

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY  
UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des sciences de l'ingénieur

Année : 2010

Département d'informatique

### MEMOIRE

Présenté en vue de l'obtention du diplôme de **MAGISTER**

## Reconnaissance de l'écriture Arabe par Systèmes Flous

### Option

Texte, parole et Images

### Par

**NEMOUCHI Soulef**

**DIRECTEUR DE MEMOIRE :** FARAH Nadir      MC      Université d'Annaba

### DEVANT LE JURY

**PRESIDENT :** M<sup>me</sup> BAHI Halima      MC      Université d'Annaba

**EXAMINATEURS:** M<sup>me</sup> SOUCI Labiba      MC      Université d'Annaba

M<sup>r</sup> KHADIR Mohamed Tarek      MC      Université d'Annaba

# ملخص

الهدف الرئيسي من العمل المقدم في هذه المذكرة هي دراسة كيفية قراءة الكلمات العربية المكتوب يدويا بطريقة آلية. وذلك باستعمال البرامج الحاسوبية. فلقد وجدنا بان المشكل الكبير في القراءة الآلية للكتابة اليدوية هي تجزئة الكلمات و في هذا الإطار اهتمنا بمعرفة الكلمة كاملة لتفادي مشكلة التجزئة.

الطرق الكلاسيكية المختارة تسمى الاحصائية وهي فعالة لكنها تستعمل في حالة مجموعة محدودة من الكلمات لتكوين برنامج خاص بها يحتاج إلى مرحلة التعرف على هذه الكلمات. نستخلص بأنه للحصول على برنامج فعال يجب الاهتمام بجميع المراحل المكونة لهذا البرنامج من الرقمنة إلى التعرف. ولقد ركزنا في هذا العمل على مرحلتين أساسيتين هما التشفير و التعرف. كما ركزنا على استعمال الطرق المنطقية التشكيكية و ذلك بتطبيق تقنيات المسماة « فسم » و ذلك بحثا عن حلول لسلبيات التقنيات الكلاسيكية مثل « الكامنز » للحصول على نتائج افضل. استعملنا ايضا تقنية « كاببيفي » . هذا الخير لا يحتاج إلى فرضيات تبنى على المعطيات الموجودة لكنه في نفس الوقت يحتاج إلى مجموعة كبيرة من الامثلة. بالإضافة إلى ذلك استعملنا تقنية الشبكة العصبية « بيانان » التي تستعمل في غالب الأحيان في مثل هذه المجالات. نتائج هذه التقنيات تجمع بطريقتين إما بالتساوي ولا أفضلية لاي منهنم أو بتفضيل احدهم على الآخر.

**كلمات المفاتيح :** المعالجة الآلية للكتابة العربية اليدوية, تقنية الافسم, تقنية الكامنز, تقنية الكاببيفي, الشبكة العصبونية بيانان.

# ABSTRACT

---

The main goal of our study is to address the problem of automatic recognition of Arabic handwritten by statistic methods. We have found that the major problem of the automatic reading of cursive manuscript characters is the segmentation in elements. In this case we are interested by the global approach with avoiding the problem of segmentation.

The chosen classical methods, said statistical are efficient. But they can't be applied just for writing with limited vocabulary, in the case of a system with training phase. We include that the performance of system for the recognition of Arabic handwritten words depends of all recognition phases; from acquisition to classification. We have oriented our research to approaches based fuzzy logic concepts: implementation of FCM algorithm for automatic classification of Arabic handwritten words, trying to find solutions to disadvantage of K-means algorithm, and to obtain more preferment results. We have implemented a KNN algorithm, this one necessitate no hypotheses in the form of training classes. However it necessitates a great number of observations. In addition, it has been greatly used in handwritten recognition. The last classifier chosen is the PNN network; this type of network is generally used for classification problems. Their results are combined then using two methods of combination.

**Keywords:** Recognition of Arabic handwritten words, Fuzzy C-Means algorithm, K-means algorithm, KNN algorithm, PNN network.

# RESUME

---

L'objectif de notre travail de recherche est d'adresser la problématique de la reconnaissance automatique de l'écriture arabe manuscrite par les méthodes statistiques. Nous avons trouvé que le problème majeur de la lecture automatique de caractères manuscrits cursifs étant la segmentation d'un tracé en éléments constituants, dans ce cadre, nous nous sommes intéressés à l'approche globale en évitant le problème de segmentation.

Les méthodes classiques choisies, dites statistiques, sont efficaces mais ne peuvent s'appliquer qu'à des écritures à vocabulaire restreint dans le cadre d'un système avec une phase d'apprentissage. Nous constatons que la performance d'un système de reconnaissance optique de l'écriture arabe manuscrite dépend de toutes les phases du processus de reconnaissance depuis l'acquisition jusqu'à la classification. Nous avons mis l'accent sur deux phases les plus étudiées dans les travaux existants : l'extraction des primitives et la classification. Nous avons orienté notre recherche vers les approches basées sur les concepts de la logique floue : application de l'algorithme FCM pour la classification automatique des mots arabe manuscrit, en essayant ainsi de trouver des solutions aux inconvénients de l'algorithme classique K-Means, et d'obtenir des résultats plus performants. Nous avons implémenté aussi l'algorithme KPPV, ce dernier ne nécessite aucune hypothèse sur la forme des classes d'apprentissage par contre, il nécessite un grand nombre d'observations pour être précis. De plus il a été beaucoup utilisé en reconnaissance de l'écriture manuscrite. Et le dernier classifieur choisi, c'est le réseau PNN ; ce type de réseaux est généralement utilisé pour des problèmes de classification. Leurs résultats sont ensuite combinés en utilisant deux méthodes de combinaison sans apprentissage : la méthode de vote simple et celle de la somme pondérée.

**Mots clés :** Reconnaissance de mots arabes manuscrits, algorithme Fuzzy C-Means, algorithme K-Means, algorithme KPPV, Réseau de neurone PNN.

# Dédicace

---

*A ceux qui me sont chers...*

# REMERCIEMENTS

---

*Au terme de ce travail je voudrais tout d'abord exprimer mes remerciements à DIEU de me donner le courage et la volonté pour finir ce mémoire.*

*Après Dieu, je tiens à exprimer toute ma reconnaissance et mon profond respect à mon encadreur Mr. FARAH Nadir, maître de conférences à l'université d'Annaba, d'avoir accepté de diriger mon travail, de m'avoir fait confiance et de m'avoir laissée m'exprimer.*

*Je souhaite adresser mes sincères remerciements à madame BAHI Halima, maître de conférences à l'université d'Annaba, pour m'avoir fait l'honneur de présider le jury de ma soutenance.*

*Je tiens aussi à remercier particulièrement les personnes qui ont accepté d'examiner et de juger mon travail et qui sont : monsieur KHADIR Mohamed Tarek et madame SOUCI Labiba, maîtres de conférences à l'université d'Annaba. Je les remercie pour leurs conseils, pour leurs disponibilités à tout moment, pour leurs gentillesse, et pour leurs aides illimitées.*

*Enfin, mes remerciements vont à tous les membres de ma famille qui m'ont tant soutenue, particulièrement à mes parents pour leur soutien moral et pour leur encouragement durant ma période de scolarité. A mes sœurs surtout Fatiha et Nadia pour leur aide, à mes petites Lina et Aya. Merci à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce travail.*

# Liste des Tableaux

Tableau	Titre	Page
Tableau 1.1	Avantages et inconvénients des l'approches globale et analytique	13
Tableau 3.1	Alphabet Arabe	41
Tableau 3.2	Hamza et Madda et les positions qu'elles occupent en association avec Alif, Waw et Ya	41
Tableau 3.3	Différents formes d'un caractère arabe	42
Tableau 3.4	Exemple de mots composés de la droite vers la gauche de 1, 2, 3,4 et 5 PAWs respectivement	43
Tableau 3.5	Le PAW « فر » dans différents mots et différentes positions	43
Tableau 3.6	Tableau récapitulatif précisant les caractéristiques et les performances de certains systèmes AOCR (approche globale	50
Tableau 4.1	Taxonomie des méthodes d'extraction de caractéristiques selon la représentation de l'image	64
Tableau 4.2	Taxonomie des caractéristiques selon la méthode de classification	65
Tableau 4.3	Les Différents Type Des Points Diacritiques	71
Tableau 4.4	Description globale des noms de wilayas en mots	77
Tableau 4.5	Taux de reconnaissance obtenue en faisant varier le coefficient flou m et en utilisant les caractéristiques de zonage	82
Tableau 4.6	Taux de reconnaissance obtenue en faisant varier le coefficient flou m et en utilisant les caractéristiques de freeman.	82
Tableau 4.7	Taux de reconnaissance obtenue en faisant varier le coefficient flou m et en utilisant les caractéristiques de moment de Zernike	82
Tableau 4.8	Résultats de classification simple (a,..., d)	87
Tableau 4.9	Résultats de combinaison des classifieurs	88
Tableau 4.11	. L'influence de paramètre K sur le résultat de la classification KPPV	89
Tableau 4.12	Résultats de K-means sans découpage de vocabulaire	89

# Liste des Figures

<b>Figure</b>	<b>Titre</b>	<b>Page</b>
Figure 1.1	Rand Tablet ou Grafacon	9
Figure 1.2	Différents types d'écritures	14
Figure 1.3	Styles d'écritures selon TAPPERT	15
Figure 1.4	Graphe de complexité des systèmes de RAED d'après BELAID	16
Figure 1.5	systèmes, représentations et approches de reconnaissance	22
Figure 2.1	Schéma général d'un système de reconnaissance de l'écriture hors-ligne	26
Figure 2.2	Effets de certaines opérations de prétraitement	27
Figure 2.3	Le voisinage en carreau et carré	28
Figure 2.4	Exemples de Redressement de mots manuscrits	29
Figure 2.5	Détection de ligne de base	29
Figure 2.6	Les différents types de segmentation	31
Figure 3.1	Racine trilére 'KTB' en arabe	42
Figure 3.2	Exemple de formes de boucles dans des styles différents	43
Figure 3.3	Le dédoublement de lettre « chadda »	44
Figure 3.4	Les voyelles de la langue arabe	44
Figure 3.5	La composition des mots en langue arabe	45
Figure 3.6	Différentes combinaisons verticales possibles de caractères	46
Figure 3.7	Ecriture voyellée	47
Figure 3.8	Exemples d'écritures arabes reflétant ses caractéristiques morphologiques	48
Figure 4.1	Schéma général du système de reconnaissance de mot manuscrit	53
Figure 4.2	Prétraitement de l'image	54
Figure 4.3	Positionnement de notre travail par rapport à la dichotomie des types d'écriture	55
Figure 4.4	Détermination du seuil de binarisation	56
Figure 4.5	Résultat d'une binarisation	56
Figure 4.6	Résultat de lissage	57
Figure 4.7	Traitement d'image d'haut niveau	58
Figure 4.8	Ligne de base du mot « Tamenraset ».	59
Figure 4.9	Zone médiane du mot « Oum-el-bouaghi ».	59
Figure 4.10	Illustration de la hauteur de la zone médiane en pixels dans un mot échantillon	60
Figure 4.11	Zone supérieure du mot « Oum-el-bouaghi ».	60
Figure 4.12	Zone inférieure du mot « Oum-el-bouaghi ».	61
Figure 4.13	Les 11 composantes connexes qui forment le mot « Oum-el-bouaghi ».	62
Figure 4.14	Illustration du résultat de l'extraction des composantes connexes	62

Figure 4.15	Résultat de squelettisation	63
Figure 4.16	Extraction de contour	63
Figure 4.17	Extraction morphologique de contour externe, (a) image originale, (b) image dilatée, (c) contour externe	67
Figure 4.18	Illustrations des différentes informations nécessaires pour l'analyse d'une composante connexe	69
Figure 4.19	Topologie De L'écriture Arabe Illustrée Dans Le Mot « Oum-El-Bouaghi ».	70
Figure 4.20	Organigramme de détection des diacritiques (Catégorie 1)	71
Figure 4.21	Types de points diacritiques (catégorie2 et catégorie3)	72
Figure 4.22	Organigramme de détection des diacritiques (les 3 catégories)	73
Figure 4.23	Trois Fonds Correspondent Aux Trois Boucles	74
Figure 4.24	Illustration du niveau de projection verticale pour obtenir le nombre de hampes	76
Figure 4.25	Algorithme de classification par C-moyenne flou	81
Figure 4.26	L'algorithme de résolution pour la méthode K-Means	83
Figure 4.27	Architecture du réseau PNN	84
Figure 4.28	Structure d'un réseau PNN	85

# Liste des Symboles

AOCR: Arabic Optical Character Recognition.

ASCII: American Standard code for information interchange

ASMO: Arabic Standard Metrology Organization

CC: composante connexe

COROC: Cognitive Optical Recognition Old Characters

FCM: Fuzzy C-Means Clustering.

KPPV: K Plus Proche Voisins

LAD: Lecture Automatique des Documents

MMC: Modèles de Markov Cachés

OCR: Optical Character Recognition.

PAW: Peace of Arabic Word.

PNN: Probabilistic Neural Network

RAED : Reconnaissance Automatique de l'Écriture et du Document

REAM : Reconnaissance de l'Écriture Arabe Manuscrite.

RBF: Radial Basic Functions.

RdF: Reconnaissance de Forme

RECAM: Reconnaissance de l'Écriture Cursive Arabe Manuscrite

REM: Reconnaissance de l'Écriture Manuscrite

SBR : Systèmes à Base de Règles

SREM : Système de Reconnaissance de l'Écriture Manuscrite

# Table des matières

ملخص.....	i
ABSTRACT.....	ii
RESUME .....	iii
Liste des Tableaux .....	vi
Liste des Figures .....	vii
Liste des Symboles .....	ix
Table des matières.....	x
<b>INTRODUCTION GENERALE .....</b>	<b>1</b>
Problématique et objectifs.....	3
Présentation du mémoire.....	5
<b>Chapitre1.Reconnaissance Automatique De L'écriture Manuscrite .....</b>	<b>7</b>
<b>I.1. Introduction .....</b>	<b>7</b>
<b>I.2. Historique de LAD (Lecture Automatique des Documents) .....</b>	<b>8</b>
<b>I.3. Différents aspects de la Reconnaissance Automatique de l'Écriture et du Document ou RAED .....</b>	<b>8</b>
I.3.1. Mode d'acquisition (En-ligne / Hors-ligne) .....	9
I.3.1.1. Reconnaissance en-ligne .....	9
I.3.1.2. Reconnaissance hors-ligne .....	10
I.3.1.2.1. Reconnaissance de texte ou analyse de documents .....	11
I.3.1.2.2. Reconnaissance de l'imprimé ou du manuscrit.....	11
I.3.2. Approches de la Reconnaissance .....	11
I.3.2.1. Approche globale .....	12
I.3.2.2. Approche analytique .....	12
I.3.2.3. Approche basée sur la lecture humaine.....	13
I.3.2.4. Avantages et Inconvénients de l'approche globale et l'analytique .....	13
<b>I.4. Complexité d'un système de RAED.....</b>	<b>14</b>
I.4.1. Disposition spatiale du texte .....	14
I.4.1.1. Types d'écriture .....	14
I.4.1.2. Styles d'écriture.....	14
I.4.2. Nombre de scripteurs .....	15
I.4.3. Taille du vocabulaire.....	15
<b>I.5. Méthodes de classification de la reconnaissance .....</b>	<b>16</b>
I.5.1. Méthodes statistiques .....	16
I.5.1.1. Méthode bayésienne.....	17
I.5.1.2. Méthode du plus proche voisin .....	18
I.5.1.3. Méthode connexionniste .....	18
I.5.1.3.1. Réseaux non bouclés .....	19
I.5.1.3.2. Réseaux bouclés .....	19
I.5.2. Méthodes stochastiques.....	19
I.5.3. Méthodes linguistiques.....	20
I.5.3.1. Méthode structurelle.....	20
I.5.3.1.1. Méthodes de tests .....	21
I.5.3.1.2. Comparaison de chaînes.....	21
I.5.3.2. Méthode syntaxique .....	21
I.5.4. Méthodes hybrides .....	22
<b>I.6. Conclusion.....</b>	<b>23</b>

<b>Chapitre2. Processus De Reconnaissance D'écriture Manuscrite.....</b>	<b>25</b>
<b>II.1. Introduction.....</b>	<b>25</b>
<b>II.2. Organisation générale d'un SREM.....</b>	<b>25</b>
I.2.1. Phase d'Acquisition d'image .....	26
II.2.2. Phase de Prétraitement.....	27
II.2.2.1. Réduction du bruit .....	27
II.2.2.1.1. Seuillage.....	28
II.2.2.1.2. Lissage .....	28
II.2.2.2. Redressement de l'écriture.....	29
II.2.2.3. Homogénéisation des données.....	29
II.2.2.4. Normalisation.....	30
II.2.3. Squelettisation.....	30
II.2.4. Phase de Segmentation .....	31
II.2.5. Phase d'Analyse ou Extraction des Caractéristiques .....	32
II.2.5.1. Niveaux des caractéristiques.....	32
II.2.5.2. Représentation des caractéristiques .....	32
II.2.5.3. Types de caractéristiques .....	33
II.2.5.3.1. Caractéristiques topologiques ou métriques .....	33
II.2.5.3.2. Caractéristiques locales ou structurelles .....	34
II.2.5.3.3. Caractéristiques statistiques .....	34
II.2.5.3.4. Transformations globales.....	34
II.2.5.3.5. Superposition des modèles (template matching) et corrélation .....	35
II.2.6. Phase de Classification .....	35
II.2.6.1. Apprentissage.....	35
II.2.6.1.1. Apprentissage supervisé .....	35
II.2.6.1.2. Apprentissage non supervisé.....	36
II.2.6.2. Reconnaissance .....	36
II.2.6.2.1. Les systèmes à base de règles SBR (y compris les systèmes experts).....	36
II.2.6.2.2. Les arbres de décision.....	36
II.2.6.2.3. Les techniques agglomératives .....	37
II.2.7. Poste-traitement .....	37
<b>II.3. Conclusion .....</b>	<b>38</b>
<b>Chapitre 3. Ecriture Arabe Et L'OCR .....</b>	<b>40</b>
<b>III.1. Origine de l'écriture arabe .....</b>	<b>40</b>
<b>III.2. Réforme de l'écriture arabe .....</b>	<b>40</b>
<b>III.3. Caractéristiques de l'écriture arabe .....</b>	<b>40</b>
<b>III.4. Vocalisation de l'écriture arabe .....</b>	<b>43</b>
<b>III.5. Problèmes posés par l'écriture arabe .....</b>	<b>45</b>
III.5.1. Multitude de graphismes .....	45
III.5.2. Ligatures.....	45
III.5.3. Vocalisation.....	46
III.5.4. Normalisation des caractères.....	47
<b>III.6. Difficultés de la reconnaissance de l'écriture arabe.....</b>	<b>48</b>
<b>III.7. Application de la reconnaissance hors-ligne de l'écriture.....</b>	<b>48</b>
<b>III.8. Conclusion.....</b>	<b>50</b>
<b>Chapitre 4. Approche globale pour la reconnaissance de mots arabes manuscrits .....</b>	<b>52</b>

<b>IV.1. Introduction .....</b>	<b>52</b>
<b>IV.2. Architecture du système.....</b>	<b>52</b>
<b>IV.3. Acquisition et Prétraitement .....</b>	<b>54</b>
IV.3.1. Binarisation .....	56
IV.3.2. Lissage.....	57
IV.3.2.1. Bouchage.....	57
IV.3.2.2. Nettoyage .....	57
<b>IV.4. Segmentation de textes (Traitement d'images de haut niveau) .....</b>	<b>58</b>
IV.4.1. Localisation de la ligne de base et la zone médiane .....	58
IV.4.1.1. Localisation de la ligne de base.....	58
IV.4.1.2. Localisation de la zone médiane .....	59
IV.4.2. Localisation de la zone supérieure et la zone inférieure.....	60
IV.4.2.1. Localisation de la zone supérieure .....	60
IV.4.2.2. Localisation de la zone inférieure .....	61
IV.4.3. Détection des composantes connexes.....	61
<b>IV.5. Squelettisation.....</b>	<b>62</b>
<b>IV.6. Traçage de contours .....</b>	<b>63</b>
<b>IV.7. Extraction des caractéristiques globales et description du mot .....</b>	<b>63</b>
IV.7.1. Choix des caractéristiques .....	65
IV.7.2. Descripteurs globaux.....	65
IV.7.3. Descripteurs externes .....	66
IV.7.3.1. Extraction de contour .....	66
IV.7.3.2. Détermination des contours des mots.....	67
IV.7.4. Zonage.....	67
IV.7.5. Caractéristiques structurelles d'un mot arabes.....	67
IV.7.5.1. Détection des différents types des points diacritiques.....	70
IV.7.5.2. Détection des boucles (occlusions) .....	74
IV.7.5.3. Détecter le nombre d'ascendants (hampes) et de descendants (jambages) .....	74
IV.7.5.3.1. Détection des jambages .....	75
IV.7.5.3.2. Détection des hampes.....	75
<b>IV.8. décomposition du vocabulaire considéré.....</b>	<b>76</b>
<b>IV.9. Reconnaissance .....</b>	<b>77</b>
IV.9.1. Sélection des attributs.....	78
IV.9.2. Classification par les méthodes implémentées .....	79
IV.9.2.1. Classification par la méthode KPPV .....	79
IV.9.2.2. Classification par la méthode C-Moyenne Floue .....	80
IV.9.2.3. Classification par méthode K-Means .....	82
IV.9.2.4. Classification par réseaux de neurones probabilistes (PNN) .....	84
III.9.2.4.1. Propriétés générales d'un PNN.....	84
IV.9.2.4.1. Avantages et inconvénients de PNN .....	85
IV.9.3. Résultats et discussion.....	85
IV.9.3.1. Classification simple .....	86
IV.9.3.2. Combinaison de classifieurs .....	88
IV.9.3.3. Influence de paramètre K sur les résultats de la classification KPPV .....	89
IV.9.3.4. Utilité de décomposition du vocabulaire.....	89
<b>IV.10. Conclusion .....</b>	<b>90</b>
<b>CONCLUSION GENERALE .....</b>	<b>91</b>
<b>BIBLIOGRAPHIE.....</b>	<b>93</b>
<b>ANNEXE A.....</b>	<b>104</b>
<b>ANNEXE B.....</b>	<b>108</b>



## *Introduction générale*

---

Comment l'être humain fait-il pour raisonner, parler, calculer, apprendre? Comment s'y prendre pour provoquer: la réflexion, la pensée, le jugement, le discernement et l'intellect...?

Ces interrogations ont menées à des recherches sur le lien entre l'esprit humain, son cerveau et la machine. Ces recherches ont données naissance à des fondations de ce que nous appelons aujourd'hui l'intelligence artificielle. Parmi les domaines de cette dernière nous trouvons la reconnaissance des formes RdF qui consiste à identifier des formes ou des objets tout en se basant sur certaines informations représentatives de la forme.

La reconnaissance optique de caractères « Optical Character Recognition OCR», au sens large du terme, est une discipline aussi vieille que l'ordinateur. Elle pose la question de l'intelligence humaine et des possibilités de l'intelligence artificielle. Pourquoi nous semble-t-il si facile de déchiffrer un texte manuscrit (même relativement mal écrit) alors que cette tâche devient si complexe à automatiser ? [ARR 02].

Alors où se situe la difficulté ? D'après SIMON [SIM 84], il y a deux niveaux de difficulté. La première difficulté réside dans la complexité calculatoire. En effet la construction d'un espace de représentation pour une application donnée, et plus particulièrement pour la reconnaissance de caractères manuscrits, se heurte hier, comme aujourd'hui, à des problèmes informatiques de temps de calcul, de temps d'accès ou de stockage. La seconde difficulté se situe dans la sémantique même du problème :

- Existe-t-il une technique générale pour construire un opérateur de RdF ?
- Peut-on construire un système auto-apprenant et auto-organisateur capable de reconnaître n'importe quelle forme pour peu qu'elle ait été apprise ? [SIM 84].

Les techniques liées au traitement de l'information connaissent actuellement un développement très actif en liaison avec l'informatique et présentent un potentiel de plus en plus important dans le domaine de l'interaction Homme-machine. L'Homme veut communiquer avec l'ordinateur de la façon la plus simple, la plus naturelle pour faciliter et accélérer l'interaction et l'échange d'informations. Il cherche à rendre ces machines accessibles par la voix, capables de lire, de voir, de se déplacer...et de traiter et d'analyser rapidement l'information reçue.

Ecrire pour communiquer a été de tous les temps une préoccupation première de l'Homme. L'écrit a été, et restera, l'un des grands fondements des civilisations et le mode par excellence de conservation et de transmission du savoir. Malgré les avancées d'autres moyens de communication tel que l'audio visuel, nombreuses sont les applications dont l'existence commence sur le papier, plus particulièrement dans la bureautique (saisie, indexation et archivage automatique de documents), en

publication assistée par ordinateur (pour faciliter la composition à partir d'une sélection de plusieurs documents), dans la poste (lecture des adresses et tri automatique), dans les banques (traitement des chèques, des factures)... Cependant malgré les progrès technologiques, le clavier reste encore un moyen obligé de communication avec l'ordinateur.

Quelques systèmes de reconnaissance de l'écriture manuscrite SREM ont été réalisés et sont opérationnels à ce jour. Cependant, ils sont spécifiques à un domaine précis et sont encore limités. Par exemple, en ce qui concerne la poste, la reconnaissance de l'écriture manuscrite REM, contrairement à celle des caractères imprimés, se limite au code postal en chiffres ainsi qu'à la ville en caractères majuscules [IMP 97, KAU 00, KB 00, GAA 01]...etc.

Contrairement au latin, la reconnaissance de l'écriture arabe manuscrite REAM ou imprimée reste encore aujourd'hui au niveau de la recherche et de l'expérimentation, le problème n'est pas encore résolu bien que l'on sache atteindre des taux assez élevés dans certaines applications pour lesquelles soit le vocabulaire est limité [ALE 90], soit la fonte est unique [ALB 95a]. Ce retard est dû principalement à la complexité de la morphologie de l'écriture arabe, le manque d'échange entre les chercheurs du domaine et celui d'infrastructures adéquates dans le sens budgétaire, de bases de données, de dictionnaires et d'outils de test et de validation. Mais dernièrement plusieurs travaux ont récemment vu le jour.

