilité de notre système, nous avons effectué des tests sur la base d'apprentissage elle-même, nous avons obtenu des taux de reconnaissance de 100% pour un seuil de rejet égal à 0.0016, Ensuite nous avons effectué les tests reconnaissance sur une base de teste test de cinquante Chiffres de Cinque scripteurs. Les résultats sont illustrés aussi dans le tableau suivant :

	Taux de reconnaissance (en%)	Taux de rejet (en%)
Base d'apprentissage	100%	0%
Base de test	80%	20%

Tableau. IV.2 Taux de reconnaissance sur les ensembles d'apprentissage et de test.

CONCLUSION GENERALE

Le travail effectué dans le cadre de ce projet de fin d'étude nous a permis d'élargir et de mettre en valeur nos connaissances, acquises tout au long des années d'études universitaires, et de découvrir le monde merveilleux de l'électronique spécifiquement la reconnaissance d'écriture des chiffres manuscrite

Malgré les efforts et les travaux réalisés dans le domaine de la reconnaissance aucun système n'est jugé fiable à 100%. Aussi, c'est un domaine qui reste ouvert aux propositions et aux expérimentations.

C'est dans ce cadre que se place notre travail. En effet, les travaux présentés dans ce mémoire abordent les différentes étapes nécessaires à la construction d'un système de reconnaissance des chiffres manuscrite, à savoir : les prétraitements, la segmentation, l'extraction des primitives et la classification.

Notre orientation s'est focalisée sur la segmentation et la classification. Ainsi, dès l'étape d'extraction des primitives, la sélection des primitives pertinentes. Cette sélection consiste à réduire les entrées du classifieur (réseau de neurones de type MLP) tout en améliorant ou en maintenant le taux de reconnaissance de la classification.

Cette modeste expérience s'est avérée intéressante et encourageante mais elle n'a pas la prétention d'être parfaite car on peut toujours envisager des possibilités d'extensions futures, telles que :

- L'utilisation d'une base réelle d'images.
- L'automatisation de la segmentation des chiffres manuscrits.
- L'automatisation de l'étape de compilation : règles du réseau neuronal.

Enfin par cette étude, nous ne pouvons en aucun cas prétendre avoir découvert entièrement le problème de la reconnaissance d'écriture des chiffres manuscrits. Par contre nous pensons vraiment que nos objectifs sont atteints et que ce travail sera un apport certain pour les futurs étudiants de département qui souhaitent s'investir dans ce domaine.

BIBLIOGRAPHIE

- [1] Drablia Salima, Maaref Assia « la reconnaissance de l'écriture arabe par réseaux de neurone » 2009. Thèse de magistère
- [2] Mekhalfia salima, Chouieb Samia « la reconnaissance des chiffres manuscrits par l'analyse en composant principale et les réseaux de neurone » 2007. Thèse de magistère
- [3] Meddour Ammar. « Reconnaissances des empreint digitale »2010. Master II
- [4] keziz et khmessi « reconnaissance de l'écriture arabe manuscrite par un approche hybride » Mémoire d'ingénieur 2005.
- [5] E.davalo , P.naim « les réseaux de neurones » eyrolles 1993.
- [6] M. Leroux, J-C. Salome et J. Badard «Recognition of Cursive Script Words in A Small Lexicon». in Proc. Int'l. Conf. Document Analysis and Recognition, pp.774, Saint Malo, France, Sep. 1991.
- [7] B .Lemarieé, M.Gilloux, M.Leroux « Un modèle neuro-markovien contextuel pour la reconnaissance de l'écriture manuscrite » Actes du congrès AFCET'96 de reconnaissance des formes et intelligence artificielle, 1996.
- [8] H.Miled « Stratégies de reconnaissance de l'écriture semi cursive: Application aux mots manuscrits arabes » Thèse de doctorat, Université de Rouen, 1998
- [9] M.Chrieriet et C.Y. Suen « Un système neuro-flou pour la reconnaissance de montants numériques de cheques arabes » .Pattern Recognition letters 14(1993), pp, 1009-1017.
- [10] E.Marisa « Automatic recognition of handwritten dates on brazilian bank cheques ». Thèse docteur école de technologie supérieure université du Québec, 2003.
- [11] A.Nazif,F.T.Yarmam-Vural «Optical Character Recognition for Cursive Handwriting» IEEE Transactions on Pattern Analysis and Machine Intelligence VOL 24, NO 6, JUIN 2002.
- [12] Alceu de S.Britto, R. Sabourin, F.Bortolozzi,C.Y.Suen «The recognition of handwritten numeral string a two-stage HMM-based method» International Journal on document analysis and recognition, 2003.
- [13] Morita présente un système hybride HMM-MLP pour la segmentation et la reconnaissance de l'écriture manuscrite et plus particulièrement des dates et les montants littéraux présentent sur les chèques brésiliens. Le bilan de cette technique a été positif, un taux de reconnaissance de 82.3% pour les dates et 96.8%pour les mots.