La reconnaissance de l'écriture arabe date des années 80 [AMI 85, SAA 85]. Depuis, les recherches se sont multipliées dans ce domaine. Certains chercheurs se sont intéressés à la reconnaissance en temps réel en utilisant des tablettes graphiques, ce qui simplifie en partie le problème en restituant le sens du tracé, d'autres se sont penchés sur l'imprimé et/ou le manuscrit en "Off-line" en utilisant un scanner ou une caméra pour la saisie des documents. Depuis, plusieurs outils de prétraitement ont été développés pour la squelettisation de l'image, de son lissage, pour la détermination du contour et l'extraction de primitives...et différentes techniques de reconnaissance ont été élaborées, parmi lesquelles figurent des méthodes statistiques, structurelles ou géométriques, avec ou sans segmentation en caractères [ELD 90, ROM 95]. Généralement toutes ces méthodes et bien d'autres, tendent à extraire, chacune à sa façon, une catégorie de caractéristiques et d'évaluer par la suite la vraisemblance entre les primitives extraites et celles de formes prototypes déjà apprises par le système de reconnaissance.

La reconnaissance des chiffres a bénéficié de très nombreux travaux ces dernières années, notamment dans le cadre de la reconnaissance de montants numériques de chèques, de champs numériques dans les formulaires, ou encore de reconnaissance de codes postaux dans les adresses postales et la reconnaissance des 10 chiffres manuscrits écrits par n'importe quel usager. L'objectif principal de toute recherche est d'aboutir à de bons résultats, ce qui conduit à des systèmes performants, mais en focalisant sur le choix de la méthode celle qui donne au mieux que possible la bonne décision, telles que les méthodes statistiques, structurelles, les réseaux de neurones,..., avec possibilité de les combinées.

Depuis les années 90, la combinaison de classifieurs a donc été une des directions de recherche les plus soutenues dans le domaine de la RdF. Les méthodes de combinaison ont ainsi été appliquées avec succès dans les domaines de la

reconnaissance de l'écrit. L'amélioration des performances des systèmes de reconnaissance est finalement le principal enjeu des recherches menées ces dernières années sur les systèmes de combinaison.

Un système de reconnaissance de l'écriture doit idéalement, localiser, reconnaître et interpréter n'importe quel texte ou nombre écrit sur un support de qualité arbitrairement variable tel que des formulaires ou des vieux manuscrits pour être enregistré dans des bases de données que tout le monde peut l'interroger.

En effet, il existe dans le domaine de la reconnaissance de l'écriture arabe, et particulièrement l'écriture manuscrite, beaucoup de problèmes intéressants à étudier aussi bien du point de vue scientifique (évaluation des méthodes de reconnaissance) que du point de vue économique (objectifs applicatifs réalistes et rentables). Le travail présenté dans ce mémoire se déroule dans le cadre du développement d'un système complet de reconnaissance des mots arabes manuscrits qui sont les 48 mots de wilayas algériennes arabes manuscrits hors ligne avec une approche multiclassifieurs.

### ***Problématique et objectifs***

Jusqu'à présent, l'établissement d'un système universel pour la REM reste un sujet irréel. Puisque les êtres humains sont les meilleurs lecteurs, il semble naturel d'étudier leur manière de lire avant d'essayer d'automatiser le processus de lecture. Par ailleurs, les différents travaux élaborés en AOCR, montrent que la cursivité de l'écriture arabe, la complexité de la morphologie des caractères, les élongations des ligatures horizontales ainsi que les combinaisons verticales de certains caractères, constituent les problèmes majeurs liés au traitement de cette écriture. En effet, ces problèmes engendrent une forte inertie à différents niveaux notamment dans :

- Le choix de primitives pertinentes décrivant la variabilité de la morphologie des caractères, sachant que certaines caractéristiques topologiques sont sensibles à la dégradation, notamment les points diacritiques et les boucles.
- La méthode de segmentation en caractères ou même en pseudo-mots (qui peuvent se chevaucher surtout dans le cas du manuscrit).

Tous ces problèmes et bien d'autres, se trouvent accentués dans le cas du manuscrit où d'autres facteurs interviennent (variabilité intra et inter-scripteurs, conditions de l'écriture, fusion de points diacritiques, chevauchement de pseudo-mots, graphismes inégalement proportionnés...).

Tous les types de caractéristiques (statistiques, structurelles, géométriques, transformations globales) et des méthodes de classification (statistique, structurelle, stochastique, connexioniste, logique flou,..) qui existent dans la littérature sont pratiquement utilisées dans la reconnaissance de l'écriture arabe.

Il y a trois façons d'aborder cette problématique. Soit le système reconnaît le mot comme une entité entière et indivisible, il s'agit d'une approche *globale* ou *holistique*. Soit il reconnaît le mot à partir de ses caractères préalablement segmentés, il s'agit d'une approche *analytique*. Soit il n'utilise que certaines propriétés et raffine sa

description du mot par rebouclage, nous parlerons alors de *systèmes basés sur la lecture humaine*.

A partir ce que nous avons vue dans le paragraphe ci-dessus, nous focalisons notre travail à la première approche (globale), nous allons aborder la reconnaissance de mots arabes en appliquant des méthodes statistiques : FCM, K-MEANS, KPPV et un réseau de neurone probabiliste PNN. Nous obtenons un système qui combine parallèlement les sorties de ces classifieurs.

Afin de simplifier la problématique, et donc, d'envisager des solutions plus ou moins convenables, il est important de spécifier le cadre de reconnaissance puis de construire l'approche qui va de paire avec ce cadre. Les critères qui influencent le processus de reconnaissance sont le mode d'acquisition d'écriture, la taille et le type du vocabulaire considéré, le nombre de scripteurs, le style et le type d'écriture.

De nombreux travaux montrent que la combinaison de classifieurs (séquentielle, parallèle ou hybride) améliore nettement les performances du système de reconnaissance par rapport à chacun des classifieurs pris isolément. Cependant l'architecture parallèle est celle qui a donné lieu aux travaux les plus importants.

En résumé que tout système de reconnaissance de l'écriture nécessite généralement une étape très importante qui est l'étape d'extraction de caractéristiques. Dans la littérature, plusieurs travaux ont porté sur l'élaboration de nouvelles caractéristiques qui ont un pouvoir discriminant inter-classes accru tout en minimisant la variabilité intra-classes. Ces caractéristiques sont généralement classées en deux familles : les caractéristiques structurelles (boucles, concavités, dépassements, points extrêmes du tracé, intersections, ...etc) et les caractéristiques statistiques qui dérivent des mesures de distribution spatiale des pixels (zonage, moments invariants, descripteurs de Fourier...etc).

La variabilité de l'écriture manuscrite permet de configurer les algorithmes de classification et d'apprentissage à des problèmes difficiles. Les réseaux de neurones ont montré des résultats remarquables dans ce domaine, mais la nécessité de performances élevées dans des applications réelles ont poussé la recherche vers d'autres modèles de classification de plus en plus complexes.

Nous avons comme objectif principal de réaliser un système de reconnaissance de mots arabe manuscrits. Nous nous sommes focalisé dans notre travail sur la combinaison parallèle car cette dernière a prouvé son efficacité dans de nombreux problèmes de classification, ce succès est dû à sa simplicité de mise en œuvre, sa capacité à exploiter les réponses des différent classifieurs à combiner, en prenant en compte ou non le comportement de chacun des classifieurs.

Dans ce travail nous avons centralisé notre but à l'étape de l'extraction des caractéristiques car elle est la plus intéressante dans le processus de reconnaissance surtout dans le cas de l'écriture non-contrainte manuscrite.

Le système réalisé c'est un système statistique qui nécessite un apprentissage et est donc relativement dédié. Il a été testé avec une base multi-scripteurs et n'est, en l'état, applicable qu'aux écritures à vocabulaire réduit et dédié à notre base. Nous

présentons dans ce manuscrit les apports d'un tel système par rapport à d'autres approches.

### ***Présentation du mémoire***

Ce mémoire est structuré en quatre chapitres comme suit :

#### **Chapitre 1. Reconnaissance automatique de l'écriture manuscrite**

Nous décrivons en premier lieu notre domaine d'application à travers une présentation de l'historique de LAD et les différents aspects de la reconnaissance de l'écriture. Nous avons mis l'accent sur la reconnaissance de l'écriture manuscrite en se basant à l'approche globale, que nous avons adoptée dans notre travail. Suivi par une description générale des principales méthodes de reconnaissance.

#### **Chapitre 2. Processus de reconnaissance d'écriture manuscrite**

Dans ce chapitre, nous abordons en détail les étapes nécessaires au processus de développement d'un système de reconnaissance hors-ligne de l'écriture manuscrite.

#### **Chapitre 3. Ecriture arabe et l'OCR**

Ce chapitre présente des généralités sur l'écriture arabe, son origine ; ces caractéristiques, et les principales difficultés rencontrées dans le domaine de reconnaissance de mots arabe manuscrits. Nous présentons par la suite quelques travaux concernant la reconnaissance globale de mots dans des vocabulaires limités.

#### **Chapitre 4. Approche globale pour la reconnaissance de mots arabes manuscrits**

Dans le dernier chapitre, nous représentons plus clairement notre travail, nous discutons en détail des différentes méthodes appliquées dans notre processus de reconnaissance, par la représentation des concepts de base. Nous présentons aussi les résultats obtenus sur notre base de mots utilisée dans le cadre de ce projet, nous montrons également par étude comparative entre les trois types de caractéristiques choisies ainsi que les classifieurs implémentés, en améliorant ces résultats indépendants par une combinaison entre eux.

Enfin, cette thèse est clôturée par une conclusion dans laquelle nous dressons un bilan et une synthèse de travail effectué, et nous discutons un ensemble de perspectives que nous jugeons nécessaire à l'achèvement de notre travail.

*CHAPITRE 1*

*RECONNAISSANCE  
AUTOMATIQUE DE  
L'ÉCRITURE  
MANUSCRITE*

## I.1. Introduction

De nos jours, l'écriture reste 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 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 RdF. 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 RdF.

La LAD (Lecture Automatique des Documents) et plus généralement la RdF sont des domaines de recherche actifs depuis la fin des années soixante. Des systèmes ont vu le jour pour des applications très spécifiques, mais comparées aux fantastiques progrès réalisés en médecine par exemple, et malgré une recrudescence des travaux de recherche depuis les années 80, les avancées restent lentes et encore décevantes. Pourtant le champ de la RdF est considérable et les publications scientifiques dans ce domaine sont nombreuses. Elles atteignaient déjà le millier par an au début des années 80 [SIM 84] et sont en explosion depuis. De plus, d'un point de vue mathématique, le problème posé par la RdF est trivial, comme le soulignait SIMON en 1984 [SIM 84] ; [ARR 02]: « Soit  $X$  un espace de représentation - de préférence un « bon espace topologique » - et un ensemble fini de noms -l'espace d'interprétation-. Une reconnaissance - une identification - est une application  $E : X \rightarrow \Omega$ , que l'on va pourvoir de « propriétés » ; on en déduira de jolis théorèmes... ».

L'OCR (Optical Character Recognition) est une opération informatique rapide permettant de réaliser la transformation d'un texte écrit sur papier en un texte sous forme d'un fichier informatique en représentation symbolique (par exemple pour les écritures latines, le codage opéré est le code ASCII (American Standard Code for Information Interchange), tandis que pour l'arabe on utilise généralement le code ASMO (Arabic Standard Metrology Organization). Il existe à l'heure actuelle un certain nombre d'OCR qui fonctionnent de façon tout à fait satisfaisante sur les documents imprimés. Mais le traitement des documents manuscrits pose toujours des problèmes. Cette difficulté liée à la modalité elle-même est encore accrue par les problèmes résultants de la qualité du support.

Dans ce chapitre, nous allons présenter l'historique de LAD, puis nous expliquons les différents aspects de la reconnaissance de l'écriture, nous poursuivrons par une description du modèle général de la REM. Nous terminerons en exposant nos conclusions et la direction choisie pour notre travail.

## **I.2. Historique de LAD (Lecture Automatique des Documents)**

Commençons par survoler l'historique de la LAD et plus spécifiquement de la reconnaissance d'écriture. Nous ne reviendrons pas sur les débuts de l'informatique dont les capacités de stockage et de représentation ne permettent pas de vrais traitements optiques des textes. Tout commence véritablement vers le milieu des années 60 avec la création de l'Extended ASCII en 1965 qui permet la représentation de 256 caractères [ARR 02]. Pendant cette période, les premiers systèmes de lecture automatique du texte imprimé ont vu le jour. Toutefois, des systèmes fiables étaient restreints à quelques fontes seulement [AYA 04]. Ensuite il y a, grossièrement, trois périodes distinctes dans le développement de l'OCR décrite dans [ARR 02].

La reconnaissance de l'écriture arabe remonte aux années 70, depuis plusieurs solutions ont été proposées. Elles sont aussi variées que celles utilisées pour le latin [BEN 01b].

Entre les années 1980 et 1990, Les réseaux de neurones ont montré des résultats remarquables dans ce domaine. Jusqu'à très récemment, le perceptron multicouche a été rapidement reconnu comme le classifieur par excellence dans beaucoup de problèmes de reconnaissance de caractères. Par la suite, le besoin d'automatisation massive a donné lieu à toute une multitude d'applications dont la lecture de chèques bancaires, des adresses postales, des documents imprimés, etc. [AYA 04].

Quelques années plus tard, des systèmes de REM ont été réalisés et sont opérationnels à ce jour. Cependant, ils sont spécifiques à un domaine précis et sont encore limités [DAR 94], parmi ces systèmes ceux qui sont orientés vers l'écriture pour la reconnaissance d'adresses manuscrites ou de montants de chèques, on trouve [GAA 01, KB 00, KHP 05, LLG 95]. Parmi eux qui sont orientés vers l'écriture arabe manuscrite citant à titre d'exemple le système Reconnaissance de l'écriture Cursive Arabe Manuscrite RECAM.

On a d'autres travaux qui sont focalisés à développer un logiciel « auto-apprenant » de reconnaissance de caractères manuscrits cursifs (documents anciens notamment). Ce système OCR veut s'affranchir des bibliothèques de langues et de la forme des caractères. Il doit reconnaître après une phase d'apprentissage automatique la plupart des caractères. Plusieurs langues ont été traitées parmi eux la langue arabe ; chinoise, etc. En explorant les méthodes de reconnaissance générique. Ils ont orientés donc vers les approches structurelles. Et cela dans le cadre du projet Cognitive Optical Recognition Old Characters COROC [ARR 02].

## **I.3. Différents aspects de la Reconnaissance Automatique de l'Écriture et du Document ou RAED**

Devant la difficulté d'établir un système universel traitant tous les cas d'écriture, la solution à court terme consiste à particulariser le cadre puis à construire l'approche la plus adaptée pour ce cadre. On peut répartir les systèmes de reconnaissance de

l'écriture dans plusieurs catégories différentes : mode d'acquisition « En-ligne » ou « Hors-ligne », Approche de reconnaissance « Globale » ou « Analytique » et enfin Méthode de reconnaissance « Statistiques » ou « Structurelle » [SAO 98].

### **I.3.1. Mode d'acquisition (En-ligne / Hors-ligne)**

Il est d'usage de distinguer deux activités en RAED, les reconnaissances *en-ligne* et *hors-ligne*. Dans le cas de la REM, il existe deux champs d'applications suivant le mode de saisie. Si le mode de saisie est dynamique, on parle de reconnaissance en temps réel appelée *en-ligne* ou *on-line*. Dans le cas de la saisie statique on parle alors de reconnaissance en temps différé appelée *hors-ligne* ou *off-line* [AYA 04].

#### **I.3.1.1. Reconnaissance en-ligne**

La reconnaissance en-ligne s'effectue à partir d'une acquisition spatio-temporelle des caractères ou mots cursifs sur une tablette électronique (version moderne de la Rand Tablet présentée sur la figure I.1).

L'écriture en ligne est obtenue lors de sa réalisation par une saisie en continu du tracé. Les données se présentent alors sous la forme d'une séquence de points ordonnés dans le temps. Dans ce cas, le signal est de type 1D (Unidimensionnel) et le système de reconnaissance peut bénéficier de la représentation temporelle. De ce fait il existe une analogie avec la reconnaissance de la parole. Il n'est donc pas surprenant de voir des chercheurs appliquer les techniques développées pour la parole à l'écriture [SAO 98]. Les applications concernées par ces systèmes de reconnaissance sont principalement les interfaces orientées stylo qui cherchent à intégrer l'écriture manuscrite comme une nouvelle modalité d'interaction entre l'homme et la machine.



*Figure I.1. Rand Tablet ou Grafacon.*

Dans les applications de REM en-ligne, l'utilisateur écrit sur une table spéciale, le système va reconnaître l'écriture et envoyer le résultat à l'ordinateur. Ces systèmes sont utilisés dans plusieurs équipements électroniques comme PDA, Pocket PC ou Tablet PC. Il y a pas mal d'avantages de REM en-ligne, par rapport à celle hors-ligne. Par exemple, car l'utilisateur écrit sur une table spéciale, il y a moins de bruit. De plus, on peut déterminer comment un caractère est écrit, c'est à dire, l'ordre de traits constituants ce caractère. D'ailleurs, la contrainte du temps de reconnaissance n'est pas stricte, on peut utiliser des algorithmes complexes. C'est pour quoi le taux de reconnaissance de ces systèmes est assez élevé [ALA 05].

### I.3.1.2. Reconnaissance hors-ligne

La reconnaissance hors-ligne, celle qui nous intéresse particulièrement privée de l'information spatio-temporelle, est plus délicate par rapport à la reconnaissance en-ligne. Elle s'applique une fois que l'écriture est sur un support papier qui est numérisé puis enregistrée sur un format d'image. Cette image contient des pixels soit de type binaire (pixels noirs et blancs) soit de type entier (image niveaux de gris), l'information est alors bidimensionnelle (2D). L'objectif n'est plus ici d'établir une interaction directe entre un scripteur et la machine, mais de traiter des documents de manière automatique.

Cette reconnaissance hors-ligne regroupe deux thèmes : la reconnaissance de caractères manuscrits isolés (numériques ou alphanumériques) et la reconnaissance de mots. Les applications actuelles visent de plus en plus la lecture automatique de documents manuscrits de type chèques bancaires, courrier postal, des formulaires,...

Dans les applications de REM hors-ligne, l'écriture de l'utilisateur est acquise par un scanner. L'entrée de ces systèmes est une image. Alors, le problème devient plus difficile. Par exemple, comment on peut enlever le bruit comme des éléments du fond d'image, comment on peut traiter le manque des traits etc. De plus, il n'y a pas d'information supplémentaire comme le cas d'en-ligne. Les applications typiques sont la classification automatique des lettres, la lecture des chèques de banque etc. [ALA 05].

En résumé la difficulté de la reconnaissance hors-ligne en trois points suivants :

- ❖ La détection de tracés : un système hors-ligne doit chercher dans l'image les tracés décrivant le message écrit. Par contre, dans les systèmes en-ligne le tracé est déjà représenté comme une suite ordonnée de point.
- ❖ L'absence d'ordre et d'information temporelle : dans le système en-ligne, les informations temporelles fournissent des connaissances précieuses sur la dynamique, la vitesse et la morphologie de l'écriture et aident à enlever en grande partie l'ambiguïté de la reconnaissance de certaines lettres.
- ❖ La variabilité de l'épaisseur du tracé : ce problème devient une contrainte supplémentaire à prendre en compte dans les systèmes hors-ligne alors qu'il ne se présente pas dans les systèmes en-ligne car, dans ce cas, les tracés sont filiformes.

Récemment sont apparus des systèmes *mixtes* capables de prendre en compte les caractéristiques statistiques et dynamiques de l'écriture. On peut classer ces systèmes en deux types [OLL 99, SOU 06] :

- ❖ Un système de reconnaissance en-ligne auquel on adjoint un système de reconnaissance hors-ligne, travaillant en parallèle sur l'image reconstituée du tracé.
- ❖ Un système de reconnaissance hors-ligne auquel on adjoint un système de reconnaissance en-ligne, travaillant en parallèle sur des caractéristiques extraites du tracé. La difficulté du problème réside dans ce cas dans la reconstitution de la temporalité du tracé.

La reconnaissance hors-ligne peut être classée en plusieurs types :

### ***1.3.1.2.1. Reconnaissance de texte ou analyse de documents***

Dans le premier cas il s'agit de reconnaître un texte de structure limitée à quelques lignes ou mots. La recherche consiste en un simple repérage des mots dans les lignes, puis à un découpage de chaque mot en caractères [BEN 99], c'est le cas traité dans notre étude.

Dans le second cas (analyse de document), il s'agit de données bien structurés dont la lecture nécessite la connaissance de la typographie et de la mise en page du document. Ici la démarche n'est plus un simple prétraitement, mais une démarche experte d'analyse de document il y'a localisation des régions, séparation des régions graphiques et photographique, étiquetage sémantique des zones textuelles à partir de modèles, détermination de l'ordre de lecture et de la structure du document [TRE 97].

L'être humain arrive avec aisance à reconnaître n'importe quel document. Cette faculté est assez basique pour un être humain ; en revanche elle pose jusqu'à ce jour encore des problèmes pour l'ordinateur. L'analyse et la reconnaissance d'images de documents englobent un ensemble de techniques informatiques avec comme but la reconstitution du contenu du document sous la forme de documents structurés, selon une forme définie par l'application en question. Les documents structurés couvrent deux catégories de documents : les documents imprimés et les documents manuscrits. Parmi les documents imprimés nous distinguons les documents à structures simples et les documents à structures complexes [KAR 06].

La reconnaissance de documents s'applique à plusieurs langues écrites. La langue latine a reçu la plus grande attention de la part de chercheurs. En revanche, malgré le nombre de personnes qui parlent la langue arabe, peu de travaux de recherche sur la reconnaissance de documents ont été consacrés à cette langue [KAR 06].

### ***1.3.1.2.2. Reconnaissance de l'imprimé ou du manuscrit***

Les approches diffèrent selon qu'il s'agisse de reconnaissance de caractères imprimés ou manuscrits. L'écriture imprimée présente des caractères bien alignés horizontalement et séparés verticalement, ce qui simplifie la phase de lecture [BEN 01b, BEN 99]. La forme des caractères est définie par un style calligraphique (fonte) qui constitue un modèle pour l'identification, par conséquent, le problème de la reconnaissance peut se concevoir au départ comme un problème de reconnaissance de caractères.

En revanche, pour le cas du manuscrit, les caractères sont souvent ligaturés et leur graphisme est inégalement proportionné provenant de la variabilité intra et inter scripteurs. Cela nécessite généralement l'emploi de techniques de délimitation spécifiques et souvent des connaissances contextuelles pour guider la lecture [FAH 01]. Dans ce cas, le problème de REM sera donc plus un problème de reconnaissance de mots ou fraction de mots qu'un problème de reconnaissance de caractères [BEL 92].

## **I.3.2. Approches de la Reconnaissance**

Il y a trois façons d'aborder la reconnaissance hors-ligne de mots manuscrits [VIN 02]. Soit le système reconnaît le mot comme une entité entière et indivisible, il s'agit d'une approche *globale* ou *holistique*. Soit il reconnaît le mot à partir de ses caractères

préalablement segmentés, il s'agit d'une approche *analytique*. Soit il n'utilise que certaines propriétés et raffine sa description du mot par rebouclage, nous parlerons alors de *systèmes basés sur la lecture humaine*.

### **I.3.2.1. Approche globale**

Les approches globales reposent sur une description du mot du même type que celles utilisées en reconnaissance de caractères. Comme le mot est une forme plus complexe que le caractère, la description contiendra plus d'informations et sera ainsi moins sensible aux changements de scripteurs. L'approche globale considère le mot comme une seule entité et le décrit indépendamment des caractères qui le constituent. Elle contourne le problème délicat de la segmentation en se basant sur une description globale de toute l'image du mot manuscrit. Cette approche présente l'avantage de garder le caractère dans son contexte avoisinant, ce qui permet une modélisation plus efficace des variations de l'écriture et des dégradations qu'elle peut subir.

L'approche globale est généralement utilisée pour la reconnaissance de vocabulaires statiques réduits tels que celui des montants littéraux de chèques bancaires [GUI 95, IMP 97, PAQ 92, SIM 92]. Elle est souvent utilisée pour d'autres tâches. Elles servent, par exemple, à réduire la taille du lexique [MAD 97] ou à valider une reconnaissance analytique [COH 94].

### **I.3.2.2. Approche analytique**

Contrairement à l'approche globale, le mot est segmenté en caractères ou en fragments morphologiques significatifs inférieurs au caractère appelés graphèmes. La reconnaissance du mot consiste à reconnaître les entités segmentés puis tendre vers une reconnaissance du mot, ce qui constitue une tâche délicate pouvant générer différents types d'erreurs [AMI 96] ; [LEC 93]. Un processus de reconnaissance selon cette approche est basé sur une alternance entre deux phases : la phase de segmentation et la phase d'identification des segments.

Deux solutions sont alors possibles : la segmentation explicite (externe) ou la segmentation implicite (interne) (voir section II.2.4). Par ailleurs, les méthodes analytiques par opposition aux méthodes globales, présentent l'avantage de pouvoir se généraliser à la reconnaissance d'un vocabulaire sans limite a priori, car le nombre de caractères est naturellement fini. De plus l'extraction des primitives est plus aisée sur un caractère que sur une chaîne de caractères [AIB 94a].

Les approches analytiques reposent sur l'identification de parties de l'image du mot comme étant des parties de l'ensemble des modèles prédéfinis et connus du classifieur. La difficulté de cette approche a été explicitée par SAYRE [SAY 73] : « pour reconnaître les lettres, il faut segmenter le tracé et pour segmenter le tracé, il faut reconnaître les lettres ». Il convient donc d'adopter une méthode itérative qui raffine la segmentation en fonction des hypothèses de classification. Il y a deux façons de résoudre ce problème. La première consiste à « disséquer » [CAS 96] le mot en graphèmes et utiliser la reconnaissance de ces graphèmes pour corriger les erreurs de segmentation. La seconde consiste à scanner l'image du mot pour y reconnaître des lettres puis à utiliser un dictionnaire afin de valider la reconnaissance.

### I.3.2.3. Approche basée sur la lecture humaine

Les systèmes basés sur la lecture humaine reposent sur le principe de la supériorité du mot (*word superiority effect*) [REI 69]. Ce principe veut qu'une lettre soit plus facile à reconnaître dans un mot que seule. Un effet secondaire de ce principe, est la capacité humaine de reconnaître un mot alors même que quelques unes de ses lettres sont inversées. Il apparaît donc que la perception de formes particulières dans un mot suffit à sa lecture.

### I.3.2.4. Avantages et Inconvénients de l'approche globale et l'analytique

Le tableau I.1 résume les principaux avantages et inconvénients de l'approche globale et analytique.

	<b>Approche globale</b>	<b>Approche analytique</b>
<b>Avantages</b>	<ul style="list-style-type: none"> <li>• La prise en compte dans le même modèle des variations au niveau de l'écriture du mot, dues par exemple à des erreurs d'orthographe ayant comme résultat des insertions, des délétions ou des substitutions locales. De même, certaines abréviations qui ne sont pas trop éloignées de la forme globale du mot peuvent être prises en compte.</li> <li>• Le contexte des caractères avoisinants est mieux modélisé. Ainsi, le modèle sera sensible aux différentes ligatures inter-lettres qui dépendent de la nature des caractères qui se succèdent dans le mot.</li> <li>• Aucune décision partielle n'est prise avant la reconnaissance finale au sens où aucune hypothèse n'est émise quant à la nature d'une partie du mot.</li> <li>• Pour un lexique réduit, le calcul est généralement rapide.</li> <li>• Elle ne nécessite pas, de par sa nature, de post-traitement lexical tel qu'une recherche dans un dictionnaire pour déterminer la validité de la chaîne de caractères reconnue.</li> </ul>	<ul style="list-style-type: none"> <li>• C'est la seule approche envisageable pour une reconnaissance dans des vocabulaires ouverts.</li> <li>• Une fois les modèles de lettres appris, elle peut s'adapter facilement à un changement de vocabulaire.</li> <li>• Au niveau pratique, peut disposer de beaucoup plus d'échantillons pour un même corpus, ce qui améliore la qualité de l'apprentissage.</li> <li>• La complexité en mémoire ne dépend pas de la taille du vocabulaire, celle du calcul croît avec le logarithme de cette taille si le dictionnaire est organisé sous forme arborescente.</li> <li>• Permet théoriquement une discrimination plus fine des mots car elle se base sur la reconnaissance des lettres qui les composent. Il est possible de récupérer l'orthographe du mot reconnu.</li> </ul>
<b>Inconvénients</b>	<ul style="list-style-type: none"> <li>• La discrimination de mots orthographiquement ou visuellement proches est très difficile. L'approche est limitée à des vocabulaires réduits allant jusqu'à quelques centaines de mots, au delà, plusieurs mots peuvent se retrouver avec une description proche ou à la limite identique.</li> <li>• Elle est coûteuse en place et temps de calcul car chaque mot a un modèle qui lui est propre; la complexité croît linéairement avec la taille du lexique.</li> <li>• L'apprentissage des modèles nécessite généralement une grande quantité d'échantillons de mots. Il est en pratique plus difficile de réunir une grande quantité de mots qu'une grande quantité de lettres.</li> <li>• Un changement de vocabulaire nécessite l'apprentissage de tous les mots nouveaux.</li> </ul>	<ul style="list-style-type: none"> <li>• Les phénomènes contextuels inter-lettres ne sont pas pris en compte.</li> <li>• La reconnaissance finale du mot se base sur la reconnaissance des lettres individuelles; si plusieurs lettres ne sont pas présentes ou trop dégradées, le mot risque d'être mal reconnu.</li> <li>• Nécessite une étape de segmentation (explicite, externe ou implicite, interne) avec les problèmes de sous- ou de sursegmentation que cela implique.</li> <li>• Pour un lexique de taille réduite, l'approche analytique requiert plus de calculs.</li> </ul>

Tableau I.1 : Avantages et inconvénients des l'approches globale et analytique.

## I.4. Complexité d'un système de RAED

D'une manière générale la complexité d'un système de RAED s'évalue suivant trois critères orthogonaux [BEL 01] :

### I.4.1. Disposition spatiale du texte

La présentation d'un texte varie globalement entre deux formats : l'écriture *contrainte* correspondant à une écriture guidée par des cadres (les formulaires par exemple) et l'écriture *non-contrainte* correspondant à une écriture guidée exclusivement par le scripteur donc extrêmement variable. Les écritures externes ou internes détachées (écriture en bâtons) sont, bien entendu, les plus aisées à traiter du fait de la séparation plus ou moins immédiate des lettres (figure I.2, extraits de la Base CEDAR5).

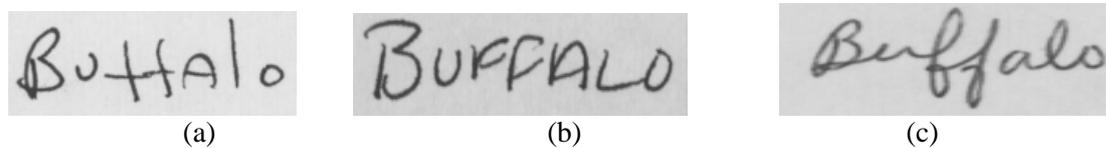


Figure I.2. Différents types d'écritures : (a) écriture en bâtons, (b) écriture majuscule, (c) écriture attachée.

Les types de contraintes peuvent caractériser l'écriture : externes correspondant aux types d'écriture et internes correspondant aux styles d'écriture propres à chaque personne. Dans ce qui suit nous exposons les différents types et styles d'écriture dans l'ordre croissant de difficulté de traitement :

#### I.4.1.1. Types d'écriture

- **Précasée** : écriture dans des cases prédéfinies (exemple : cas de bordereaux et formulaires).
- **Zonée** : écriture dans des zones bien délimitées (exemple : montant numérique dans un chèque).
- **Guidée** : écriture sur une ligne support (exemple : montant littéral d'un chèque).
- **Libre** : écriture d'emplacement quelconque (exemple : notes manuscrites).

#### I.4.1.2. Styles d'écriture

Selon la classification de Tappert en 1984 des styles d'écriture (voir figure I.3), on trouve :

- **Boxed discrete characters** : des caractères majuscules complètement disjoints, le plus souvent dans des formulaires précasés, évitant ainsi l'opération de segmentation.
- **Spaced discrete characters** : des caractères majuscules ou minuscules disjoints (script- séparé).
- **Run-on discrete characters** : le plus souvent des caractères minuscules, mais majuscules aussi qui peuvent être jointes ou disjointes rendant la reconnaissance plus complexe grâce aux liaisons fréquentes entre les lettres.

- **Pure cursive script** : écriture purement cursive avec des lettres complètement connectées, rendant ainsi la localisation des lettres difficilement dissociable de la tâche de reconnaissance.
- **Mixed cursive and discrete** : combinaison des styles précités, c'est le style d'écriture le plus utilisé vu sa simplicité et rapidité de tracé, cependant c'est le style d'écriture le plus difficile à reconnaître vu les différentes variations permises aux scripteurs.

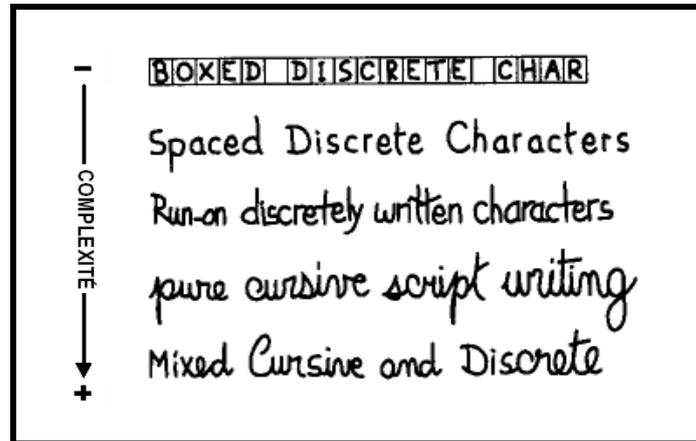


Figure I.3. Styles d'écritures selon TAPPERT.

Il est évident que la reconnaissance d'une écriture mixte naturelle est un problème beaucoup plus complexe, du fait de l'ambiguïté des limites entre les lettres, que la reconnaissance d'une écriture détachée, où la séparation des lettres est quasi immédiate.

#### I.4.2. Nombre de scripteurs

La difficulté de traitement croît avec le nombre de scripteurs. Trois catégories d'écritures se distinguent : les écritures *monoscripteurs*, *multiscripteurs* et *omniscripteurs*. En mode multiscripteurs, le système doit être capable de reconnaître l'écriture de plusieurs personnes prédéfinies, alors qu'en omniscripteur il doit s'adapter à n'importe qui.

#### I.4.3. Taille du vocabulaire

On distingue essentiellement deux types d'applications : celles qui sont à vocabulaire limité (<100 mots) et celles qui sont à vocabulaire très étendu (> 10 000 mots). La reconnaissance sera bien plus aisée dans le premier cas puisqu'il sera possible de comparer un mot inconnu avec la totalité des mots du dictionnaire. Dans le second cas il faudra bien souvent mettre en œuvre des décisions hiérarchiques pour réduire le temps de calcul et l'encombrement mémoire.

La figure I.4 présente un schéma synthétique de la complexité des systèmes RAED [BEL 01]. Plus l'application s'éloigne du centre du repère, plus la reconnaissance ne devient difficile.

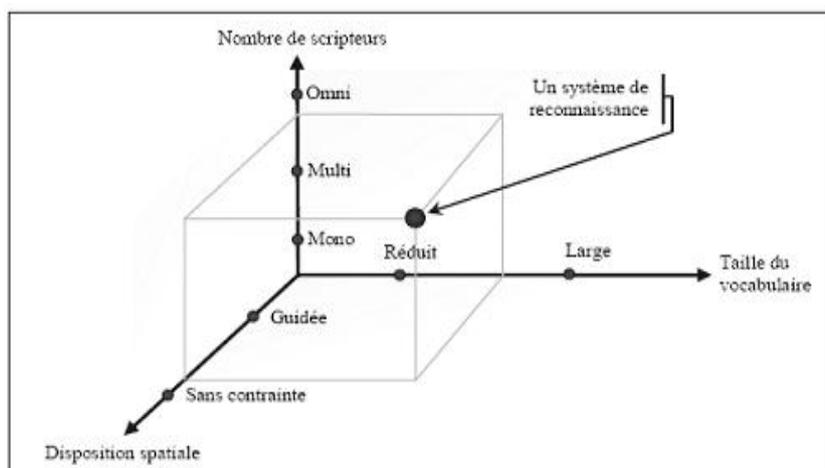


Figure I.4. Graphe de complexité des systèmes de RAED d'après BELAID [BEL 01].

Aujourd'hui les systèmes d'OCR contraints et à vocabulaire restreint donnent des résultats étonnants. Les taux de reconnaissance affichés par ce genre de logiciel avoisinent les 100% pour les documents imprimés ou mono-scripteurs de bonne qualité [ARR 02].

En attendant, d'avoir les mêmes performances pour les documents manuscrits non contraints écrits par n'importe quels scripteurs. Ainsi, pas mal de recherches sont faites pour traiter l'écriture manuscrites que se soient latines ou arabes bien que la dernière est rarement traitée [SEN 98, PAR 89, ZAD 00] etc.

## I.5. Méthodes de classification de la reconnaissance

Toutes les méthodes de la reconnaissance de formes ont été appliquées à la reconnaissance de l'écriture [BEL 92].

Cependant, les capacités de l'homme à reconnaître des objets, des personnes ou l'écriture d'une personne inconnue par exemple, quel que soit le contexte (on peut reconnaître une personne que l'on connaît de dos, ou encore lire l'écriture de tierces personnes), ce qui a conduit les chercheurs à trouver des méthodes et des techniques de reconnaissance souples et intelligentes. Parmi ces méthodes de classification on trouve :

- Les méthodes statistiques.
- Les méthodes stochastiques.
- Les méthodes linguistiques.
- Les méthodes hybrides.

### I.5.1. Méthodes statistiques

Dites aussi méthodes géométriques, se fondent sur une caractérisation des paramètres des formes étudiées (caractères, mots,...). Ces méthodes permettent de prendre une décision de classification d'une forme inconnue.

Elle repose sur une description extensive, plutôt que compréhensive, des classes [MER 04]. Bien entendu, il est assez rare que l'on puisse prévoir a priori toutes les

occurrences possibles de formes observées pour les différentes classes. On considère plus précisément qu'une classe est décrite par un ensemble représentatif d'exemplaires de formes appartenant à cette classe.

Le concept de classification peut être exprimé en termes de partitionnement d'espace de traits, où la forme est transformée en vecteur de caractéristiques. Ce dernier présente les caractéristiques mesurées. Elle est fondée sur l'étude statistique des mesures que l'on effectue sur les formes à reconnaître. L'étude de leur répartition dans un espace métrique et la caractérisation statistique des classes, permettent de prendre une décision de reconnaissance du type « plus forte probabilité d'appartenance à une classe » [BEN 99].

Les méthodes statistiques bénéficient des méthodes d'apprentissage automatique qui s'appuient sur des bases théoriques fondées, telles que la théorie de la décision bayésienne, les méthodes de classification non supervisées ... etc.

Dans le cas d'une description statistique de la forme, le problème est ramené à un exercice d'analyse de données classique. La classification est basée sur une distance (arbres de décisions, k plus proches voisins...) [JAI 00] et peut être de nature probabiliste ou neuronale [ARR 02].

La méthode statistique comprend les étapes suivantes [MER 04] :

- Choix des paramètres,
- Construction d'une description statistique des classes à partir d'un ensemble d'apprentissage,
- Calcul de la distance d'une forme à chacune des classes à partir de la description statistique précédemment élaborée,
- Choix de la classe la plus proche.

Parmi ces nombreuses théories et méthodes on peut citer à titre d'exemple:

### **I.5.1.1. Méthode bayésienne**

La méthode bayésienne consiste à choisir parmi un ensemble de caractères, celui pour lequel la suite de primitives extraites a la plus forte probabilité à posteriori par rapport aux caractères préalablement appris [ANI 92]. Elle a été élaborée dans le cadre des statistiques mathématiques, son application à la RdF a été formalisée par CHOWEN 1970 [SOU 97]. Cette technique est une méthode de base pour la RdF dans un cadre probabiliste. Qui sert de référence pour les autres méthodes. En particulier pour l'évaluation du taux d'erreur. Elle peut être formulée de la façon suivante :

Soit un problème caractériser par un ensemble de  $N$  observations:  $\mathbf{X} = \{X_1, X_2, X_3, \dots, X_n\}$  réparties  $C$  classes  $\{C_1, C_2, C_3, \dots, C_c\}$ .

On cherche à estimer la probabilité d'appartenance à posteriori de  $\mathbf{X}$  à chacune des  $C$  classes. Le théorème de Bayes fournit un cadre théorique pour la problématique de la classification. Si l'on considère un problème de classes  $C_1, \dots, C_k$ , le théorème de Bayes permet de calculer les probabilités a posteriori, connaissant les distributions des observations a priori (1.1) :

$$p(C_i / x) = \frac{p(x / C_i)P(C_i)}{P_i(x)} \quad 1.1$$

$p(C_i / x)$  est la probabilité a posteriori d'appartenir à la classe  $C_i$  connaissant le vecteur des descripteurs  $x$ ,  $p(x / C_i)$  est la densité de probabilité du vecteur d'exemples  $x$  ( $x = [x_1, x_2, \dots, x_n]$ ) dans la classe  $C_i$ ,  $P(C_i)$  est la probabilité a priori de la classe  $C_i$  et  $P_i(x)$  est la densité de probabilité non conditionnelle définie par (1.2):

$$p_i(x) = \sum_{i=1}^k p(x / C_i)P(C_i) \quad 1.2$$

Dans le cas d'un problème de classification, la formule (1.2) définit une règle de décision. La probabilité de mauvaise classification est minimisée en sélectionnant la classe qui a la plus grande probabilité a posteriori.

### **I.5.1.2. Méthode du plus proche voisin**

L'algorithme K Plus Proche Voisins KPPV affecte une forme inconnue à la classe de son plus proche voisin en le comparant aux formes stockées dans une classe de références nommée prototypes. Il renvoie les K formes les plus proches de la forme à reconnaître suivant un critère de similarité. Une stratégie de décision permet d'affecter des valeurs de confiance à chacune des classes en compétition et d'attribuer la classe la plus vraisemblable (au sens de la métrique choisie) à la forme inconnue [BEN 99, BUR 04].

Cette méthode présente l'avantage d'être facile à mettre en œuvre et fournit de bons résultats. Son principal inconvénient est lié à la faible vitesse de classification due au nombre important de distances à calculer.

### **I.5.1.3. Méthode connexionniste**

Dans un modèle statistique, la connaissance, c'est-à-dire la distribution des classes, est représentée par un modèle mathématique dont les paramètres doivent être estimés. Ces modèles constituent une limitation puisqu'ils ne seront jamais qu'une approximation de la forme des classes. Le modèle connexionniste surmonte ce problème en représentant la connaissance sous la forme d'un réseau d'unités élémentaires reliées par des arcs pondérés. C'est dans ces connexions que réside la connaissance, et celle-ci peut prendre une forme plus variée qu'avec un modèle mathématique prédéfini. Et Les nœuds de ce graphe sont des automates simples appelés neurones formels. Les neurones sont dotés d'un état interne, l'activation, par lequel ils influencent les autres neurones du réseau. Cette activité se propage dans le graphe le long d'arcs pondérés appelés liens synaptiques [AMA 96].

En OCR, les primitives extraites sur une image d'un caractère (ou de l'entité choisie) constituent les entrées du réseau. La sortie activée du réseau correspond au caractère reconnu.

Le choix de l'architecture du réseau est un compromis entre la complexité des calculs et le taux de reconnaissance [SOU 97].

Par ailleurs, le point fort des réseaux de neurones réside dans leur capacité de générer une région de décision de forme quelconque, requise par un algorithme de classification, au prix de l'intégration de couches de cellules supplémentaires dans le réseau [LIP 87].

Il existe de nombreuses topologies de réseaux de neurones, signalons que suivant la topologie et l'algorithme d'apprentissage utilisé. Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Elles peuvent être quelconques, mais le plus souvent il est possible de distinguer une certaine régularité. Ce qui donne deux structures de réseaux de neurones :

#### ***1.5.1.3.1. Réseaux non bouclés***

Les neurones sont rangés en couches distinctes où l'information se propage de couche en couche depuis la couche d'entrée jusqu'à la couche de sortie. De plus à l'intérieure d'une même couche les neurones ne sont pas connectés entre eux. L'apprentissage utilisé est supervisé. Ils sont utilisés en classification, modélisation de processus, RdF (caractères, parole,...), en prédiction,....

Les réseaux de ce type sont parfois appelés réseaux de type perceptron. Le perceptron monocouche, multicouches et le réseau Radial Basic Functions RBF en font partie. Ils sont des réseaux de type « feedforward » c'est-à-dire à propagation avant et ils sont divisés en deux familles Réseaux multicouches et Réseaux à connexions locales.

#### ***1.5.1.3.2. Réseaux bouclés***

Ce type de réseaux permet un retour en arrière de l'information. Ils sont caractérisés principalement par les rebouclées des sorties en entrées. L'apprentissage utilisé est non supervisé. Ils sont utilisés couramment pour des tâches de traitement de signal, de modélisation, de commande... Ils sont divisés en deux familles : Réseaux à connexions récurrentes, Réseaux à connexion complète.

### **1.5.2. Méthodes stochastiques**

Contrairement aux méthodes précédemment décrites, l'approche stochastique utilise un modèle pour la reconnaissance, prenant en compte la grande variabilité de la forme. La distance communément utilisée dans les techniques de « comparaison dynamique » est remplacée par des probabilités calculées de manière plus fine par apprentissage. La forme est considérée comme un signal continu observable dans le temps à différents endroits constituant des états « d'observations ». Le modèle décrit ces états à l'aide de probabilités de transitions d'états et de probabilités d'observation par état. La comparaison consiste à chercher dans ce graphe d'états, le chemin de probabilité forte correspondant à une suite d'éléments observés dans la chaîne d'entrée. [BEN 99]. Ces méthodes sont robustes et fiables du fait de l'existence d'algorithmes d'apprentissage efficaces [SEY 99]. Si l'apprentissage est lent, la reconnaissance est par contre très rapide car les modèles comprennent généralement peu d'états et le calcul est relativement immédiat. Les méthodes les plus répondues dans cette approche sont les méthodes utilisant les Modèles de Markov cachés (MMC).

Les MMC sont aujourd'hui couramment employés en RdF, dans certaines de ces approches pour la reconnaissance de l'écriture, les images des mots sont transformées en séquences de segments d'image au moyen d'une procédure de segmentation. Ces segments sont ensuite transmis à un module chargé d'estimer la probabilité selon laquelle chaque segment apparaît lorsque l'état correspondant de la chaîne de Markov est un certain état.

Par ailleurs, l'écriture arabe semi-cursive, aussi bien dans sa forme imprimée que manuscrite, se prête naturellement à une modélisation stochastique, en occurrence markovienne, à tous les niveaux de reconnaissance. Ces modèles peuvent prendre en charge le bruit et la variabilité inhérente à l'écriture manuscrite et d'éviter le problème de segmentation explicite des mots.

Cependant, l'utilisation des MMCs en reconnaissance de l'arabe est relativement limitée [BEN 01b].

On peut définir un modèle stochastique comme un processus aléatoire qui change d'état si,  $i=1, \dots, N$  au hasard aux instants  $t = 1, 2, \dots, T$ . A chaque état, il possède une variable aléatoire  $x(t) = S_i$ ,  $t \in [1 \dots T]$  qui prend ses valeurs dans l'ensemble des observations que l'on fait dans le temps du phénomène analysé [RAB 89], [BEL 92]. Une évolution du système est donc une suite de transitions d'états à partir d'un état de départ  $X(1) = S_1 : S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow \dots \rightarrow S_T$

La loi d'évolution du système consiste à avoir les chaînes  $S_1, S_2, \dots, S_m$ ,  $\forall m \leq T$  avec la probabilité  $\mathbf{P}(S_1, S_2, \dots, S_T)$ . Cette probabilité est définie de proche en proche ainsi, pour calculer  $\mathbf{P}$ , il suffit de se donner la probabilité initiale  $\mathbf{P}(S_1)$  et celles où la loi de probabilité des états à un certain instant dépend de la totalité du système. On dira que "celle-ci garde la mémoire de son passé" [BEL 92].

### **I.5.3. Méthodes linguistiques**

Ces méthodes consistent à mettre en relation la structure des formes analysées et la syntaxe d'un langage formel. La description des formes est réalisée par l'intermédiaire de phrases et le problème de classification est ramené à un problème d'analyse de grammaire (parsing). De manière générale, les méthodes linguistiques (syntaxiques ou structurelles) permettent la description de formes complexes à partir de formes élémentaires. Ces dernières, encore appelées caractéristiques, sont extraites directement des données présentes en entrée du système. La différence principale entre ces méthodes et les méthodes statistiques est que ces caractéristiques sont des formes élémentaires et non pas des mesures. Une autre différence est qu'elles introduisent la notion d'ordre dans la description d'une forme. Les méthodes les plus répandues utilisent le calcul de distance d'édition entre deux chaînes et la programmation dynamique [GRA 03].

Nous allons présenter dans ce qui suit deux variantes de la méthode linguistique, la méthode structurelle et la méthode syntaxique.

#### **I.5.3.1. Méthode structurelle**

Dans le cas d'une description structurelle, il faudra utiliser des formalismes plus complexes. Les formes sont décomposées en primitives simples qui peuvent être des graphèmes ou même les pixels de l'image. Elles sont ensuite représentées par un objet complexe, composé des primitives, comme une chaîne ou un graphe. Le processus de

reconnaissance (grammatical [GON 78], stochastique [BEL 97] ou graphique [BUN 01]) est propre à la représentation utilisée.

Les méthodes structurelles reposent sur la structure physique des caractères. Elles cherchent à trouver des éléments simples ou primitifs, et à décrire leurs relations. Les primitives sont de type topologiques telles que : une boucle, un arc... et une relation peut être la position relative d'une primitive par rapport à une autre [ANI 92, HA 96]. Parmi les méthodes structurelles nous pouvons citer :

#### ***1.5.3.1.1. Méthodes de tests***

Elles consistent à appliquer sur chaque caractère traité des tests de plus en plus fins sur la présence ou l'absence de primitives (décomposer l'image en primitives), de manière à répartir les échantillons en classes. Le processus le plus habituel consiste à diviser à chaque test l'ensemble des choix en deux jusqu'à n'obtenir qu'une seule forme correspondant au caractère entré. Ce choix dichotomique est très rapide et très simple à mettre en œuvre, mais il est très sensible aux variations du tracé [BEN 99].

#### ***1.5.3.1.2. Comparaison de chaînes***

Les caractères sont représentés par des chaînes de primitives. La comparaison du caractère traité avec le modèle de référence, consiste à mesurer la ressemblance entre les deux chaînes et à se prononcer sur celui-ci. La mesure de ressemblance peut se faire par calcul de distance ou par examen de l'inclusion de toute ou une partie d'une chaîne dans l'autre [BEN 99].

Les algorithmes à mettre en œuvre pour effectuer une classification à partir des primitives structurelles relèvent le plus souvent de l'algorithmique géométrique. On trouve ainsi typiquement [MER 04]:

- Exploration d'arbres de décision : c'est la forme la plus générale (la moins structurée) de vérification de prédicats.
- Mise en correspondance de graphes ou sous-graphes, ou encore recherche de cliques maximales dans des graphes : l'accent est ici mis non seulement sur la présence de primitives, mais sur les relations qui les lient. Une forme est décrite par un graphe encodant ces relations, et la reconnaissance s'effectue en cherchant un appariement entre le graphe décrivant une classe et celui construit à partir des primitives décrivant une forme à reconnaître.

#### **1.5.3.2. Méthode syntaxique**

Une forme apparaît comme l'assemblage d'un certain nombre d'éléments de base (primitives ou traits pertinents). La méthode syntaxique est basée sur la recherche de lois d'assemblage d'éléments de base pour former un ensemble construit qui représente la forme. Chaque caractère est représenté par une phrase dans un langage où le vocabulaire est constitué de primitives. Les caractères d'une même famille sont représentés par une grammaire.

La reconnaissance d'une forme inconnue se fait alors par une analyse syntaxique de la phrase qui la décrit. Elle consiste à déterminer si la phrase de description du caractère peut être générée par la grammaire. L'inconvénient de cette méthode est l'absence d'algorithmes efficaces pour l'inférence grammaticale directe [BEN 99].

### I.5.4. Méthodes hybrides

Il existe aussi des méthodes dites hybrides, les modèle de Markov cachés en font partie car ils emploient une approche statistique tout en ayant la possibilité d'utiliser des descriptions structurelles [GRA 03, SOU 06]. Depuis le milieu des années 90, l'utilisation des MMC pour la reconnaissance de mots latins s'est intensifiée [AVI 96, BEL 97, CHE 94, MOR 03], particulièrement dans des applications comme la lecture de chèques ou d'adresses postales pour lesquelles on dispose de bases de données importantes. L'approche par modèle discriminant en cas de petits vocabulaires; un modèle pour chaque mot, et l'approche par chemin discriminant en cas de grands vocabulaires ; un modèle pour plusieurs mots et c'est le découpage de Viterbi qui effectue la reconnaissance, en cherchant le chemin optimal dans un treillis de mots [SOU 06].

La figure 1.5 illustre les différents systèmes de reconnaissance, leurs représentations et les différentes approches de reconnaissance qu'on a citée dans le paragraphe ci-dessus.

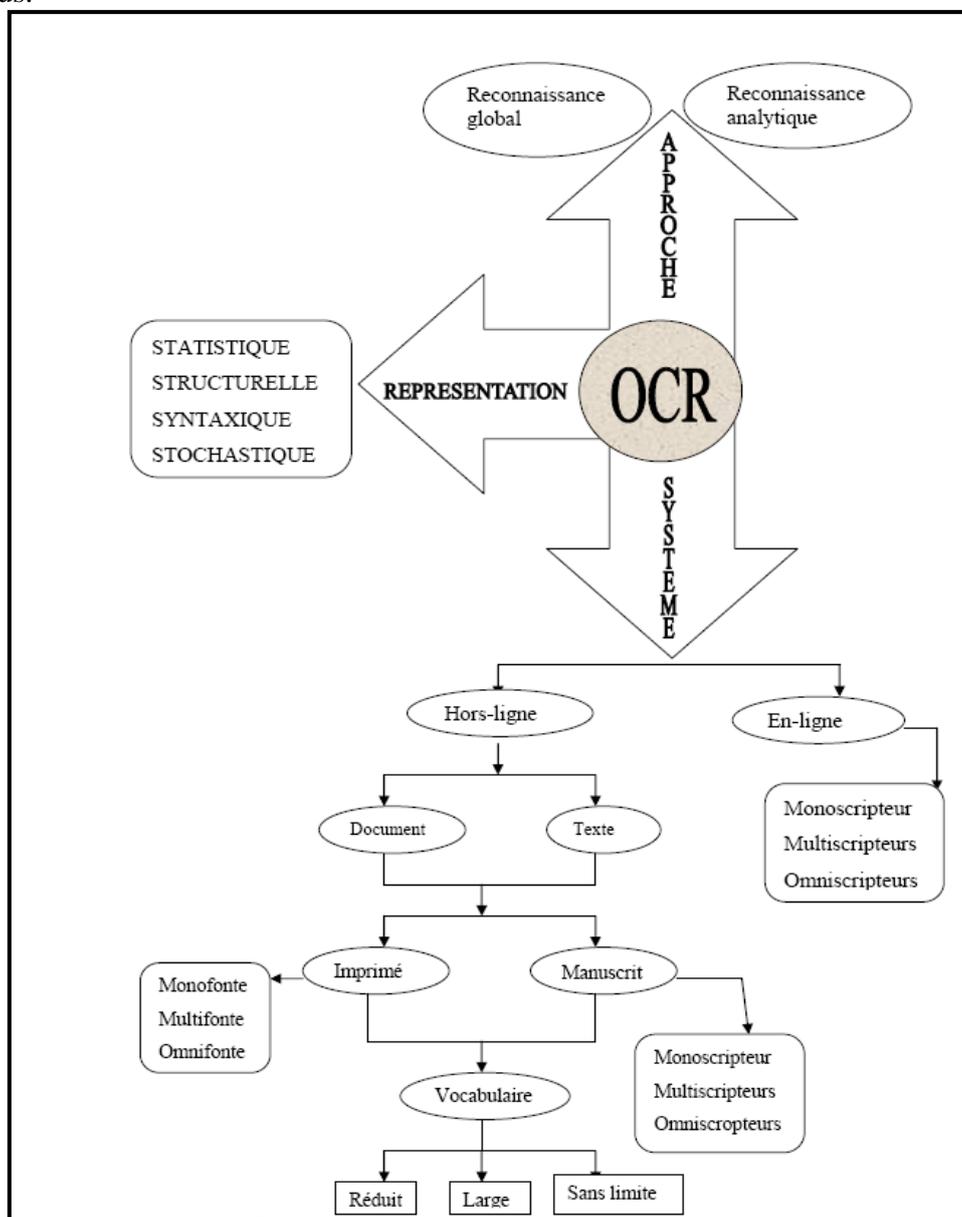


Figure I.5. Différents systèmes, représentations et approches de reconnaissance [ARR 02].

## **I.6. Conclusion**

Ce chapitre a permis de présenter notre domaine d'intérêt : la REM. Après avoir passé en revue un certain nombre de techniques utilisées pour la mise en œuvre d'un système de reconnaissance de l'écriture, nous avons présenté celui sujet de notre étude. Dans un premier temps nous avons parlé un peu sur l'historique de LAD et déterminer leur complexité puis nous avons présenté les différents aspects de la RAED. Nous avons présenté aussi certains concepts généraux liés à l'OCR, en précisant les principales méthodes de reconnaissance.

Ce domaine de recherche est devenu l'intérêt des chercheurs durant la décennie. L'arriver à concevoir un système idéal capable d'approcher les performances humaines reste jusqu'à aujourd'hui un rêve pour les chercheurs.

*CHAPITRE 2*

*PROCESSUS DE  
RECONNAISSANCE  
D'ECRITURE  
MANUSCRITE*

## II.1. Introduction

La reconnaissance de l'écriture arabe s'intègre dans le domaine de la reconnaissance de l'écriture semi cursive, avec bien sur, des spécificités et de nouveaux problèmes qui lui sont propres. Les premières recherches dans ce domaine remontent aux années 60, mais compte tenu des performances matérielles et logicielles limitées à cette époque, elles n'ont pas abouti à des résultats satisfaisantes.

En général, l'objectif de la REM 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 antérieure 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 [BEN 02].

## II.2. Organisation générale d'un SREM

La RdF se considère comme une science d'élaboration d'algorithmes permettant de classer des objets dans l'aspect varie par rapport à un objet type. Elle s'intéresse à la conception et à la réalisation des systèmes (Matériels ou logiciels) capable de percevoir à une certaine mesure d'interpréter des informations captées dans le monde physique [BEL 92].

Depuis plusieurs années, de nombreux travaux de recherche ont porté sur les systèmes de reconnaissances *hors-ligne* ont vue une progression remarquable. Ainsi plusieurs méthodes ont été développées en sorte d'aider à atteindre des systèmes satisfaisants qui sont existés dans des domaines d'applications variés tels que la lecture automatique des adresses postales, les chèques bancaires et le traitement des documents ; etc.

Il faut distinguer deux types d'écriture manuscrite hors-ligne. Dans le premier cas, chaque caractère est séparé par l'utilisateur. Dans le deuxième cas, les caractères sont écrits normalement, c'est à dire, les caractères se sont touchés. C'est le système qui doit séparer ces caractères. Le problème devient plus difficile. Dorénavant, quand on parle d'écriture manuscrite, on aborde le deuxième cas.

On s'appuie sur le schéma classique d'un processus de reconnaissance de l'écriture hors-ligne pour décrire les principaux traitements à effectuer. La construction d'un SREM comprend plusieurs étapes distinctes. Les étapes sont représentées par la figure suivante [SOU 06]. Dans un premier temps une phase d'acquisition est réalisée afin de numériser l'information manuscrite à reconnaître, suivi d'un ensemble de prétraitements permettant de réduire au maximum la variabilité intrinsèque à l'écriture ainsi que les bruits introduits lors de l'acquisition. Une seconde étape optionnelle, est celle de la segmentation. Ensuite une étape d'extraction des caractéristiques sera

réalisée ; son but est la réduction de la quantité d'information et d'extraction des caractéristiques les plus pertinentes pour la reconnaissance. Nous passons à la phase d'apprentissage. Elle permet d'estimer les différents paramètres de la modélisation choisie, à partir du corpus de données d'apprentissage. Cette étape permet d'obtenir un dictionnaire d'hypothèses possibles. Un module de post-traitement peut-être ajouté, en introduisant des informations supplémentaires et : ou complémentaires, pouvant être de différentes nature : contextuelles, lexicale, syntaxique, sémantique. Elle a pour but d'améliorer le taux de la reconnaissance.

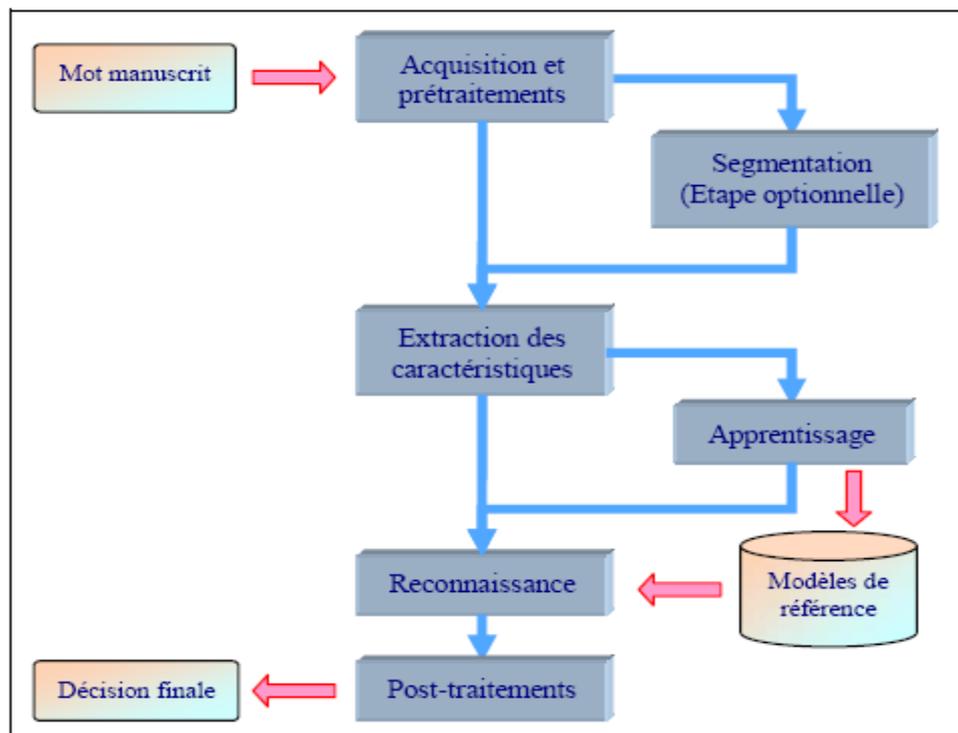


Figure II.1. Schéma général d'un système de reconnaissance de l'écriture hors-ligne.

Pour arriver à concevoir un bon SREM, il faut bien saisir et comprendre le rôle de chacune des composants qui sont illustrés dans la figure II.1. Il est bon de noter :

- qu'il n'y a pas nécessairement de composants fonctionnels dans tous les SREM. Un SREM peut prendre charge de la fonction de classification sans avoir auparavant explicitement extrait les caractéristiques au moyen, par exemple, d'une certaine forme quelconque d'appariement par référence.
- que des composants fonctionnels ne sont pas toujours mis en place à titre de composants incompatibles. Ainsi, un objet logiciel peut extraire des caractéristiques et les classifier simultanément.
- que des composants fonctionnels n'interviennent pas nécessairement en séquence. De fait, dans de nombreuses applications, un nombre important d'extractions de caractéristiques sont effectuées avant la segmentation.

### I.2.1. Phase d'Acquisition d'image

L'opération d'acquisition constitue l'interface entre le système et le monde réel. Cette étape consiste à capter l'image d'un texte au moyen des capteurs physiques (scanner,

caméra,...) et de la convertir en grandeurs numériques adaptés au système de traitement, avec un minimum de dégradation possible.

Cette étape est assez simple mais très importante car elle influence sérieusement les étapes suivantes. Il y a deux paramètres importants :

- Résolution: la résolution normale est 300 dpi. Pourtant, quand la taille de l'écriture est petite, il faut augmenter la résolution. La résolution du scanner (nombre de points par pouce : ppp ou dots per inch: dpi) désigne sa capacité à digitaliser les traits fins.
- Niveau d'éclairage : si on ajuste le scanner pour que l'image soit plus claire, le bruit est réduit mais des traits minces disparaissent aussi.

## II.2.2. Phase de Prétraitement

L'étape de prétraitement présente un niveau de complexité plus important en mode d'acquisition hors-ligne que pour le mode en-ligne. Elle a pour but de réduire au maximum les bruits dans l'image et d'éliminer autant que possible les variabilités liées au style d'écriture [BEL 92], qui peuvent provenir de différentes sources : le matériel, le scripteur ou encore le contexte [BRO 83]. En ce qui concerne le matériel, certains sont dû, d'une part au positionnement du papier sur le scanner, créant des inclinaisons, des bombages et des pliures du papier, et d'autre part à son contenu hétérogène, le processus de vieillissement fait apparaître des tâches d'humidité, la transparence de l'encre sur les rectos, la fragmentation des contours fins, ... etc. Pour toutes ces raisons la phase de prétraitement est indispensable dans un SREM.

Cette phase consiste à préparer les données issues du capteur à la phase suivante. Il inclut toutes les fonctions effectuées avant l'extraction des caractéristiques pour produire une version « nettoyée » de l'image d'origine. Il s'agit essentiellement de réduire le bruit superposé aux données et essayer de ne garder que l'information significative de la forme représentée, de corriger les erreurs, d'homogénéiser, normaliser et réduire les données [BEL 92].

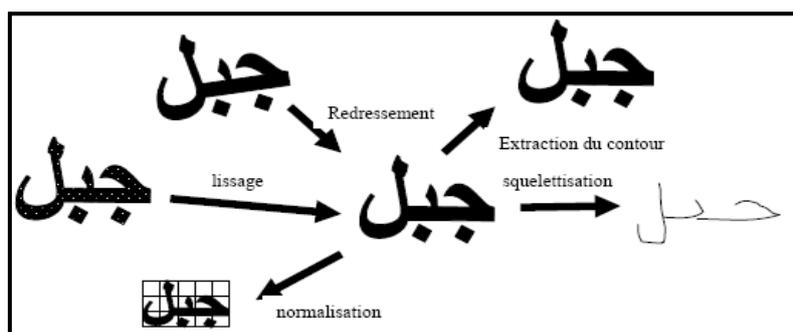


Figure II.2. Effets de certaines opérations de prétraitement.

### II.2.2.1. Réduction du bruit

Le bruit se manifeste par la présence d'informations résiduelles qui perturbent les données appartenant à la forme. Il est souvent dû aux capteurs (appareils de mesure ...). L'atténuation du bruit est un traitement préliminaire souvent difficile à effectuer car il n'est pas forcément bien localisé.

Dans le cas des images, parmi les méthodes de suppression de bruit les plus courantes, on retrouve le seuillage et le lissage.

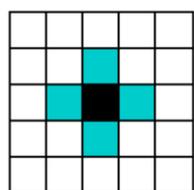
### *II.2.2.1.1. Seuillage*

L'image entrée est une image en couleur et les algorithmes de reconnaissance courants travaillent souvent sur des images binaires. Donc, il faut faire le seuillage.

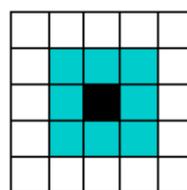
L'arrière-plan étant généralement plus clair que le tracé, la séparation la plus implicite consiste à considérer tous les pixels dont le niveau de gris est inférieur à un certain seuil comme faisant partie du tracé et vice-versa. La détermination du seuil peut être globale, c'est-à-dire que le seuil a la même valeur pour tous les pixels du document, ou locale [BEL 92], pour laquelle le seuil varie d'une position à une autre. Les méthodes de seuillage global ne sont pas trop consommatrices en terme de temps de calcul, par contre, elles ne donnent de bons résultats que si le document est uniformément éclairé. Les méthodes de seuillage local sont plus robustes à de telles dégradations mais leur temps de calcul est plus important. La méthode de seuillage local de base est celle de Niblack, elle fut ensuite améliorée par Sauvola [SAU 97]. Ces méthodes calculent le seuil de chaque pixel en fonction de la moyenne et de la variance des niveaux de gris de ces pixels voisins.

### *II.2.2.1.2. Lissage*

Il permet de réduire au maximum les discontinuités introduites dans l'image au cours des différentes transformations et ainsi de rétablir la régularité et la continuité du contour du mot. Le lissage consiste à examiner le voisinage d'un pixel et de lui attribuer la valeur 1 si le nombre de pixel noir dans cette zone est supérieur à un seuil [GRA 03]. La figure ci-dessous représente deux voisinages du pixel p :



Le voisinage en carreau  
(Quatre voisins pairs)



Le voisinage carré  
(Huit voisins)

*Figure II.3. Le voisinage en carreau et carré.*

L'image des caractères peut être entachée de bruits dus aux artefacts de l'acquisition et à la qualité du document, conduisant soit à une absence de points ou à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales qu'on appelle opérations de bouchage et de nettoyage [BUR 04].

L'opération de nettoyage permet de supprimer les petites tâches et les excroissances de la forme. Pour le bouchage il s'agit d'égaliser les contours et de boucher les trous internes à la forme du caractère en lui ajoutant des points noirs.

### II.2.2.2. Redressement de l'écriture

L'un des problèmes rencontrés en OCR est l'inclinaison des lignes du texte, qui introduit des difficultés pour la segmentation. Le redressement de l'écriture comprend généralement deux étapes : la première vise le redressement de la ligne de base du mot, la seconde le redressement des caractères. Il convient alors de le redresser afin de retrouver la structure de lignes horizontales d'une image texte.

Cette ligne de base est définie comme la ligne sur laquelle reposent les lettres ne possédant pas de dépassement bas. L'idée est de rendre horizontaux les mots à l'aide d'une transformation géométrique de type rotation isométrique des points de l'image (voir Figure II.4.a) [BEL 01]. Et cela se fait comme suit :

Si  $\alpha$  est l'angle d'inclinaison, pour redresser l'image, une rotation isométrique d'angle  $-\alpha$  est opérée grâce à la transformation linéaire suivante [STE 99]:

$$\begin{cases} x' = x \cos \alpha + y \sin \alpha \\ y' = y \cos \alpha - x \sin \alpha \end{cases}$$

*Le redressement des écritures penchées* : Cette technique facilite la segmentation préalable des mots en caractères. L'idée est de trouver l'angle moyen d'inclinaison puis de faire une transformation géométrique de type cisaillement de l'image (voir Figure II.4.b) [BEL 01].

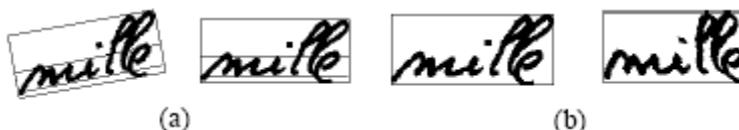


Figure II.4. Exemples de Redressement de mots manuscrits.

Pour détecter la ligne de base d'une image, on peut utiliser la méthode de la projection horizontale où la ligne de base correspondra à la ligne horizontale ayant la plus grande densité de pixels noirs.



Figure II.5. Détection de ligne de base.

Les deux opérations de redressements (ligne de base et de caractères) pouvaient être faites dans l'ordre inverse [HEU 03, SOU 06].

### II.2.2.3. Homogénéisation des données

Cette opération consiste à débarrasser la forme d'informations redondantes, superflues et inutiles pour l'application. Il existe deux principaux types de techniques :

- Amélioration de la qualité des données pour faire ressortir les détails significatifs et éliminer le reste, l'augmentation du contraste d'une image en est un exemple.
- Sélection directe d'informations pertinente appelé *filtrage*. Un filtre passe-bas, par exemple, laisse passer les basses fréquences et atténue les hautes pour écraser les contours.

#### II.2.2.4. Normalisation

Etant donné que la taille des caractères arabes est très variable, la normalisation de la taille est souvent utilisée pour échelonner les caractères à une taille fixe et pour centrer les caractères avant d'entamer toute extraction car le classificateur va effectuer plus efficacement sur les images de taille homogène. Cette opération est très utile dans les méthodes de reconnaissances qui sont sensibles aux petites variations dans la taille et la position. Mais si la taille fixée est très petite, on peut perdre d'information, si elle est très grande, l'étape de reconnaissance va opérer lentement [ALA 05].

Dans certains travaux sur la reconnaissance de l'écriture, cette opération se fait par le calcul du rectangle englobant (i.e., Bounding Box : le plus petit rectangle contenant le caractère) de chaque caractère, en appliquant des fonctions d'expansion (indépendantes) dans les directions x et y pour rendre le caractère encastré dans un rectangle de certaine taille. Pour reconnaître les différentes tailles ou estimer la largeur moyenne d'un caractère nécessaire à sa segmentation, les chercheurs mesurent la hauteur moyenne d'un caractère à partir de sa projection horizontale [AIB 95b].

Après la normalisation de la taille, les images de tous les caractères se retrouvent définies dans une matrice de même taille, Pour faciliter les traitements ultérieurs. Cette opération introduit généralement de légères déformations sur les images. Cependant certains traits caractéristiques tels que la hampe dans les caractères (ل ط ظ) par exemple peuvent être éliminés à la suite de la normalisation, ce qui peut entraîner à des confusions entre certains caractères [STE 99].

#### II.2.3. Squelettisation

La squelettisation sert à obtenir une épaisseur égale à 1 du trait d'écriture et de se ramener ainsi à une écriture linéaire. Le squelette doit préserver la forme, connexité, topologie et extrémités du tracé, et ne doit pas introduire d'éléments parasites [BEL 01].

La squelettisation du tracé d'écriture peut être utilisée pour : simplifier l'extraction de caractéristiques ou pour atténuer la variabilité inhérente aux styles d'écriture [ALM 87, ZAH 90, GOR 92, ABU 94, ABU 95, AMI 96, ABU 98, ALM 02, ALM 04, ALO 02, KHO 00, AMI 03, ALM 06, ALR 06] ; la localisation de la ligne de base et la normalisation [PEC 03, PEC 06] ; la segmentation du texte en lignes et les lignes en mots [ABU 94], [ABU 96] et la segmentation de mots en graphèmes [ALM 87, KHO 00, KHO 03] ; et la restauration des informations dynamiques dans l'écriture hors-ligne [ABU 93].

Il existe deux types d'algorithmes de squelettisation: les algorithmes séquentiels [YU 90, LAR 98, STE 99] qui s'appuient sur un ordre de balayage (les pixels sont parcourus l'un à la suite de l'autre) pour effectuer la squelettisation ; et les algorithmes parallèles [PAV 81, GUO 92, JAN 92] où le calcul effectué en un pixel

est indépendant de celui effectué sur les autres pixels, ce qui permet d'examiner en même temps tous les pixels pour générer le squelette.

## II.2.4. Phase de Segmentation

Dans cette phase les différentes parties logiques d'une image sont extraites. A partir d'une image acquise il y'a d'abord séparation des blocs de texte et des blocs graphiques, puis à partir d'un bloc de texte il y'a extraction des lignes, ensuite à partir de ces lignes sont extraits les mots puis les caractères (ou parties du caractère).

La segmentation permet d'isoler les éléments textuels, mots et caractères, pour la reconnaissance. Elle se base sur des mesures de plages blanches (interlignes et inter caractères) pour faire la séparation. La multiplicité des polices et la variation des justifications empêchent de stabiliser les seuils de séparation, conduisant à la génération de blancs inexistantes ou au contraire à l'ignorance de blancs séparateurs de mots. Ce type d'erreur est très fréquent, d'après une récente étude réalisée par Nagy & al décrite dans [BEL 01].

Il existe deux techniques permettant la mise en œuvre de la segmentation : segmentations explicite et implicite.

- **Segmentation explicite** : segmentation sur des critères topologiques, elle consiste à utiliser des points caractéristiques dans le mot [BOZ 89], tels que les minima locaux du contour supérieur, les espaces ou encore les points d'intersection. Cette approche est antérieure à la reconnaissance et n'est pas remise en cause pendant la phase de reconnaissance. Elle doit d'être d'une grande fiabilité car la moindre erreur remet en cause la totalité des traitements ultérieurs.
- **Segmentation implicite** : segmentation d'après les modèles de lettres, elle consiste à effectuer un découpage a priori de l'image en intervalles de grandeur régulière [MOH 96]. Cette technique est similaire à celle utilisée en reconnaissance de la parole, où le signal est divisé en intervalle de temps régulier. Contrairement à la segmentation explicite, il n'ya pas de pré-segmentation du mot. La segmentation s'effectue pendant la reconnaissance et est guidée par cette dernière. Le système cherche dans l'image, des composantes ou des groupements de graphèmes qui correspondent à ses classes de lettres [SOU 06].
- **Sans segmentation** : détection de la présence d'une lettre, certains systèmes se basent sur la forme globale des mots. Elles s'avèrent cependant indispensable lorsque la taille du vocabulaire est importante.

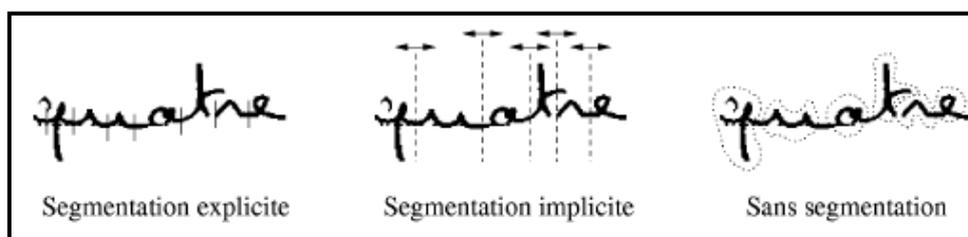


Figure II.6. Les différents types de segmentation.

## II.2.5. Phase d'Analyse ou Extraction des Caractéristiques

En REM, les données à traiter sont des images. Cette représentation de l'information n'est pas adaptée pour les systèmes de reconnaissance. Une étape d'extraction de caractéristiques est réalisée de manière à extraire l'information la plus discriminante pour la tâche de reconnaissance et également pour réduire le volume d'informations qui sera fourni au système. D'après Jain et al. [JAI 00], la performance d'un système de classification dépend fortement des relations entre le nombre d'échantillons utilisés, le nombre de caractéristiques considérées et la complexité du système.

Pour un problème de classification donné, la principale qualité recherchée pour un ensemble de caractéristiques est sa faculté de rassembler les objets appartenant à une même classe dans une partition de l'espace de représentation, tout en éloignant autant que possible les autres. Cette qualité est souvent appelée *pouvoir discriminant* de l'ensemble de caractéristiques [SOU 06]. Cette phase est une phase critique lors de la construction d'un système de reconnaissance. L'une des raisons pour laquelle cette étape pose un problème est qu'une grande majorité des techniques d'extraction s'accompagne d'une perte d'information irrémédiable. De ce fait, il faut effectuer un compromis entre quantité et qualité de l'information.

### II.2.5.1. Niveaux des caractéristiques

Les caractéristiques peuvent être extraites à partir des mots, des lettres ou des sous-lettres, donnant ainsi lieu aux trois niveaux de caractéristiques :

- **Caractéristiques de bas niveaux** : extraites à partir des sous- lettres, ayant des formes élémentaires tel que les petites lignes, les courbes, les traits, les barres,..., et des particularités géométriques simples, ce qui fait qu'elles soient très appréciées.
- **Caractéristiques de niveaux moyen** : extraites à partir des lettres, généralement utilisées dans les systèmes de reconnaissance des caractères cursifs basés sur la segmentation explicite, citons, à titre d'exemple de caractéristiques de niveau moyen, les distributions de transitions entre le fond et l'écriture.
- **Caractéristiques de haut niveau** : ce sont les caractéristiques perceptuelles, facilement visibles, consistant en la détection d'éléments structurels, elles sont indépendantes des styles d'écritures évitant ainsi le problème de la variabilité des formes. Parmi les caractéristiques de haut niveau on peut citer : les boucles, les ascendants, les descendants, en plus des jonctions, les points finaux et traits et les points diacritiques pouvant être utilisés pour trouver une représentation approximative du mot, ceci permet de se débarrasser d'une partie du lexique ou de rejeter un résultat du processus de reconnaissance dont la représentation n'est pas compatible avec celle détectée.

### II.2.5.2. Représentation des caractéristiques

Le schéma de représentation des caractéristiques varie en fonction des contraintes d'implémentation et des éventuelles stratégies utilisées.

- **Vecteur et matrice** : ce schéma est généralement utilisé pour représenter les caractéristiques de niveau bas et intermédiaire. L'image du mot est segmentée en

portions dont les caractéristiques sont extraites et représentés par des valeurs booléennes, entières ou réelles dans une matrice ou un vecteur, cette représentation est moins convenable pour représenter les caractéristiques de haut niveau.

- **Comptage** : généralement utilisé pour représenter les caractéristiques de haut niveau en calculant le nombre de caractéristiques existantes (nombre des ascendants, nombre des descendants, nombre des boucles,...).
- **Séquence** : utilisé pour représenter les caractéristiques de niveau haut et intermédiaire qui permettent d'approximer le mot par une séquence de symboles représentant un ensemble de primitives structurales.
- **Structure de graphe** : l'image est représentée par un graphe où les nœuds correspondent aux différentes caractéristiques, et les relations entre elles sont illustrées par des arcs, cette représentation graphique est très puissante car elle montre la position des caractéristiques et leurs relations.

### II.2.5.3. Types de caractéristiques

Il existe une grande diversité des caractéristiques utilisées en reconnaissance de mots. Le choix d'un type particulier de caractéristiques se révèle toujours difficile pour le concepteur d'un système de reconnaissance de mots manuscrits car les performances du système dépendent avant tout d'une bonne définition des caractéristiques. Les caractéristiques retenues doivent permettre de décrire de façon non équivoque toutes les formes appartenant à une même classe tout en les différenciant des autres classes. La difficulté du choix de ces caractéristiques réside dans le compromis à établir entre des contraintes telles que la rapidité de détection des caractéristiques, leur facilité de mise en œuvre et leur insensibilité aux distorsions (styles d'écriture différents, bruits dans l'image,...).

Dans la littérature, il existe un grand nombre de travaux concernant l'extraction de caractéristiques pour la reconnaissance de l'écriture. Afin de les présenter il est préférable de les regrouper en catégories.

Les types de caractéristiques peuvent être classés en cinq groupes principaux : caractéristiques structurales, caractéristiques statistiques, transformations globales, et superposition des modèles et corrélation [KER 00, AIB 95b], et Les caractéristiques topologiques ou métriques

#### II.2.5.3.1. Caractéristiques topologiques ou métriques

Le terme métrique désigne la mesure d'une distance. La topologie est : l'étude des propriétés de l'espace (et des ensembles) du seul point de vue qualitatif [DAR 94]. Concrètement, la topologie consiste, à l'aide de sondes appliquées directement sur l'image "brute", à effectuer par exemple sur l'échantillon les mesures et les tests suivants :

- Compter dans une forme le nombre de trous,
- Evaluer les concavités,
- Mesurer des pentes et autres paramètres de courbures et évaluer les orientations principales,
- Mesurer la longueur et l'épaisseur des traits,
- Détecter les croisements et les jonctions des traits,
- Mesurer les surfaces, les périmètres,

- Déterminer le rectangle délimitant l'échantillon, ou le polygone convexe,
- Evaluer le rapport d'élongation (ou allongement) longueur/largeur, ...
- Rendre compte de la disposition relative de ces caractéristiques.

#### ***II.2.5.3.2. Caractéristiques locales ou structurelles***

A la différence des caractéristiques topologiques, les caractéristiques structurelles sont généralement extraites non pas de l'image brute, mais à partir d'une représentation de la forme par le squelette ou par le contour [DAR 94]. Ainsi, on ne parle plus de trous, mais de boucles ou de cycles dans une représentation filiforme du caractère. Cependant, pour le reste, les caractéristiques structurelles correspondent à peu près aux caractéristiques topologiques.

Parmi ces caractéristiques on peut citer [KER 00]:

- Les traits et les anses dans les différentes directions ainsi que leurs tailles.
- Les points terminaux, jonctions et croisements.
- Les points d'intersections.
- Les boucles.
- Le nombre de points diacritiques et leur position par rapport à la ligne de base.
- Les voyellations et les zigzags (hamza).
- La hauteur et la largeur du caractère.
- La catégorie de la forme (partie primaire ou point diacritique, etc).

L'inconvénient de ces primitives est que leur extraction nécessite une squelettisation préalable du caractère, puisque l'épaisseur du trait ne contient pas d'information.

#### ***II.2.5.3.3. Caractéristiques statistiques***

Ces caractéristiques sont dérivées de la distribution des pixels. [Huette et al] suggèrent 3 familles de caractéristiques telles que : les moments invariants, les projections, et les profils. Elles sont extraites en considérant la distribution des pixels noirs de l'objet (caractère, mot, chiffres). Le processus d'identification de la meilleure méthode d'extraction de caractéristiques n'est pas évident. Par exemple, Trier et al rapportent que les moments de Zernike s'appliquent bien sur des images à niveaux de gris et que la projection s'applique souvent sur des caractères segmentés pour résoudre leur problème de REM. Cependant, il est nécessaire d'effectuer pour chaque problème de reconnaissance une évaluation expérimentale de quelques méthodes d'extraction de primitives les plus prometteuses. Ces expériences permettront de faire un choix judicieux des primitives à extraire car souvent, l'utilisation d'une seule méthode d'extraction de caractéristiques n'est pas suffisante pour obtenir de bonne performance du système de reconnaissance [BEN 04].

#### ***II.2.5.3.4. Transformations globales***

Il est intéressant pour le domaine de la reconnaissance de forme d'essayer d'extraire à partir des images, des informations non visible.

Elles sont naturellement basées sur une transformation globale de l'image. Le principe d'une caractéristique globale est de dépendre de la totalité des pixels d'une image. Elle sert à changer d'espace de représentation et ainsi faciliter l'extraction de

caractéristiques pertinentes [SOU 06]. La transformation consiste à convertir la représentation en pixels en une représentation plus abstraite pour réduire la dimension des caractères, tout en conservant le maximum d'informations sur la forme à reconnaître.

La transformée de Fourier est certainement une des plus utilisée en reconnaissance de forme et de caractères. Les caractéristiques extraites sont en fait les descripteurs de Fourier basés sur les coefficients complexes des séries de Fourier. Elles sont invariantes aux rotations et aux changements d'échelles.

Une des transformations les plus simples est celle qui représente le squelette ou le contour d'un caractère sous forme d'une chaîne de codes de directions [AIB 95b]. La chaîne de code obtenue est souvent simplifiée pour réduire les redondances et les changements brusques de direction [HAI 07].

#### ***II.2.5.3.5. Superposition des modèles (template matching) et corrélation***

La méthode de 'template matching' appliquée à une image binaire (en niveaux de gris ou squelettes), consiste à utiliser l'image de la forme comme vecteur de caractéristiques pour être comparé à un modèle (template) pixel par pixel dans la phase de reconnaissance, et une mesure de similarité est calculée [KER 00].

### **II.2.6. Phase de Classification**

Le choix du classifieur est très important. Il constitue l'élément de décision dans un système de RdF. Le but des traitements antérieurs était d'extraire l'information essentielle caractérisant les objets et de la présenter sous la forme la plus condensée possible au classifieur.

La classification dans un SREM hors-ligne consiste à déterminer la classe d'appartenance de l'objet en entrée, elle regroupe deux tâches : l'apprentissage et la reconnaissance et décision. A cette étape les caractéristiques de l'étape précédente sont utilisées pour identifier un segment de texte et l'attribuer à un modèle de référence.

#### **II.2.6.1. Apprentissage**

Cette étape permet de construire un dictionnaire de prototype. Il s'agit de regrouper en classes plusieurs prototypes dont les caractéristiques se rapprochent. Il existe 2 types d'apprentissages : supervisé et non supervisé.

##### ***II.2.6.1.1. Apprentissage supervisé***

L'apprentissage est dit supervisé si les différentes familles des formes sont connues à priori et si la tâche d'apprentissage est guidée par un superviseur ou professeur. On choisit un sous-ensemble de formes, chacune est analysée puis le professeur indique la classe dans laquelle il souhaite la voir rangée. L'apprentissage consiste alors à analyser les ressemblances entre formes de la même famille et les dissemblances entre formes de familles différentes pour en déduire les classes avec les meilleures séparatrices possibles. Les paramètres décrivant cette partition sont stockés dans une base d'apprentissage.

### ***II.2.6.1.2. Apprentissage non supervisé***

Il s'agit de construire automatiquement les classes, sans intervention de professeur, à partir d'échantillons de référence et de règles de regroupement. Ce mode nécessite un nombre élevé d'échantillons et des règles de construction précises et non contradictoire, mais n'assure pas toujours une classification correspondant à la réalité de l'utilisateur.

En fait le rôle du module d'apprentissage consiste à ajuster ses paramètres afin de donner une réponse lors de la phase de décision.

### **II.2.6.2. Reconnaissance**

Cette étape consiste à déterminer les classes d'appartenance des différents fragments extraits de l'image originale par l'étape de segmentation, cette classification s'appuie sur un certain nombre de connaissances à priori comme le type des images à reconnaître (chiffre, mot, lettre, caractère, graphème), le nombre de classes possibles. La décision de reconnaissance suppose généralement un apprentissage préalable et peut être effectuée en utilisant une multitude de méthodes de classification (Voir section I.5). Nous citons quelques techniques de classification :

- Systèmes à base de règles.
- Arbres de décision
- Techniques agglomératives
- Réseaux de neurones artificiels
- Modèles de Markov cachés

#### ***II.2.6.2.1. Les systèmes à base de règles SBR (y compris les systèmes experts)***

Utilisent habituellement les règles SI... ALORS pour établir jusqu'à quel point les conditions dans la partie SI sont conformes au modèle. Dans les SBRs, il est possible que deux règles ou plus (comportant des recommandations de classification différentes) soient applicables à la même forme d'entrée. Il en découle des conflits et il faut alors avoir recours à des mécanismes de résolution de conflits. Là encore, dans un système de reconnaissance de caractères comportant un alphabet « 0/1 », on peut s'attendre à avoir une règle telle que : SI (la forme à grande boucle), ALORS (la classe = « 0 »).

#### ***II.2.6.2.2. Les arbres de décision***

Les arbres de décision sont composés d'une structure hiérarchique en forme d'arbre. Cette structure est construite grâce à des méthodes d'apprentissage par induction à partir d'exemples. L'arbre ainsi obtenu représente une fonction qui fait la classification d'exemples, en s'appuyant sur les connaissances induites à partir d'une base d'apprentissage. En raison de cela, ils sont aussi appelés arbre d'induction (Induction Decision Trees). Une définition un peu plus formelle des arbres de décision est la suivante : un arbre de décision est un graphe orienté, sans cycles, dont les nœuds portent une question, les arcs des réponses, et les feuilles des conclusions, ou des classes terminales [OSO 98].

### II.2.6.2.3. Les techniques agglomératives

Tentent essentiellement de chercher des points dans l'espace de caractéristiques qui sont près les uns des autres et de les placer dans la même classe. Un moyen de le faire consiste à attribuer arbitrairement une classe à chaque point dans un ensemble de points non classés, puis de trouver le centre (ou la moyenne) de chaque groupe de points similaires classés. De là, chaque point est réattribué à la classe du point-centre qui s'en rapproche le plus. S'ensuit une prénumérisation des points-centres des différentes classes, et le processus est répété jusqu'à ce qu'il n'y ait plus d'attributions à faire [KHA 99].

Les RNA et les MMC sont décrites en détails dans le chapitre précédent (section I.5.1.3 et section I.5.2).

La reconnaissance peut conduire à un *succès* si la réponse est unique (un seul modèle répond à la description de la forme du caractère). Elle peut conduire à une *confusion* si la réponse est multiple (plusieurs modèles correspondent à la description). Enfin elle peut conduire à un *rejet* de la forme si aucun modèle ne correspond à sa description. Dans les deux premiers cas, la décision peut être accompagnée d'une *mesure de vraisemblance*, appelée aussi *score* ou *taux de reconnaissance*.

La validation de méthodes de reconnaissance se fait par comparaison des résultats de la reconnaissance automatique aux étiquettes données par le professeur. On en tire donc :

- Les taux de reconnaissance qui représentent le pourcentage de mots reconnus parmi les mots présentés
- Les taux d'erreurs qui représentent le pourcentage de mots acceptés par le système mais classés de façon incorrecte.
- Les taux de rejet qui correspondent à la décision de ne pas classer la forme.

Le système évaluant un critère de décision peut assigner une forme à une classe mais il peut aussi déterminer avec quelle confiance il effectue cette décision. Si le critère de décision prend des valeurs très proches pour plusieurs classes, la confiance dans la décision est faible.

Le rejet est très important dans les applications, il va être modulé en fonction d'impératifs de risque et de coût. Par exemple, il est très coûteux de faire des erreurs sur la lecture des adresses postales et encore plus sur les montants des chèques, on préférera donc augmenter les taux de rejet quitte à baisser les taux de reconnaissance.

### II.2.7. Poste-traitement

Cette étape aide à réduire considérablement des erreurs. Cependant, ce n'est pas une étape complètement séparée des étapes précédentes. Comme le processus de reconnaissance l'écriture de l'humain, l'étape de poste-traitement est intégrée strictement en ces étapes.

Le post-traitement comprend la vérification, l'exécution de l'action et l'adaptation. Cette étape peut être rajouté à un système de reconnaissance de l'écriture à pour but d'améliorer le taux de la reconnaissance, en introduisant des informations contextuelles permettant de lever l'ambiguïté dans la reconnaissance de certains mots ou caractères, parmi ces informations en citant :

- Les connaissances pragmatiques sur la longueur moyenne de chacune des lettres, ou sur le nombre de lettres constituant un mot.
- Les algorithmes de correction orthographiques ou morphologiques à l'aide de dictionnaires de digrammes, tri-grammes ou n-grammes.
- Les connaissances linguistiques quand il s'agit de la reconnaissance de phrases entières, on fait intervenir des contraintes de niveaux successifs : lexical, syntaxique ou sémantique.
- Lexical : pour valider la reconnaissance effectuée en ne retenant que des mots du dictionnaire, et en rejetant les listes de lettres inconsistantes.
- Syntaxique et sémantique : pour réduire la liste des mots candidates et valider ceux qui ont été retenus à l'étape précédente. [BEL 01, SAO 98].

Il peut être aussi envisagé si la reconnaissance ne donne pas les résultats escomptés, et sachant que le problème n'est en fait pas résultant de la reconnaissance elle-même mais d'un autre module, de l'acquisition, du prétraitement ou même de l'extraction des caractéristiques, ou ça peut bien être engendré par une base d'apprentissage pas assez complète.

### **II.3. Conclusion**

La problématique de la reconnaissance des caractères constitue un sous ensemble de la RdF, la reconnaissance des caractères étant limitée aux formes à base de texte. L'objectif de tout système de reconnaissance de caractères est de tirer automatiquement un sens d'une image à deux dimensions (ou d'une trace) d'une entrée de texte. On peut dire, cependant, que tous les systèmes de reconnaissance de caractères comportent quatre parties fonctionnelles : le prétraitement, l'extraction des caractéristiques, la classification des formes et le post-traitement.

Selon les auteurs, quelles que soient les techniques utilisées, toutes les méthodes de reconnaissance de caractères doivent relever deux importants défis : la segmentation et l'adaptation. La segmentation, ou le manque de segmentation, constitue le plus grand problème auquel sont confrontés les concepteurs qui tentent de monter un système de reconnaissance de caractères libre de toute restriction.

L'autre problème important dans la reconnaissance de caractères est l'adaptation, particulièrement en l'absence de la rétroaction directe (d'un être humain) pour redresser les erreurs. C'est à dire que l'apprentissage ne serait pas surveillé et que l'incertitude persisterait. C'est que la machine aurait à décider elle-même de l'erreur et de l'emplacement de l'erreur, une tâche qui dans les meilleures circonstances peut être ardue.

Dans ce chapitre nous avons abordé les différentes étapes nécessaires à la REM ainsi que leurs objectifs à savoir : les prétraitements, l'extraction des primitives ...etc., qui doivent être choisis d'une manière rigoureuse par le concepteur, avant de développer son système de reconnaissance. Chacune des étapes abordées a ces propres avantages et inconvénients, certaines étapes engendrent des pertes d'informations. Toutes ces difficultés font que la RdF ne soit pas un processus simple, linéaire et évident. Ces étapes doivent être bien menées pour ensuite compter sur une classification efficace.

*CHAPITRE 3*

ECRITURE  
ARABE  
ET  
L'OCR

### III.1. Origine de l'écriture arabe

Dans ce chapitre, nous présentons les caractéristiques morphologiques de l'écriture arabe. Ensuite nous exposons les principaux travaux développés en OCR arabe, tout en soulevant les problèmes majeurs rencontrés dans ce domaine.

La langue arabe est née dans la péninsule Arabique aux environs du VI<sup>e</sup> siècle. Elle est de la famille des langues sémitiques (لغات سامية). La théorie la plus communément admise relie celle-ci de l'écriture nabatéenne, l'ascendante directe de l'écriture araméenne ancienne qui elle est même issue de l'alphabet phénicien. L'arrivée de l'Islam a profondément marqué l'histoire de la langue et de l'écriture arabe [ZGH 02].

Cette langue s'est étalée très rapidement et a relié un immense empire, allant du Proche-Orient, jusqu'au l'Espagne et la Sicile, en passant par toute l'Afrique du Nord. Elle ne s'est pas limitée au texte coranique, mais elle est devenue une langue de culture, de philosophie, de sciences et de techniques. La langue Arabe est une langue officielle de 19 pays. Elle a deux types d'écritures possibles :

- l'écriture classique, pour le Coran et la littérature classique,
- l'écriture universelle du monde arabe actuel.

### III.2. Réforme de l'écriture arabe

Durant la retranscription du Coran sur un support scripturaire, il s'est avéré :

- Les 22 consonnes nabatéennes étaient insuffisantes pour écrire les 28 phonèmes arabes.
- En plus l'absence de points diacritiques et des signes vocaliques qui réduisait la graphie à un simple schéma consonantique.

Ainsi la réforme de cette écriture est devenue une affaire impérative et très urgente. C'est durant califat Abbasside1 que les pionniers de la réforme, débutèrent leurs travaux. Les plus importantes modifications apportées aux nombres de deux sont [ZGH 02]:

- Ajout des six nouvelles lettres (د, ذ, غ, ظ, ض, ث, خ).
- Ajout des points diacritiques, afin de différenciés les lettres qui partageaient une forme identique. Des petits points noirs ont été utilisés comme marques de différenciation. Les points étaient placés au dessus et au-dessous du contour de la lettre, individuellement ou par groupes de deux ou trois. Exemple : ت, ب, ث.

### III.3. Caractéristiques de l'écriture arabe

L'alphabet arabe (Tableau III.1) comprend vingt-neuf lettres fondamentales (vingt-huit si l'on exclut la hamza, qui se comporte, soit comme une lettre à part entière, soit comme un diacritique). La hamza « ء » à une orthographe spéciale qui dépend de

règles grammaticales, ce qui multiplie les formes nécessaires à sa représentation, puisqu'elle peut s'écrire seule ou sur le support de trois voyelles (alif, waw et ya). De plus l'alphabet arabe comprend d'autres caractères additionnels tels que « ؤ » et « ڤ », de ce fait, certains auteurs considèrent que l'alphabet arabe comprend plutôt 31 lettres que 29.

La considération du symbole « ~ » qui s'écrit uniquement sur le support du caractère « ڤ », fait apparaître d'autres graphismes (Tableaux III.2 et III.3). L'écriture arabe a ainsi plusieurs spécificités nous citons ci-dessous.

Caractère (forme finale)	Equivalent phonétique	Nom	Caractère (forme finale)	Equivalent phonétique	Nom
ا	a	Alif	ك	k	Kaf
ظ	d	Dad	ذ	d	Thal
ب	b	Ba'	ل	l	Lam
ط	t	Tah	ر	r	Ra'
ت	t	Ta'	م	m	Mim
ظ	z	Zah	ز	z	Zayn
ث	th	Tha'	ن	n	Nun
م	'	Yayn	هـ	a	Hin
ج	j	Jan	هـ	h	Ha'
غ	gh	Ghayn	ش	sh	Shin
ح	h	Hha'	و	w	Waw
ف	f	Fa'	ص	s	Sad
ك	kh	Kha'	ي	y	Ya'
ق	q	Qaf	ـ	.	Hamzath
د	d	Dal			

Tableau III.1. Alphabet Arabe.

caractère	initiale	médiane	finale	isolé
Alif+~			ا	آ
Alif+ء			ا	آ
			ا	آ
Waw+ء			و	ؤ
Ya+ء	ن	ن	ن	ن

Tableau III.2

caractère	initiale	médiane	finale	isolé
Lamalif			لا	لا
+ء			لا	لا
Lamalif			لا	لا
+ء			لا	لا
			لا	لا

Tableau III.3

Tableaux III.2. et III.3. Hamza et Madda et les positions qu'elles occupent en association avec Alif, Waw et Ya.

Le sens d'écriture va de droite à gauche. Il n'y a pas de différence entre les lettres manuscrites et les lettres imprimées ; les notions de lettre capitale et lettre minuscule n'existent pas. En revanche, la plupart des lettres s'attachent entre elles, même en imprimerie, et leur graphie diffère selon qu'elles sont précédées et/ou suivies d'autres

lettres ou qu'elles sont isolées (variantes contextuelles). Certaines lettres, cependant, ne s'attachent jamais à la lettre suivante : de fait, un mot unique peut être entrecoupé d'un ou plusieurs espaces, lesquels sont aussi utilisés pour séparer les mots. La longueur de cet espace inter-mot est généralement supérieure à l'espace intra-mot entre caractères non attachés.

L'alphabet arabe est un abjad, le lecteur doit connaître la structure de la langue pour restituer les voyelles. Cela se traduit par le fait que toutes ces lettres, à l'exception de l'Alef, sont des consonnes. Le Waw et le Ya sont, elles, des demi-voyelles, dans la mesure où elles représentent à la fois une consonne et une voyelle : le Waw se prononce 'w' ou 'ou' alors que le Ya se prononce 'y' ou 'i'. Ainsi, la racine trilitère 'KTB' (figure III.1) peut, selon le vocalisme, être lue 'kutub' (*les livres*), 'kataba' (*il écrivit*) ou encore 'kutiba' (*qui a été écrit*).

كتب

Figure III.1. Racine trilitère 'KTB' en arabe.

La lettre arabe change de forme selon sa position dans le mot. Elle ne s'écrit donc pas de la même manière au début, au milieu ou en fin de mot. Le Tableau III.3 illustre les différentes lettres arabes ainsi que leurs formes au début, au milieu ou en fin de mot. Il est à noter la présence des points diacritiques rattachés aux lettres dont le nombre varie de un à trois points ; Ces points peuvent se situer au-dessus ou au-dessous du caractère, mais jamais en haut et en bas simultanément. Un nombre important de mots en arabes sont composés uniquement de consonnes. Néanmoins, nous notons la présence de voyelles au sein des mots. Les voyelles se subdivisent en deux catégories : voyelles longues et brèves ; qu'on va les expliquer par la suite.

Les formes correspondantes à un même caractère, souvent appelées « formes internes », présentent parfois de sensibles différences ; dans certains cas, il est même difficile d'en déduire s'il s'agit d'une même lettre. Cependant le codage ASMO attribue un seul code pour les différentes formes d'un même caractère, contrairement au latin où le code ASCII prévoit deux codes différents pour la même lettre dans sa forme majuscule et minuscule [BEN 99].

Lettre	Isolé	For.F/M/I	Lettre	Isolé	For.F/M/I
Alif	ا	ا ا ا	Dhad	ض	ض - ض ض
Ba	ب	ب - ب - ب	Tad	ط	ط ط ط
Ta	ت	ت - ت - ت	Thad	ظ	ظ ظ ظ
Tha	ث	ث - ث - ث	Ayn	ع	ع - ع - ع
Jim	ج	ج - ج - ج	Gyn	غ	غ - غ - غ
Ha	ح	ح - ح - ح	Fa	ف	ف - ف - ف
Kha	خ	خ - خ - خ	Qaf	ق	ق - ق - ق
Dal	د	د - د - د	Kaf	ك	ك - ك - ك
Dhal	ذ	ذ - ذ - ذ	Lam	ل	ل - ل - ل
Ra	ر	ر - ر - ر	Mim	م	م - م - م
Zay	ز	ز - ز - ز	Nun	ن	ن - ن - ن
Sin	س	س - س - س	Ha	ه	ه - ه - ه
Shin	ش	ش - ش - ش	Waw	و	و - و - و
Sad	ص	ص - ص - ص	Ya	ي	ي - ي - ي

Tableau III.3. Différents formes d'un caractère arabe.

F : forme **F**inale du caractère.  
M : forme **M**édiane du caractère  
I : forme **I**nitial du caractère.

Certains caractères arabes incluent une boucle qui peut avoir différentes formes (Figure III.2).

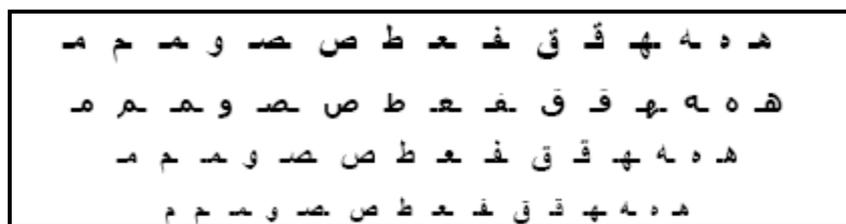


Figure III.2. Exemple de formes de boucles dans des styles différents.

Certains caractères ne peuvent être rattachés à leur gauche et de ce fait ne peuvent se trouver qu'en position isolée ou finale, ce qui donne quand ils existent, des mots composés d'une ou de plusieurs parties qu'il est convenu d'appeler généralement PAW (*piece of arabic word*) ou encore pseudo-mot. Un PAW correspond donc à une chaîne d'un ou de plusieurs caractères (voir Tableau III.4). L'écriture arabe est ainsi semi-cursive plutôt que totalement cursive.

5 PAWs /mot	4 PAWs /mot	3 PAWs /mot	2 PAWs /mot	1 PAWs /mot
السعودية, الورود	الارض, الجزائر	التاجر, الكوتر	تونس, يرد	بلد, مطر

Tableau III.4. Exemple de mots composés de la droite vers la gauche de 1, 2, 3,4 et 5 PAWs respectivement.

Comme le caractère, le PAW peut se trouver dans des mots différents à des positions différentes, mais contrairement au caractère, le PAW présente une structure morphologique stable, il garde la même calligraphie dans les différentes positions qu'il occupe (Tableau III.5).

Initiale	Médiane	Finale	Isolé
قرار	رقراق	اقر	قر

Tableau III.5. Le PAW « قر » dans différents mots et différentes positions.

### III.4. Vocalisation de l'écriture arabe

Étant donné que l'écriture arabe est une écriture purement consonantique, pour préciser la prononciation, dix signes ont été inventés. Les voyelles se subdivisent en deux catégories : voyelles longues et brèves. Il s'agit de trois voyelles brèves et de sept signes orthographiques qui s'ajoutent au-dessus et au-dessous des consonnes [ZGH 02]; de plus les trois voyelles longues.

- Les voyelles longues sont composées de trois lettres alif ( ا ), waaw ( و ) et yaa ( ي ).
- Les trois voyelles brèves sont :
  - **fatha** ( َ ) : elle surmonte la consonne et se prononce comme un « a » en français;
  - **damma** ( ُ ) : elle surmonte la consonne et se prononce comme un « u » en français ;
  - **kasra** ( ِ ) : elle se note au-dessous de la consonne et se prononce comme un « i » en français.

Les sept signes orthographiques sont les suivants :

- **sukun** ( ْ ) : ce signe indique qu'une consonne n'est pas suivie (ou mue) par une voyelle. Il est noté toujours au-dessus de la consonne ;
- les trois signes de tanwin : Nous distinguons une forme particulière d'une voyelle brève à savoir la double voyelle. Lorsque les trois voyelles brèves (fatha, kasra et damma) sont doublées, elles prennent un son nasal, comme si elles étaient suivies de " n " et on les prononce respectivement : an ( ً ) pour la **fathatan** ; in ( ٍ ) pour la **kasratan** ; un ( ٌ ) pour la **dammatan**.

À l'instar des voyelles brèves, la fathatan et la dammatan se positionnent toujours au-dessus de la consonne contrairement à la kasratan qui ne se place qu'au-dessous de la consonne ;

- **chadda** ( ّ ) : l'arabe peut renforcer une consonne quelconque. Ce renforcement est indiqué à l'aide d'un signe nommé *chadda* ou *tachdid* (intensité); La figure III.3 montre le dédoublement de lettre : *chadda* .

ل + ر + ر + ح + م + ا + ن = الرَّحْمَان

Figure III.3. Le dédoublement de lettre « chadda ».

- **wasla** ( َ ) : quand la voyelle d'un *alif* au commencement d'un mot doit être absorbée par la dernière voyelle du mot qui précède, on en indique l'élosion par le signe wasla placé au-dessus de l'*alif* ;
- **madda** ( ̣ ) : le *madda* (prolongation) se place sur l'*alif* pour indiquer que cette lettre tient lieu de deux *alifs* consécutifs ou qu'elle ne doit pas porter le *hamza*. Ce signe de contraction a la forme d'un *alif* horizontal. Le *madda* surmonte aussi les groupes de lettres exprimant une abréviation.

La figure III.4 illustre les types de voyelles de la langue arabe.

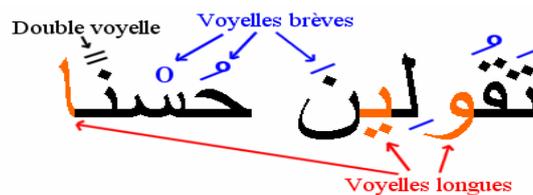


Figure III.4. Les voyelles de la langue arabe.

Les voyelles sont utilisées dans le but d'assurer qu'un texte est lu à haute voix sans erreurs de prononciation. Certains livres tels que le coran, la poésie et les livres d'apprentissage de la langue arabe utilisent les voyelles de manière systématique.

Si en français 5 signes orthographiques (les accents grave, aigu et circonflexe, le tréma et la cédille) modifient certaines lettres, en arabe toutes les formes de consonnes sont susceptible de porter chacune des huit signes de voyellation et souvent deux d'entre eux superposés (par exemple chaddah+voyelle et chaddah+tanwin). Outre ; les caractères arabes voyellés nécessitent des matrices de dimensions importantes notamment en hauteur [BEN 99].

L'écriture arabe est curviligne, composée de consonnes, dans la grande majorité et de voyelles longues, liées entre elles par des ligatures. Les mots sont constitués d'un regroupement de lettres. Chaque lettre possède une représentation selon la position dans laquelle elle se trouve : initiale, médiane ou isolée (voir Tableau III.3). La composition du mot "bienvenue" en arabe repose sur le regroupement de cinq lettres ; la figure ci-dessous illustre ce mot.

م + ر + ح + ب + ا = مرحبا

*Figure III.5. La composition des mots en langue arabe.*

Le mot "bienvenue" en arabe figurant dans la figure III.5 est composé de deux pseudo-mots, dits aussi fraction d'un mot. En effet, un mot en arabe peut comprendre un ou plusieurs pseudo-mots composés d'un certain nombre de caractères différents.

### III.5. Problèmes posés par l'écriture arabe

Parmi les caractéristiques qui distinguent le style d'écriture arabe par rapport au style lapidaire d'écriture latine et qui, par conséquent, ont engendré des difficultés lors de son introduction aux nouvelles technologies, à savoir l'imprimerie d'abord et l'informatique ensuite, nous retiendrons les caractéristiques suivantes :

#### III.5.1. Multitude de graphismes

En s'écrivant, les lettres arabes se lient les unes aux autres. Cet usage entraîne quatre morphologies différentes d'une même lettre en fonction de son emplacement dans le mot : initiale, médiane, finale et isolée à l'exception des six lettres qui ne possèdent que deux formes seulement : و, ز, ر, د, ا (voir le Tableau III.3).

#### III.5.2. Ligatures

L'écriture arabe connaît trois types de ligatures : les ligatures contextuelles, les ligatures linguistiques et les ligatures esthétiques. Une ligature contextuelle est une chaîne de caractères prenant des formes spéciales suivant leur position dans le mot en obéissant à des règles grammaticales strictes et liées uniquement à l'écriture. Les ligatures linguistiques sont indispensables pour l'écriture d'une langue donnée et obéissant à des règles grammaticales. Souvent elles ont un statut de lettre et parfois même une place à part dans le dictionnaire, ce qui les rapproche des digraphes. Les ligatures esthétiques sont des graphies optionnelles qui existent pour des raisons

esthétiques, de lisibilité et/ou de tradition. On peut les remplacer par leurs composantes sans changer la validité grammaticale, ou le sens du texte.

En arabe, les ligatures contextuelles et linguistiques sont obligatoires et indispensables pour composer un texte manuscrit, typographique ou dactylographique. Comme ligature linguistique, l'arabe ne connaît que la ligature *lâm-alif*. Bien qu'il s'agisse de la lettre *lâm* sous forme initiale suivie de la lettre *alif* sous forme finale ; لا: ا + ل = لا. Comme ligatures contextuelles on aura à titre d'exemple : لمجة = ل + م + ج + ة.

Contrairement aux ligatures contextuelles et linguistiques, les ligatures esthétiques sont optionnelles. Elles sont utilisées dans la paléographie et la typographie de qualité. Les caractères d'une même chaîne (ou pseudo-mot) sont ligaturés verticalement (dans certaines fontes deux, trois et même quatre caractères peuvent être ligaturés verticalement), occultant ainsi toute tentative de segmentation en caractères. Exemples des ligatures esthétiques au début des mots :

لمجة : Ligature esthétique entre les 2 premières lettres.

لمجة : Ligature esthétique entre les 3 premières lettres.

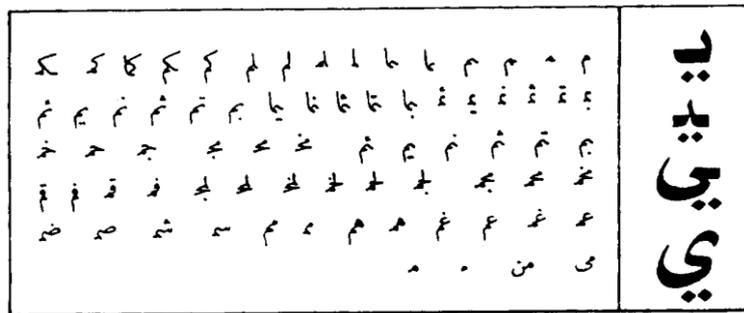


Figure III.6. Différentes combinaisons verticales possibles de caractères [BEN & al 01].

En informatique, les ligatures rendent difficile la reconnaissance optique, l'affichage et l'impression des caractères. De plus, Ce chevauchement modifie les dimensions des PAW et souvent la morphologie de certains caractères ; la ligne de base dans ce cas, n'est plus horizontale [BEN 99].

### III.5.3. Vocalisation

Pour préciser la prononciation du texte purement consonantique, les voyelles brèves et les signes orthographiques sont ajoutés au-dessus et au-dessous des consonnes (voir section III.4). Chaque forme de consonne est susceptible de porter chacun des dix signes et souvent deux d'entre eux superposés. Du point de vue technique, les ligatures esthétiques rendent pratiquement impossible la position exacte de signes de vocalisation surtout lorsqu'il s'agit d'une superposition de trois ou quatre lettres.

الَّذِينَ اتَّخَذُوا دِينَهُمْ لَهْوًا وَلَعِبًا  
وَعَرَّتْهُمْ الْحَيَاةُ الدُّنْيَا فَالْيَوْمَ نَنسَهُمْ  
كَمَا نَسُوا الْقَاءَ يَوْمَ هَذَا وَمَا كَانُوا  
بِثَانِنَا بِمُحَدِّثِينَ

Figure III.7. Ecriture voyellée.

### III.5.4. Normalisation des caractères

En s'écrivant, certaines lettres arabes montent très haut comme le *alif* surmonté du hamza ( ا ), le *kâf* ( ك ) et le *lâm* ( ل ), d'autres descendent très bas comme le *mîm* final ( م ), le *hâ* final ( خ ) ainsi que le *gîm* final ( ج ) et enfin on trouve des lettres très petites comme le *bâ* ( ب ), le *mîm* ( م ), et le *tâ* ( ت ) médianes. En typographie, pour que les petites lettres soient lisibles et distinguables, on est obligé de bien espacer les lignes. Par conséquent, on met beaucoup de mots dans une ligne mais beaucoup moins de lignes dans une page qu'avec les caractères latins. En plus l'usage des signes de vocalisation oblige d'espacer encore plus les lignes.

À ces problèmes s'ajoutent :

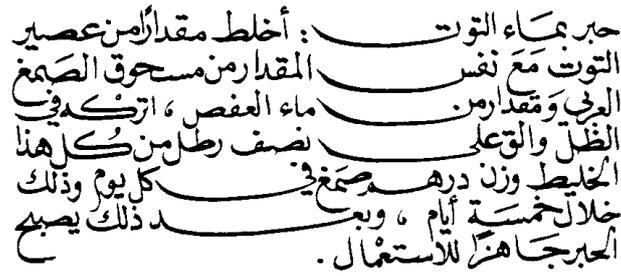
- ✓ l'écriture et la lecture de l'arabe s'effectuent de droite à gauche ;
- ✓ les formes minuscules et majuscules des lettres sont inexistantes ;
- ✓ les pays arabes font usage de deux types de graphies pour représenter les chiffres : les chiffres arabes et les chiffres hindous.
- ✓ D'autre part, le caractère arabe présente une forme cursive voyellée nécessitant, pour la majorité des lettres, des matrices de dimensions importantes. Ceci laisse jusqu'à présent les formes informatisées des caractères arabes non encore normalisées.

طَبَعَتْ هَذِهِ النسخة لِقراءة  
اللغة الفرنسية ، نَعُذِرُ للقارئ  
العربي لانه سيكون مضطراً  
لِقراءة الكتاب من جهة اليسار

بِسْمِ اللَّهِ  
قَالَ النسخ الامانة الأستاذ ابو علي النهدي رحمه  
الله تعالى هذا كتابي مختصاً به من علمي القلماً  
من الكتاب الموسوم بحمل اصول الخط بما رأيت من  
أستاذنا بفضلنا له مَكْتَبُونَ وَبِتَبَيُّنِهِ مَتَمُونَ

يعبرون على اهل الاندلس فبصرون بهم كل الاضرار  
واهل الاندلس ايضا يكابدونهم ويحاربونهم جهدهم  
الطاقة الي ان كان زمان الاسكندر و وصل الي اهل  
الاندلس فاعلموه بما هم عليه من التناكد مع اهل السوس  
فاحضر الفعلة والمهندسين وقصد مكان الزقاق وكان  
ارضا حافه فامر المهندسين بوزن الارض ووزن سطوح  
مال البحرين فوجدوا البحر الكبير يشف علوه على البحر  
الشامى بشي يسير فرفعوا البلاد التي على الساحل من  
بحر الشام ونقلها من اخفض الي ارفع ثم امر ان تحفر  
الارض التي بين بلاد طنجة وبلاد الاندلس فحفرت حجة  
وصل الحفر الي الجبال التي في اسفل الارض وبني عليها

(a) Ecriture manuscrite.



(b) Exemple d'élongations horizontales.

Figure III.8. Exemples d'écritures arabes reflétant ses caractéristiques morphologiques.

### III.6. Difficultés de la reconnaissance de l'écriture arabe

Il existe principalement deux types de difficultés dans le domaine de la reconnaissance de l'écriture arabe, le premier est lié au manque d'outils de test et de validation, alors que le second est lié à la morphologie de l'écriture arabe comme déjà mentionné au dessus (section III.5) :

- ✓ les outils de test et de validation tels que les bases de données d'images servent à valider les résultats obtenus et à évaluer de manière unifiée les divers travaux réalisés dans des environnements différents.
- ✓ D'autres, tels que les lexiques et les dictionnaires de validation sont indispensables pour améliorer les performances des systèmes de reconnaissance, suite à leur absence, le post- traitement (lexical, syntaxique ou sémantique) est très peu utilisé dans les systèmes de reconnaissance de l'écriture arabe.
- ✓ L'écriture arabe a un certain nombre de particularités telles que la forte dépendance de la calligraphie du caractère de son contexte, la complexité et la multiplicité des graphies des lettres, la variabilité des liaisons inter-caractères ou des ligatures horizontales et verticales ainsi que la présence de chevauchements. Ces particularités compliquent les tâches de choix des procédures de prétraitement, de la segmentation de textes arabes, de la sélection et de l'extraction des primitives.

### III.7. Application de la reconnaissance hors-ligne de l'écriture

La REM est le vieux rêve de tous ceux qui ont eu besoin d'entrer des données dans un ordinateur. Aujourd'hui, il existe plusieurs domaines dans lesquels la REM est attendue avec impatience, par exemple dans le tri automatique du courrier, le traitement automatique de dossiers administratifs, des formulaires d'enquêtes, ou encore l'enregistrement des chèques bancaires et postaux.

Nous présentons dans cette partie de ce chapitre quelques travaux qui ont été réalisés dans le domaine de la reconnaissance de l'écriture arabe. Nous constatons que la majorité des premiers travaux sur la reconnaissance de l'écriture arabe hors-ligne étaient consacrés aux caractères imprimés puis manuscrits [SOU 06]. Durant la dernière décennie ; plusieurs systèmes arabes ont été conçu pour reconnaître le mot en entier sans le segmenter. D'autres ont été conçus pour reconnaître les caractères

constituant le mot. En effet, une reconnaissance au niveau des caractères nécessite généralement une segmentation au préalable, un processus souvent assez délicat pour une reconnaissance fiable. Plusieurs travaux ont été réalisés dans ce domaine, nous citons quelques uns dans le tableau suivant :

Référence	Système	primitives	classification	performances
[ALM 87]	Mots hors-ligne	structurelles	Syntaxique, Comparaison de chaînes	91 %
[GOR 92]	Mots manuscrit hors-ligne	Structurelles	Syntaxique	90 %
[AME 94]	Hors-ligne, manuscrit	Structurelles	Dictionnaire (suivi de contour)	-
[ABU 94]	Caractères hors-ligne	Structurelles	Appariement de graphes flous	73.6-99.4 %
[ALB 95a]	Hors-ligne, imprimé	Structurelles	Mise en correspondance spatiale de modèle de primitives	73.13
[AMI 96]	Caractères hors-ligne	Structurelles	Réseaux de neurones MLP	92 %
[AMI 97]	Hors-ligne, imprimé, multifontes	Structurelles	Réseaux de neurones	98%
[MIL 98]	Mots hors-ligne	Structurelles, descripteurs de Fourier	HMMs combinés de façon hiérarchique	67.8-89.4 %
[SOU 99]	Hors-ligne, manuscrit	Structurelles	Dictionnaire	83.55%
[MAS 00a]	Hors-ligne, manuscrit	Structurelles	MMC	96%
[AMI 00]	Hors-ligne, imprimé, multifontes	Structurelles	Arbre de décision	92%
[AMI 00b]	Hors-ligne, imprimé, multifontes	Structurelles	Réseaux de neurones Fuzzy ARTMAP	95.25%
[KHO 00]	Hors-ligne, manuscrit	Transformation polaire suivie par transformation de Fourier	Modélisation par templates	93.2%
[ALM 02]	Mots hors-ligne	Zoning	HMMs discrets	45%
[ALO 02]	Pseudo-Mots	Zoning, Structurelles	Réseaux de neurones et HMMs discrets	69.85-73.53 %
[AZI 02]	Hors-ligne, manuscrit	Statistiques, structurelles, géométriques	Réseaux de neurones (3 types combinés en parallèle)	-
[AMI 03]	Caractères hors-ligne	structurelles	Syntaxique basée sur la logique inductive	86.65 %
[KHO 03]	Mots hors-ligne	Structurelles	HMMs discrets	72-87 %
[PEC 03]	Mots hors-ligne	Transformation Loeve Karhunen	HMMs semi- continus	89 %
[BUR 04]	Hors-ligne, manuscrit	KNN, moments	-	94%

[ALM 04a]	Mots hors-ligne	Structurelles, zoning	Syntaxique et HMMs discrets	60 %
[ELH 05]	Mots hors-ligne	Zoning, concavités	HMMs continus	74.40- 87.20 %
[FAR 05]	Hors-ligne, manuscrit	structurelles	RN ; KPPV ; KPPVFlou	94%
[SAR 05]	Mots hors-ligne	Structurelles, zoning	Réseaux de neurones MLP	84%
[SOU 06a]	Mots hors-ligne	Structurelles	Réseaux de neurones transparent, et RN neuro-symbolique	84-92%

*Tableau III.6. Tableau récapitulatif précisant les caractéristiques et les performances de certains systèmes AOCR.*

### **III.8. Conclusion**

Nous avons présenté dans ce chapitre, l'origine de l'écriture arabe et leurs principales propriétés morphologiques et typographiques.

Nous constatons que malgré, les difficultés de l'écriture arabe, on peut avoir des points de similitude avec l'écriture latine, c'est-à-dire lors de construction d'un système de reconnaissance avec une base des images de mots arabe manuscrite ou imprimée, on suit les même étapes que celle latine, sauf qu'il existe des certaines différences capitales entre eux aux niveaux de l'extraction des caractéristiques surtout. Nous avons aussi passé en revue dans la dernière partie de ce chapitre certains travaux qui ont été réalisés dans le domaine de la reconnaissance hors-ligne de l'écriture arabe. D'après ce chapitre ; nous déduisons que les problèmes majeurs dans ce domaine se ramènent à la cursivité de l'écriture et à la sensibilité de certaines caractéristiques topologiques de l'arabe ; à la variabilité des liaisons inter-caractères ou des ligatures horizontales et verticales ainsi que la présence des chevauchements. Ces difficultés rendent le choix des procédures de prétraitement, de segmentation et d'extraction des caractéristiques plus délicate.

*CHAPITRE 4*

APPROCHE GLOBALE  
POUR LA  
RECONNAISSANCE  
DE MOTS ARABES  
MANUSCRITS

### IV.1. Introduction

Dans les chapitres précédents nous avons présenté entre autres les différents aspects et les différentes approches d'un système de reconnaissance de forme, nous avons aussi discuté quelques approches et architectures concernant la reconnaissance de mots.

Nous avons consacré ce chapitre pour décrire l'architecture et les approches proposées pour mettre en œuvre un système de reconnaissance de mots arabe manuscrits dans un vocabulaire limité. Nous nous sommes basées sur l'étude des spécificités de l'écriture arabe et celles du vocabulaire considéré, notamment celui des 48 mots des willayas algériennes.

Ce chapitre est structuré en plusieurs sections, en premier lieu, nous décrivons l'architecture générale du système, qui est basé sur une combinaison parallèle entre deux approches distinctes exposées dans le présent chapitre.

Dans la deuxième section nous présentons les choix effectués et les procédures utilisées pour les étapes de prétraitement et d'extraction de caractéristiques. Ces étapes sont communes aux propositions présentées dans ce chapitre : l'approche statistique et l'approche neuronale.

La troisième section sera consacrée à la description des différentes méthodes choisies dans l'étape d'extraction de caractéristique en expliquant en détail ces méthodes, cette étape est considérée comme le noyau de notre proposition.

La dernière section, sera dédiée pour décrire les approches choisies, en expliquant le principe de fonctionnement et les opérateurs des algorithmes KPPV ; K-MEANS et FCM ainsi que l'approche neuronale PNN.

### IV.2. Architecture du système

L'objectif principal de ce travail est de concevoir et réaliser une combinaison de quatre classifieurs différents [3 statistiques et un neuronal] pour la reconnaissance hors ligne des 48 mots des willayas algérienne arabes manuscrits, multi scripteur à vocabulaire limitée. Le système proposé peut s'intégrer dans le cadre de tri automatique du courrier ou lecture d'un montant de cheque littéral algérien.

Les images sont d'abord binarisées, lissées, segmentées en mots (voir section II.2.2), puis les contours de ces mots sont extraits ainsi que leurs squelettes.

Les caractéristiques (voir section II.5) représentatives de ces mots sont extraites (d'après une image binaire, contour ou squelette) pour servir d'entrée aux classifieurs proposés.

De nombreux travaux montrent que la combinaison de classifieurs (séquentielle, parallèle ou hybride) améliore nettement les performances du système de reconnaissance par rapport à chacun des classifieurs pris séparément. Cependant l'architecture parallèle est celle qui a donné lieu aux travaux les plus importants.

Nous nous sommes focalisé dans notre travail sur la combinaison parallèle car cette dernière a prouvé son efficacité dans de nombreux problèmes de classification [ZOU 04, FAR 05, AZI 02b, NEM 09], ce succès est dû à sa simplicité de mise en œuvre, sa capacité à exploiter les réponses des différents classifieurs à combiner, en prenant en compte ou non le comportement de chacun des classifieurs. Dans cette combinaison

parallèle, chacun des ces classifieurs est responsable de connaître le mot entier ; Il doit apprendre les caractéristiques globales du mot. La différence entre eux c'est la façon de traiter les caractéristiques du mot selon le principe de fonctionnement de chacun des ces classifieurs choisis.

Cette fusion est faite soit de manière démocratique, dans le sens où elle ne favorise aucun classifieur par rapport à un autre, où bien de manière dirigée et, dans ce cas, on attribue à la réponse de chaque classifieur un poids en fonction de ses performances et ceci est en fonction généralement du taux obtenu lors de la phase d'apprentissage ; Dans ce chapitre, nous allons détailler le travail réalisé, en donnant dans un premier temps le schéma général du système proposé avant de décrire chaque étape réalisée.

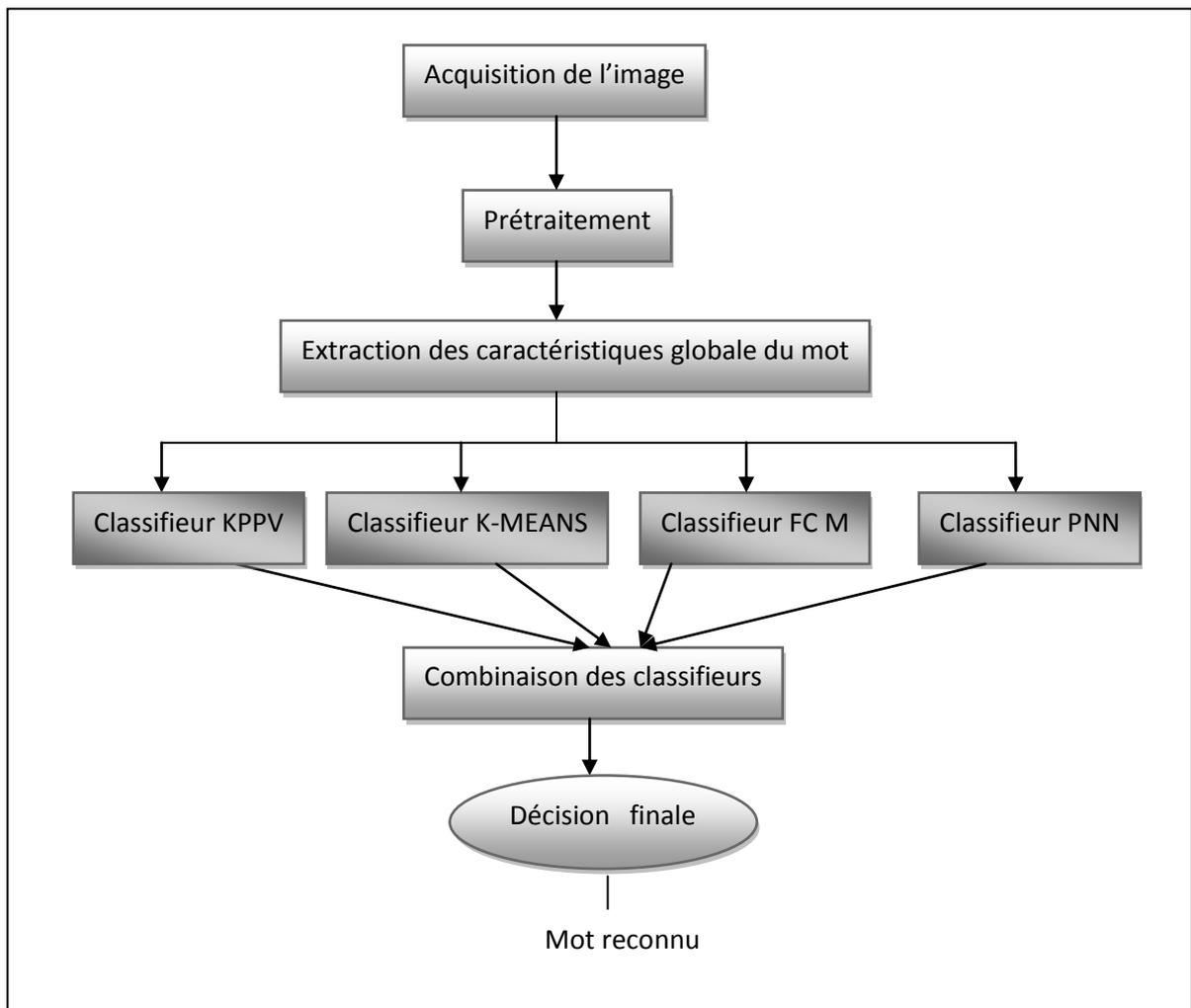


Figure IV.1. Schéma général du système de reconnaissance de mot manuscrit.

Nous avons fait le choix d'utiliser ces quatre types de classifieurs pour plusieurs raisons. Tout d'abord, chacun des classifieurs est différent de l'autre : la méthode FCM ne nécessite aucune information a priori sur les données à traiter, elle a été déjà employée avec succès dans la segmentation d'image. De plus, elle permet à une donnée d'appartenir à deux classes ou plus (une donnée à un degré d'appartenance à chaque classe). Contrairement à la méthode K-MEANS où une donnée doit appartenir ou non à une classe. Nous avons pensé d'implémenter ces deux méthodes pour

comparer les performances entre les résultats obtenus en déduisant l'importance de la logique floue dans le domaine de la classification des images (dans notre cas les images représentent des mots arabe manuscrits). La troisième méthode implémentée c'est la méthode KPPV. À l'opposition de la méthode FCM, il nécessite une base de référence (des images des mots déjà classées). De plus, ce dernier entre dans la famille des algorithmes dits « paresseux » : à l'inverse de beaucoup d'autres méthodes d'apprentissage automatique (tels que la régression logistique, les réseaux de neurones artificiels, les méthodes à noyaux, etc.) aucun apprentissage ne prend réellement place, c'est-à-dire qu'il n'y a pas de phase de détermination de paramètres d'une fonction par le biais d'une optimisation mathématique [BEL 01]. La dernière méthode choisie est la PNN qui est incluse dans la famille des réseaux de neurones, qui ont de grandes capacités de classification et qui ont montré leurs aptitudes en reconnaissance d'écriture [HAI 07, AZI 02b].

Dans la section suivante, nous allons présenter les prétraitements effectués et la phase d'extraction de caractéristiques décrite dans la section IV.7.

### IV.3. Acquisition et Prétraitement

Nous avons commencé notre travail par l'acquisition des images des mots arabes par un scanner, elles sont stockées sous l'extension .bmp ; Avant l'analyse de l'image quelques prétraitements sont effectués. Ces dernières comportent la binarisation des images numériques suivie par un lissage pour réduire le bruit issu lors de l'étape d'acquisition. La figure IV.2 illustre ces étapes.

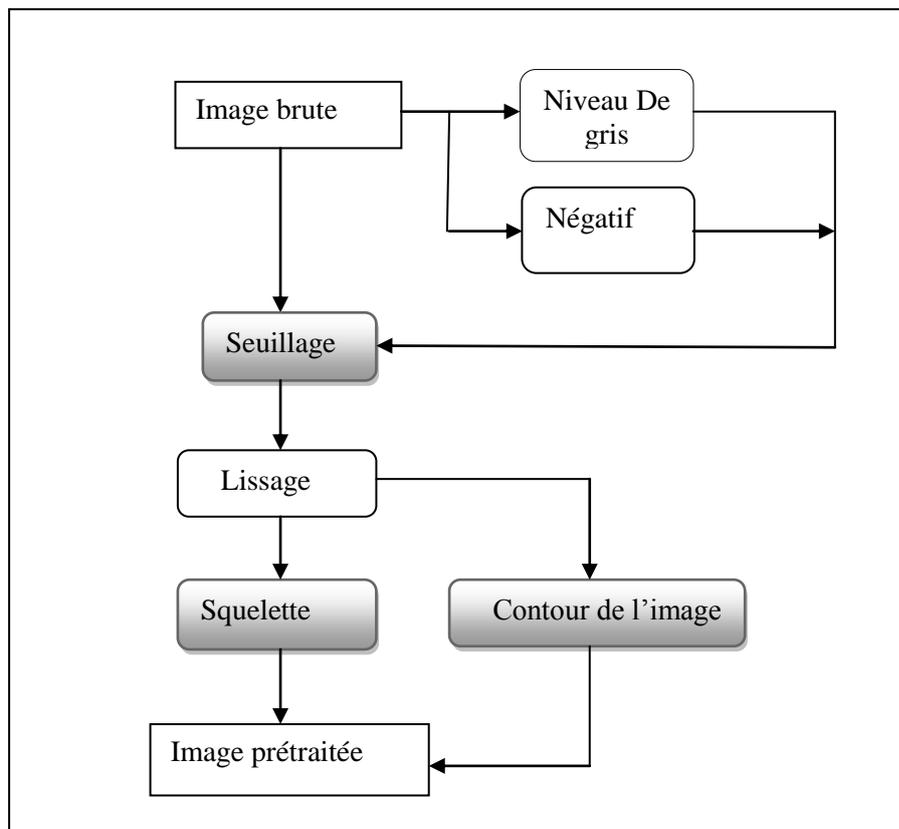


Figure IV.2. Prétraitement de l'image.

Concernant le prétraitement, un grand problème se pose ; c'est celui des boucles qui risquent d'être bouchées ou ouvertes, et aux points diacritiques qui peuvent être éliminées à la suite de certaines opérations de prétraitement, ou encore confondues avec le bruit.

En effet, le prétraitement risque d'altérer surtout la forme des points diacritiques de manière à les confondre avec le bruit s'ils sont trop amincis.

Nous avons effectué nos expérimentations sur des 48 mots de willayas algériennes écrits de manière contrainte pour pouvoir appliquer des prétraitements assez classiques. Les principales contraintes sont :

- Ecriture du mot sur une seule ligne (écriture guidée) avec peu ou pas d'inclinaison.
- Absence de fond (utilisation de fond blanc sans texture).
- Séparation claire entre les mots.

[AYA 04] présente une dichotomie des différents types d'écriture ainsi les classes de l'écriture manuscrite. Sur laquelle nous positionnons nos travaux. Il s'agit de l'écriture arabe manuscrite non contraintes (figure IV.3)

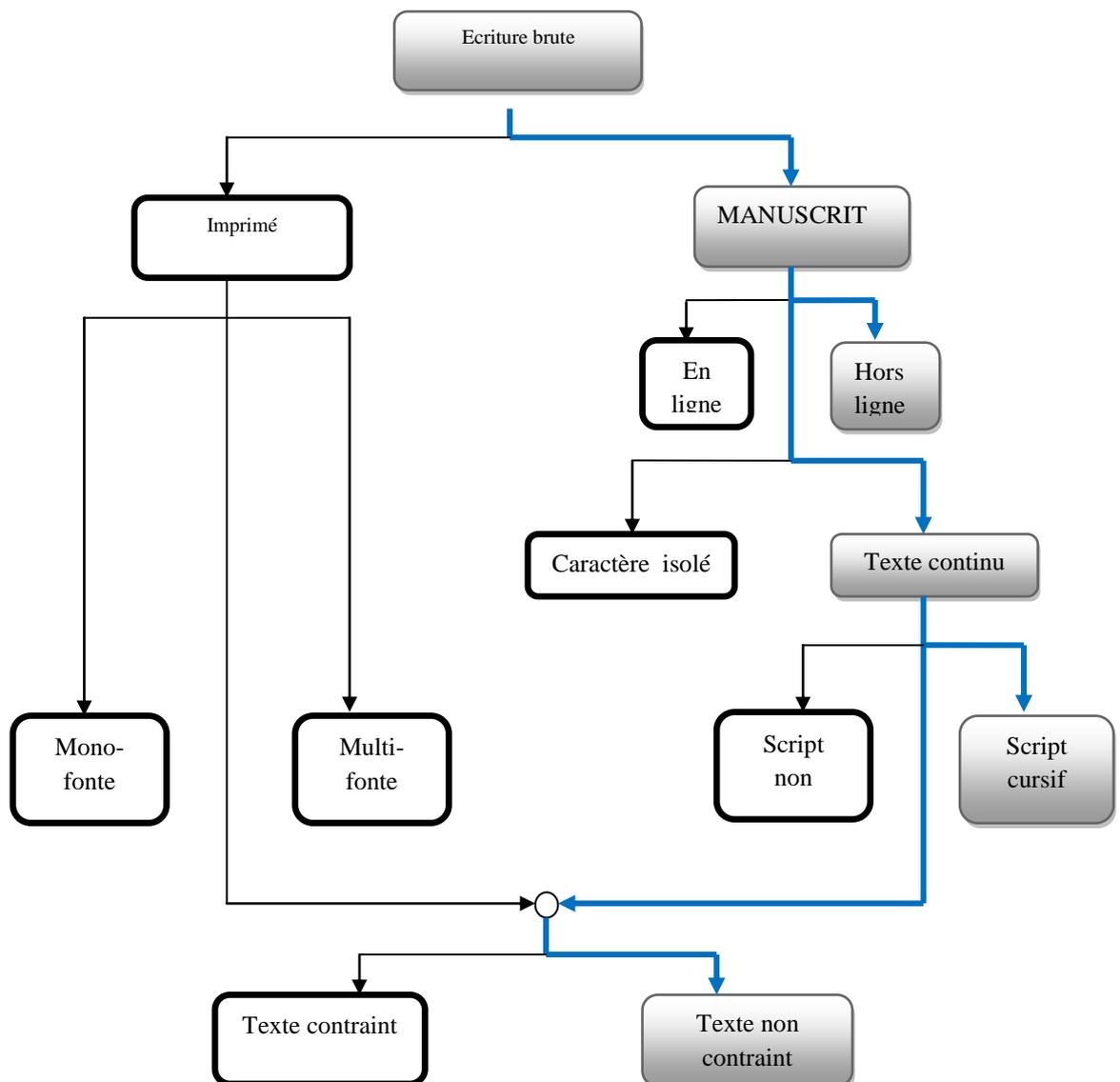


Figure IV.3. Positionnement de notre travail par rapport à la dichotomie des types d'écriture.

D'après la figure ci-dessus :  Le type d'écriture traité dans notre travail.

### IV.3.1. Binarisation

Initialement, les images sont en niveau de gris, elles sont d'abord transformées en images bimodal (**blanc et noir**). Cette méthode consiste à comparer le niveau de gris des pixels composant l'image avec un seuil à partir de l'histogramme. Dans notre cas nous avons choisi d'utiliser un seuil dynamique spécifié pour chaque image déduit à partir de son histogramme [BEN 00, OTS 79, PAV 82].

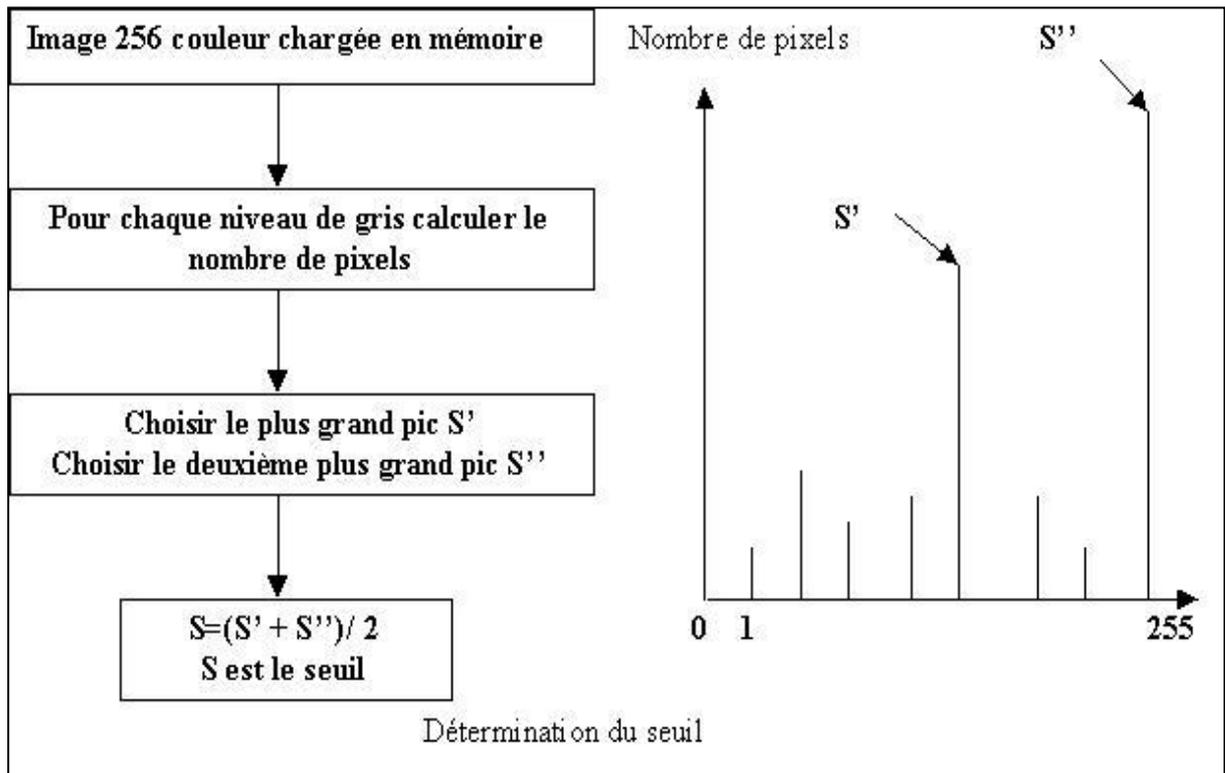


Figure IV.4. Détermination du seuil de binarisation.

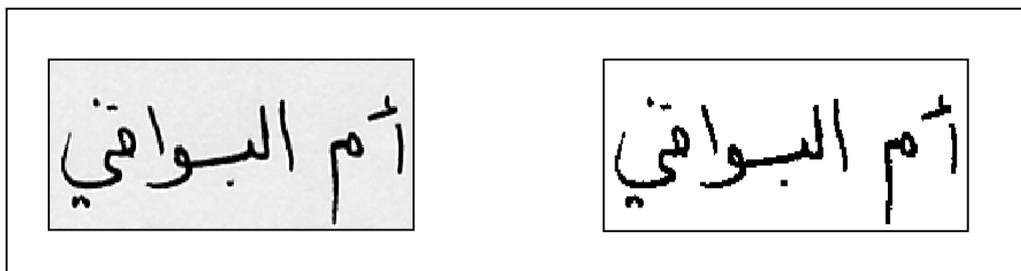


Figure IV.5. Résultat d'une binarisation.

### IV.3.2. Lissage

L'image de mots peut contenir du bruit dû aux conditions d'acquisition, par exemple. Ce bruit peut se manifester soit par les absences de points (trous), soit par des excroissances ou des empâtements. Le lissage est une transformation locale qui ne tient pas compte que du voisinage de chaque point de l'image. La transformation la plus fréquemment utilisée pour le lissage des textes hors-ligne, consiste à parcourir l'image pixel par pixel, au moyen d'un masque de dimensions 3\*3 et à modifier éventuellement la valeur du pixel examiné en fonction des ses 8 pixels voisins (nord, nord est, est, etc.). Ce masque peut être appliqué, par exemple, pour le bouchage des trous et l'élimination des pixels isolés. Les techniques de lissage permettent de résoudre ces problèmes par deux opérations :

#### IV.3.2.1. Bouchage

Cette opération consiste en un épaississement du tracé du mot pour mieux extraire les caractéristiques. Il s'agit d'égaliser les contours et de boucher les trous internes à la forme du mot en lui ajoutant des points noirs [BEL 92, SAR 00].

#### IV.3.2.2. Nettoyage

A pour but d'éliminer les petites taches qui n'a pas pu être éliminé par l'opération de lissage et les excroissances de la forme. Elle est réalisée de différentes manières suivant le type de bruit à enlever. L'algorithme qu'on a utilisé pour cette étape est le suivant :

```
Algorithme          Lissage_Image ;  
// Soit couleur_ P la couleur de P (pixel) de l'image avant lissage ;  
Pour tout Pixel P de l'image Faire  
Nbre_Voisins_Noirs ← le nombre de d-voisins noirs du P ;  
Si ( Nbre_Voisins_Noirs < 2) Alors couleur_ P ← blanc ;  
Sinon Si ( Nbre_Voisins_Noirs > 2) Alors couleur_ P ← noir ;  
      Sinon couleur_ P ne change pas ;  
FinPour ;  
Fin.
```

La Figure IV.6 montre le résultat de l'application de cet algorithme.

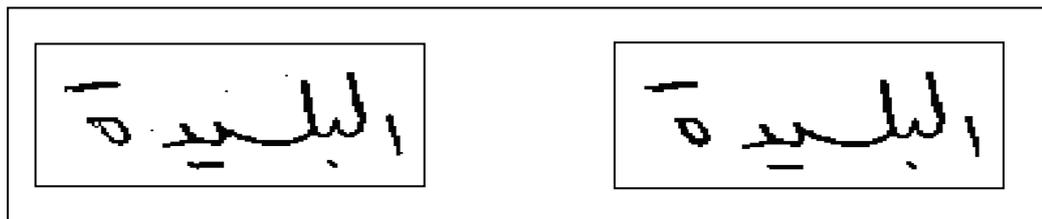


Figure IV.6 : Résultat de lissage.

## IV.4. Segmentation de textes (Traitement d'images de haut niveau)

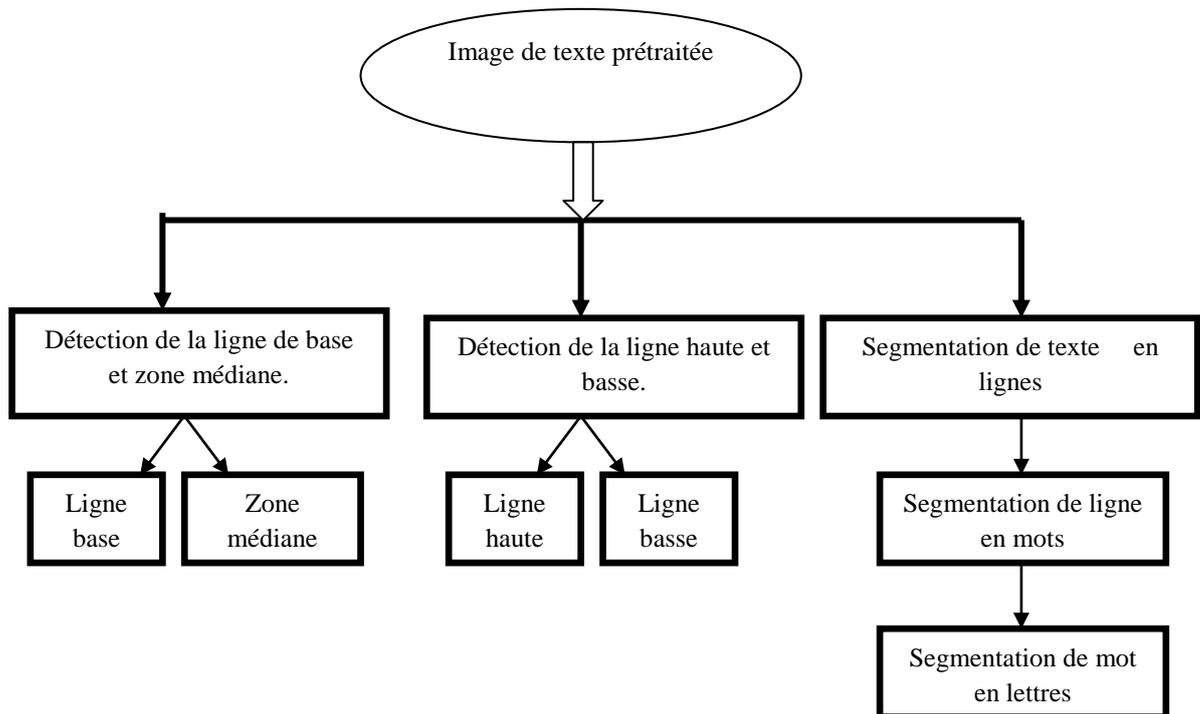


Figure IV.7. Traitement d'image d' haut niveau.

### IV.4.1. Localisation de la ligne de base et la zone médiane

Dans notre système, nous supposons que l'écriture des mots présente peu ou pas d'inclinaison (nous n'avons pas traité le problème de l'inclinaison).

#### IV.4.1.1. Localisation de la ligne de base

Dans l'écriture arabe, la ligne de base porte une information assez importante qui nous permet d'extraire facilement les différentes caractéristiques structurelles des mots. L'une des raisons pour laquelle cette étape pose un problème est qu'une localisation erronée de la ligne de base s'accompagne d'une perte d'information irrémédiable [BEN 00, SOU 06].

Il y a de nombreux méthodes pour la localisation de la ligne de base et certaines de ces méthodes sont spécifiques à la détection des inclinaisons et aux redressements des mots. La méthode que nous avons choisie pour détecter la ligne de base est celle de la ligne horizontale ayant la plus grande densité de pixels noirs dans le mot [GOW 90, TOL 90], cette méthode donne de bons résultats dans le cas où les mots sont parfaitement horizontaux ou qu'ils ont une inclinaison légère. L'algorithme de la localisation de la ligne de base de cette méthode nécessite le calcul de l'histogramme horizontal de l'image du mot binaire et lissée. Le calcul de l'histogramme horizontal se fait par une projection horizontale du mot de la manière suivante:

**Algorithme                  Calcul\_Histo\_Horz ;**

Initialiser un tableau Histo\_Horz à 0, a comme dimension le nombre de ligne dans l'image ;

Pour i Allant de 1 à nombre ligne de l'image Faire

Pour j Allant de 1 à nombre colonne de l'image Faire

    Si ( couleur\_pixel(i,j) = noir ) Alors

        incrémenter Histo\_Horz[j] de 1;

    FinPour;

FinPour;

**Fin.**

La ligne de base correspond à l'indice k du Histo\_Horz tel que Histo\_Horz[k] a la plus grande valeur.

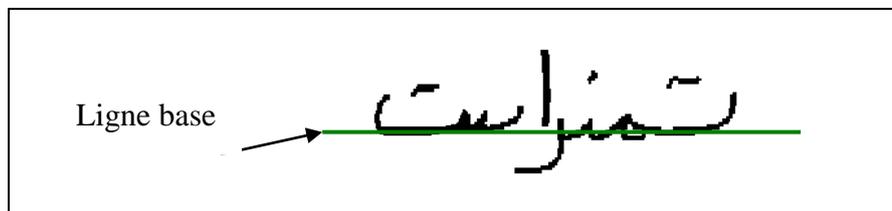


Figure IV.8. Ligne de base du mot « Tamenraset ».

**IV.4.1.2. Localisation de la zone médiane**

La zone médiane peut être définie comme le corps principal du mot qui englobe les lettres qui n'ont ni ascendant, ni descendant, voir Figure (IV.9). La localisation de la zone médiane nous a permis la détection des extensions (hautes et basses) dans l'écriture.

L'idée qu'on a utilisé pour localiser cette zone est d'estimer sa hauteur (en pixels) par rapport à la ligne de base. Nous avons réalisé une analyse statistique de sa hauteur par rapport à la ligne de base sur un ensemble d'échantillons qui contient 100 mots écrits par différents scripteurs, et qui sont homogènes, c'est-à-dire que l'ensemble des images a été scanné avec la même résolution. Les résultats qu'on a obtenus montrent qu'il existe une certaine stabilité pour l'ensemble des échantillons.

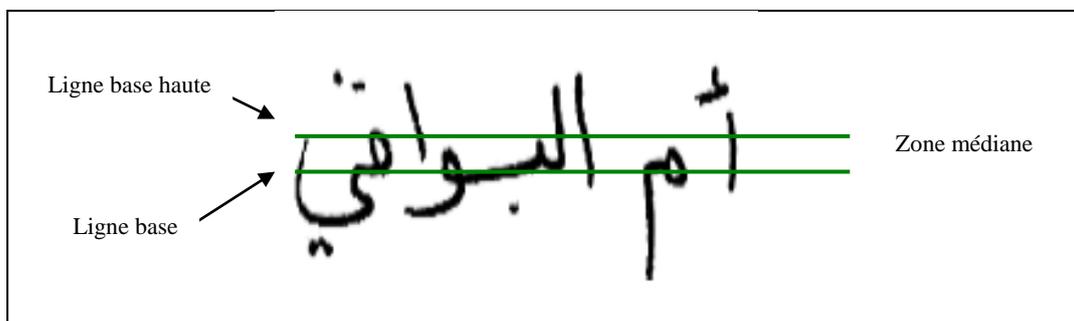


Figure IV.9. Zone médiane du mot « Oum-el-bouaghi ».

En effet, pour la plupart des échantillons la hauteur de la zone médiane est comprise entre 12 et 16 pixels, bien sûr qu'il n'existe pas une estimation parfaite, à cause de la grande variabilité de l'écriture et l'épaisseur du trait de l'écriture.

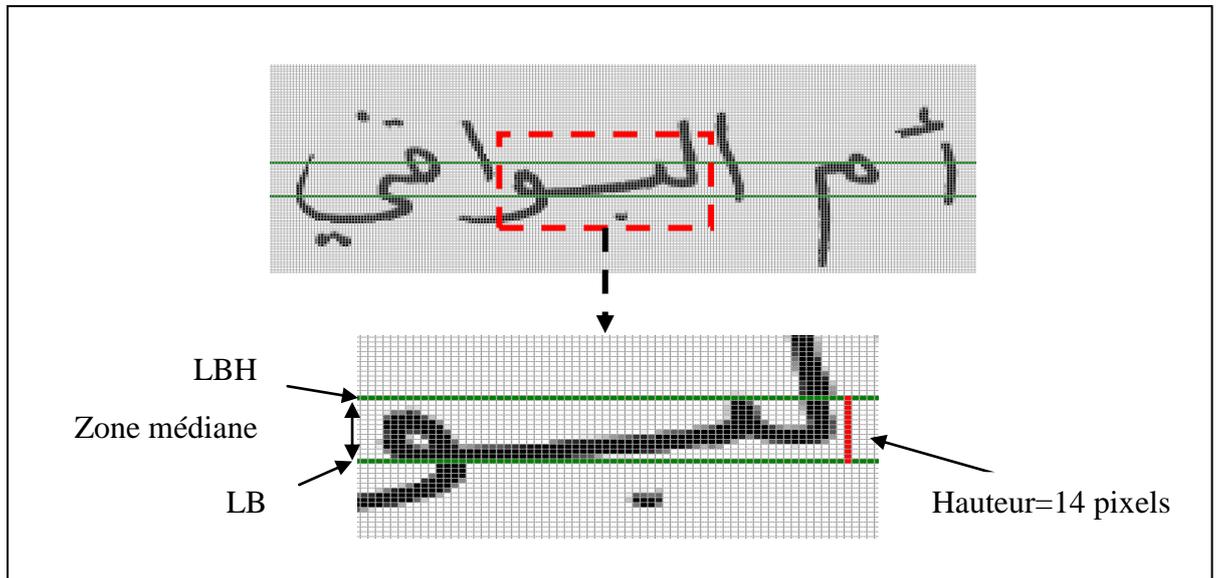


Figure IV.10. Illustration de la hauteur de la zone médiane en pixels dans un mot échantillon.

Donc après avoir localiser la ligne de base, la zone médiane peut être cernée de la manière suivante :

Soit LB : Ligne de base,

Et soit LBH : Ligne base haute, voir Figure (IV.10),

$LBH = LB - \text{seuil}$  ; où  $\text{seuil} \in [12, 16]$ .

Enfin, le corps du mot ou la zone médiane est localisée entre LB et LBH.

#### IV.4.2. Localisation de la zone supérieure et la zone inférieure

##### IV.4.2.1. Localisation de la zone supérieure

La zone supérieure est la zone qui contient les dépassements hauts dans un mot. Elle nous permet de détecter les hampes et les points diacritiques hauts.

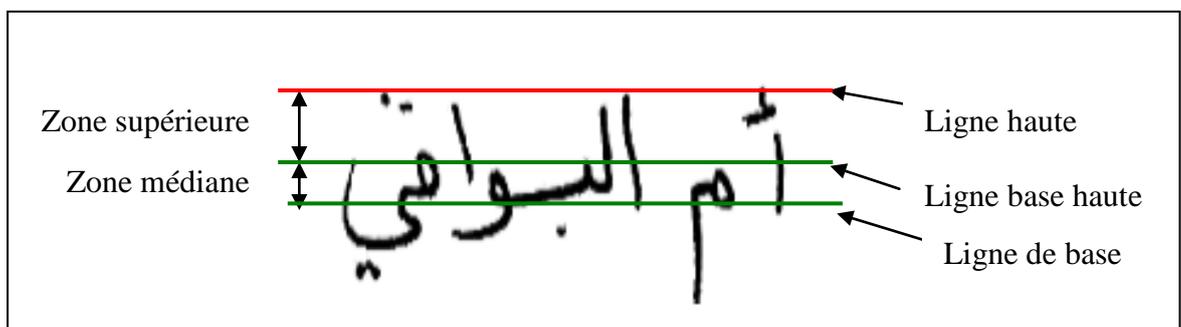


Figure IV.11. Zone supérieure du mot « Oum-el-bouaghi ».

Pour localiser cette zone il suffit de détecter la ligne haute, voir Figure (IV.11). Etant donné que l'image est binaire et lissée et qui ne contient pas des bruits (pixels noirs qui n'appartiennent pas à la forme (le mot à traiter). Nous balayons l'image ligne par ligne, colonne par colonne et de haut en bas, dès que nous trouvons le premier pixel noir nous considérons sa ligne comme une ligne haute.

Enfin la zone supérieure est cernée entre la ligne haute et la ligne base haute.

#### IV.4.2.2. Localisation de la zone inférieure

La zone inférieure est la zone qui contient les dépassements bas dans un mot. Elle nous permet de détecter les jambages et les points diacritiques bas.

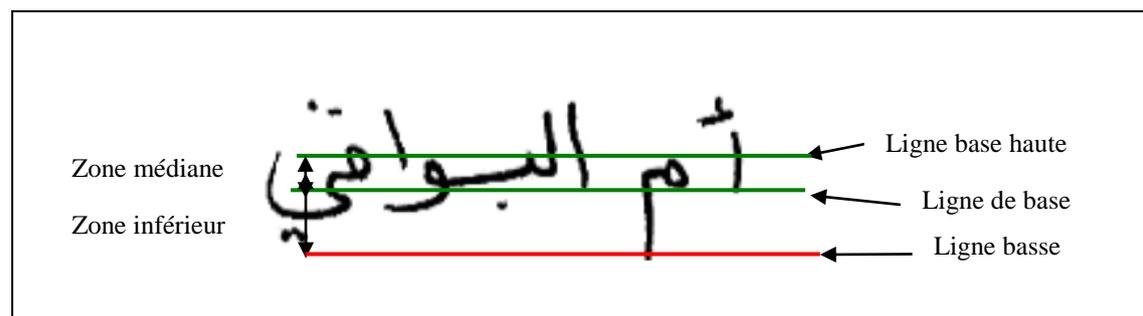


Figure IV.12. Zone inférieure du mot « Oum-el-bouaghi ».

Le principe de localisation de la zone inférieure que nous avons utilisé, est le même que pour la localisation de la zone supérieure, il suffit de détecter la ligne basse, voir Figure (IV.12). Nous supposons que l'image est binaire et lissée et qui ne contient pas des bruits (pixels noirs qui n'appartiennent pas à la forme). nous balayons l'image ligne par ligne, colonne par colonne mais cette fois ci de bas en haut, dès que nous trouvons le premier pixel noir nous considérons sa ligne comme une ligne basse.

Nous obtenons alors la zone inférieure qui est cernée entre la ligne de base et la ligne basse.

Il se peut qu'il existe des mots qui ne possèdent pas des zones supérieures et/ou des zones inférieures, ce sont les mots qui n'ayant pas des ascendants et des points diacritiques hauts et/ou des descendants et des points diacritiques bas.

#### IV.4.3. Détection des composantes connexes

La segmentation consiste à détecter les différentes composantes connexes CCs (groupe de pixels noirs liés) d'une image et représenter chacune avec une couleur différente.

Un mot arabe se compose généralement d'une ou plusieurs CCs, chacune contient un ou plusieurs caractères. Par exemple dans le mot de « Oum-el-bouaghi » nous trouvons six CCs. Mais, pour un premier temps nous considérons chaque point diacritique dans le mot comme une CC, puisqu'il constitue un ensemble de pixels qui forment une connexité. C'est dans l'étape d'analyse que nous allons les déterminer comme des points diacritiques appartenant à une CC. La figure suivante montre les différentes CCs dans le mot de « Oum-el-bouaghi ». Dans notre cas nous n'utilisons pas les CCs, nous traitons le mot entier.

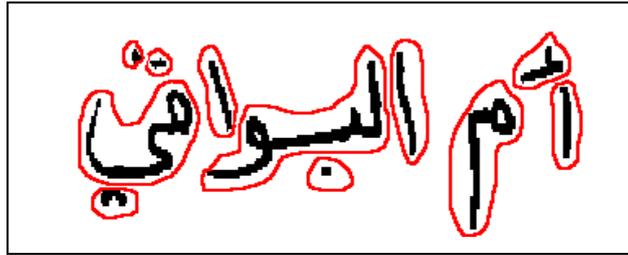


Figure IV.13. Les 11 composantes connexes qui forment le mot « Oum-el-bouaghi ».

L'idée que nous avons utilisée est l'étiquetage de chaque CC dans le mot. L'étiquetage consiste à attribuer un label, étiquette ou numéro différent pour chaque CC et identique pour tous les pixels d'une même composante. Il existe de nombreux algorithmes réalisant cette fonction. Ils dépendent de la connexité considérée et se différencient aussi par leurs approches séquentielles ou parallèles [DAR 94].

Une CC peut alors être extraite par l'intermédiaire de son étiquette pour faire un traitement spécifique. La figure illustre le résultat de l'étiquetage, chaque CC correspond une couleur (non noire) dans l'image de droite.

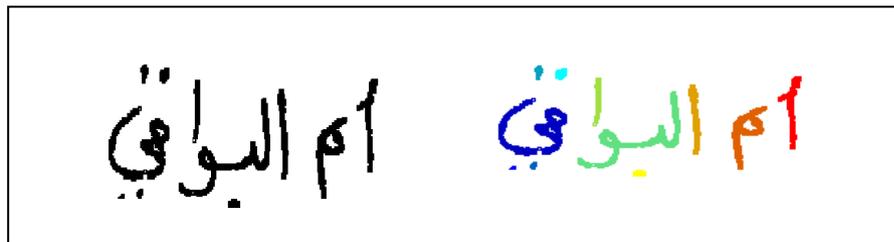


Figure IV.14. Illustration du résultat de l'extraction des composantes connexes.

## IV.5. Squelettisation

Un algorithme de squelettisation efficace doit permettre la réduction de la quantité de données et la préservation uniquement des caractéristiques essentielles de la forme, de plus, il doit avoir les propriétés suivantes :

- 1) Préservation de la connexité de la forme.
- 2) Convergence vers des squelettes d'épaisseur unitaire.
- 3) Approximation de l'axe médian de la forme.
- 4) Réduction maximale de données.

La conception des algorithmes de squelettisation dédiés à certains types d'écriture comme le latin, chinois, etc. était un domaine de recherche très actif, et la littérature en est très abondante [LAM 92]. En revanche, les algorithmes destinés spécifiquement à la squelettisation de l'écriture arabe restent relativement peu nombreux [MAH 91, ALT 95, ALP 97]. Notons que l'application directe des algorithmes de squelettisation conçus notamment pour l'écriture latine ou chinoise sur l'écriture arabe ne permet pas d'obtenir des squelettes adaptés, du fait de la présence de marques diacritiques, et la squelettisation de ces marques par ces algorithmes peut les faire disparaître ou les déformer totalement. Par exemple, un point diacritique comme dans la lettre "ب" peut se transformer après squelettisation en un petit trait, ce qui n'est plus adapté pour l'écriture arabe manuscrite, car il est très fréquent que certains scripteurs utilisent un petit trait à la place de deux points comme dans la lettre

" ي". A cause de ces problèmes, les marques diacritiques sont parfois extraites de l'image d'écriture avant la squelettisation puis sont classifiées séparément [ZAH 90], [BEN 07]

Dans notre cas, La méthode utilisée consiste à appliquer des amincissements successifs sur l'image binarisée jusqu'à ce que le mot résultant ne contienne que des lignes d'épaisseur 1 pixel. La méthode nécessite l'emploi successif de 8 masques.

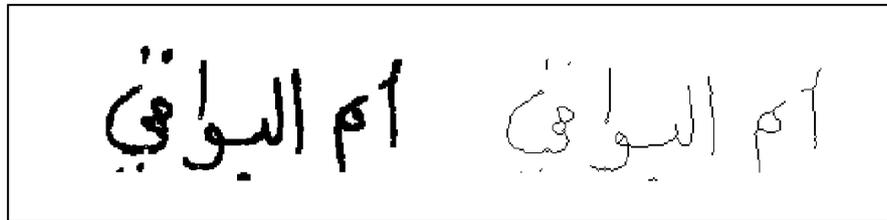


Figure IV.15. Résultat de squelettisation.

#### IV.6. Traçage de contours

Le contour, bordure ou *i-contour*, d'un ensemble connecté de pixels  $R$  est défini comment étant l'ensemble de tous les pixels de  $R$  qui possèdent au moi un  $d$ -Neighbors n'appartenant pas à  $R$ .

Le *d-contour* de  $R$  est l'ensemble de pixels de  $R$  qui possèdent au moins un voisin n'appartenant pas à  $R$  [PAV 82].

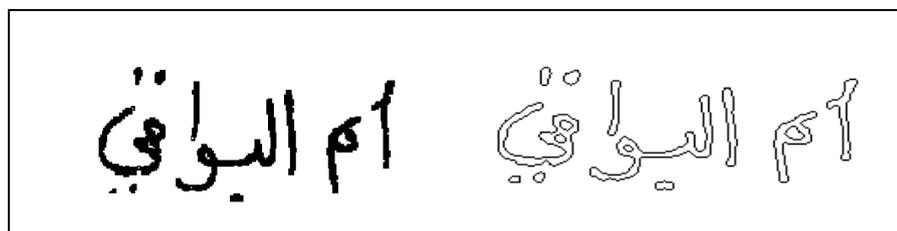


Figure IV.16. Extraction de contour.

#### IV.7. Extraction des caractéristiques globales et description du mot

Cette phase doit assurer un maximum de fiabilité, car les traitements ultérieurs ne vont plus manipuler l'image d'origine mais plutôt les résultats fournis par ce module, donc le choix des primitives à extraites est primordial pour garantir le succès d'une méthode de reconnaissance de forme, pour cette raison cette phase présente l'objet de notre travail.

La reconnaissance d'une forme est basée sur une description de celle-ci. La description peut ainsi être vue comme une modélisation mathématique de l'objet permettant un calcul de distance entre deux ou plusieurs modèles. Il existe un très grand nombre de descriptions possibles pour une forme suivant l'information à

modéliser [TRI 96] et les invariances à exprimer. LONCARIC [LON 98] distingue trois classifications possibles pour les descripteurs de forme :

- La première, la plus courante, vient de PAVLIDIS [PAV 78] qui considère d'une part les descripteurs basés sur la frontière (*descripteurs externes*) et d'autre part ceux basés sur la forme elle-même (*descripteurs globaux ou internes*).
- La seconde consiste à différencier les descripteurs numériques (scalaires ou vectoriels), des descripteurs qui ne le sont pas, également désignés comme techniques appartenant au domaine spatial.
- Enfin la troisième classe les descripteurs selon qu'ils préservent entièrement ou seulement partiellement l'information contenue dans la forme.

Nous exposons, ici, différentes approches utilisées en description de formes. Le tableau IV.1 montre quatre formats de représentation d'image arrangés en colonnes qui sont : image de gris, image binaire, le contour et le squelette. Pour chaque format, des méthodes d'extraction de caractéristiques raisonnables sont indiquées. Notons que le choix des caractéristiques et du format de représentation ne sont pas totalement dépendants ni tout à fait indépendants. Les méthodes d'extraction peuvent être distinguées selon deux classifications différentes décrites dans le tableau IV.2. Cette taxonomie s'appuie sur la relation entre les caractéristiques et la méthode de classification utilisées. Ainsi, deux catégories de classifieurs sont considérées : structurelles et statistique. Les caractéristiques sont organisées en lignes selon deux classes : heuristique et systématique.

Caractéristiques	Image de gris	Image binaire	Contour	Squelette
Appariement	X	X		X
Motif déformables	X			X
Transformation unitaire	X	X		
Transformation log-polaire	X	X		
Moments géométriques	X	X		
Moments Zernik	X	X		
Ondelettes	X	X		
Algébriques	X	X		
Histogrammes de projection		X		
Masques	X	X		
Profil du contour			X	
Code de Freeman			X	
Spline			X	
Descripteurs de Fourier			X	X
Description graphique				X
Discrets				X
zonage	X	X	X	X

Tableau IV.1. Taxonomie des méthodes d'extraction de caractéristiques selon la représentation de l'image, d'après [LAA 97].

	Structurel	Statistique
Heuristique	Caractéristiques discrètes	zonage
Systématique	Code de Freeman Spline graphe	Volumétrie Profil du contour Descripteurs de Fourier

Tableau IV.2. Taxonomie des caractéristiques selon la méthode de classification.

### IV.7.1. Choix des caractéristiques

Le concepteur d'un nouveau système de reconnaissance se trouve toujours confronté à ce choix délicat lors de la définition de son ensemble de caractéristiques. Les caractéristiques utilisées dans ce travail sont présentées par la suite, notant que ces caractéristiques sont implémentées en MATLAB.

D'après le tableau IV.1, Nous pouvons représentée l'image sous forme binaire, contour ou squelette ; et les caractéristiques peuvent êtres extraites à partir de chacun d'eux. Notre but dans ce travail est d'avoir une bonne représentation de l'image. Trois ensembles de caractéristique sont proposés, le premier ensemble basé sur l'image binaire ; le deuxième basé sur le contour et le troisième basé sur le squelette de l'image.

Les descriptions numériques résumant l'information contenue dans la forme entière par un vecteur d'attributs. Il existe une multitude de descripteurs dans la littérature nous présentons dans ce qui suit les trois descripteurs utilisés dans notre travail : les moments de Zernike et le contour et le zonage.

### IV.7.2. Descripteurs globaux

L'image I d'une forme peut être décrite par une application  $f : (x, y) \rightarrow \{0, 1\}$  où  $(x, y) \in I$  sont les coordonnées des pixels de l'image. Si le pixel  $(x, y)$  est en dehors de la forme  $f(x, y) = 0$  et  $f(x, y) = 1$  s'il appartient à la forme. l'ensemble des points de la forme est noté  $F = \{(x, y) \in I | f(x, y) = 1\}$  et  $I(n \times m)$  est le support de l'image avec m le nombre de lignes et n le nombre de colonnes. Parmi les descripteurs globaux on trouve les moments de Zernike. D'après les études récentes nous avons trouvé que ce descripteur est le meilleur dans le cas d'une image binaire :

#### *Les moments de Zernike*

Les moments sont utilisés en physique pour décrire la répartition des masses dans un corps. En analyse d'image, on peut envisager la même démarche en associant le niveau de gris d'un point de l'image à la même élémentaire en un point. Donc les moments fournissent des informations concernant l'arrangement spatial de l'image.

Les polynômes de Zernike ont été définis en 1934 dans le cadre de la théorie de la diffraction optique [ZER 34]. Dérivés de ces polynômes, les moments ont été utilisés par de nombreux auteurs en reconnaissance de caractères [TEA 80, KHO 90a, KHO 90b, BEL 91b]. Plusieurs études montrent également la supériorité de ces descriptions par rapport à d'autres approches [TEH 88, KIM 00]. Ils sont définis par les équations suivantes :

$$Z_{pq} = \frac{(p+1)}{\pi(N-1)^2} \sum_{x=1}^N \sum_{y=1}^N V_{nm}^*(r, \theta)^* f(x, y) \quad 4.1$$

Avec

$$V_{nm}(x, y) = R_{nm}(r)e^{jm\theta} \quad 4.2$$

Où

n : un entier positif ou nul

m : un entier tel que  $|m| \leq n$

p : la longueur du vecteur entre l'origine et le pixel (x,y)

$\theta$  : L'angle entre le vecteur p et x

$R_{nm}$  : un polynôme radial

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!r^{n-2s}}{s! \binom{(n+|m|)/2}{s} \binom{(n-|m|)/2}{s}} \quad 4.3$$

Les moments de Zernike est donc la projection d'une fonction  $f(x,y)$  décrivant une image sur un espace de polynômes orthogonaux engendré par  $V_{nm}^*(p, \theta)$  [ZHE 00]. On désigne par moments de Zernike une série de calculs utilisée pour transformer une image en un vecteur de composantes réelles représentant les moments  $Z_{ij}$ .

Les moments sont invariants par rotation, translation et changement d'échelle (après normalisation de la taille de la forme). De plus, grâce à l'exploitation d'une base de fonctions orthogonales, ces moments sont peu corrélés. Ces caractéristiques sont très importantes pour décrire les mots arabes à cause de ses grandes variations d'écritures comme nous avons présentés dans le chapitre 1 section III.3)

L'ordre des moments possède une grande influence sur la conservation de l'information angulaire. Plus l'ordre est élevé et plus les variations angulaires décrites sont fines. Dans notre système ; nous avons utilisé les 100 premiers moments de Zernike.

### IV.7.3. Descripteurs externes

Comme pour une forme, l'image I d'un contour peut être vue comme une application  $g(x, y) \rightarrow \{0, 1\}$  où  $(x, y) \in I$  sont les coordonnées des pixels de l'image. Si le pixel  $(x, y)$  est en dehors du contour  $g(x, y) = 0$  et  $g(x, y) = 1$  s'il appartient au contour.

#### IV.7.3.1. Extraction de contour

Il y a plusieurs méthodes pour obtenir le contour d'une forme binaire. Les plus simples sont les méthodes morphologiques (érosion, dilatation,...etc) [ARR 02].

Le contour interne d'un objet s'obtient par une érosion d'un objet A (suivant l'élément B) suivie d'une différence :

$$A - (A \ominus B) \quad 4.4$$

Le contour externe s'obtient par une dilatation la forme à partir d'un masque  $3 \times 3$  plein puis en lui soustrayant l'image de départ (figure III.17):

$$A - (A \oplus B) \quad 4.5$$

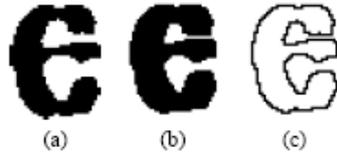


Figure IV.17. Extraction morphologique de contour externe, (a) image originale, (b) image dilatée, (c) contour externe.

#### IV.7.3.2. Détermination des contours des mots

L'opération de traçage des contours de l'image sert à décrire l'image d'un mot sous forme d'une liste de contours pouvant contenir des contours externes correspondant soit aux composantes connexes primaires du mot (corps principal) ou à ses composantes secondaires (points diacritiques), soit des contours internes (occlusions). Cette opération nous a permis de déterminer la chaîne de code de Freeman de chaque tracé du mot, qui va nous donner par son alphabet un critère de similarité nous permettant de mettre en correspondance les contours de deux images stéréoscopiques. Son principe est de décrire les arcs formant un contour par une suite de vecteurs de taille élémentaire et de direction choisie dans un ensemble fini. La direction d'un arc est codée par une valeur comprise entre 0 et 7 dans le sens trigonométrique.

#### IV.7.4. Zonage

Le principe de cette méthode est résumé en trois étapes suivantes :

- 1) calculer le nombre total de pixels blancs dans l'image.
- 2) diviser l'image en  $p \times q$  zones de même taille.
- 3) Pour chaque zone, on calcule la valeur moyenne de pixels blancs de tous les pixels blancs de l'image, en appliquant cette formule :

$$C_j = R_j / N \quad 4.6$$

Où

- $C_j$  : concentration des points blancs dans la zone  $j$ .
- $R_j$  : nombre de points blancs dans la zone  $j$ .
- $N$  : nombre total de points blancs du mot.

Le résultat obtenu sera utilisé comme un élément du vecteur de caractéristiques.

#### IV.7.5. Caractéristiques structurelles d'un mot arabes

Dans notre système, nous avons retenu aussi les caractéristiques globales structurelles suivantes qui sont en totalité 9 :

- Le nombre de hampes dans chaque composante connexe.
- Le nombre de jambages dans chaque composante connexe.
- Le nombre de boucles dans chaque composante connexe.

Et les points diacritiques qui sont :

- Le nombre de un point unique haut dans chaque composante connexe.
- Le nombre de deux points liés hauts dans chaque composante connexe.
- Le nombre de trois points liés hauts dans chaque composante connexe.
- Le nombre de un point unique bas dans chaque composante connexe.
- Le nombre de deux points liés bas dans chaque composante connexe.
- Et une caractéristique statistique qui représente le pourcentage de chaque composante connexe dans le mot (Le nombre des composantes connexes).

L'écriture arabe sous sa forme imprimée ou manuscrite est semi-cursive par nature, elle impose donc la notion de sous-mot et présente une certaine richesse en diacritiques, permettant de différencier des lettres ayant la même forme de base.

Pour décrire les différentes caractéristiques citées ci-dessus nous avons besoin de certaines informations élémentaires qui sont:

- ♣ la hauteur de la zone supérieure du mot en pixels. On la désigne par HS.
- ♣ la hauteur de la zone inférieure du mot en pixels. On la désigne par HI.
- ♣ la hauteur de la zone médiane du mot en pixels. On la désigne par HM.
- ♣ L'aire d'une composante connexe est le nombre total des pixels noirs qui forment cette connexité. On la désigne par AC.
- ♣ l'aire supérieure d'une composante connexe qui représente le nombre des pixels noirs de la composante connexe dans la zone supérieure du mot. On la désigne par ACS.
- ♣ L'aire inférieure d'une composante connexe qui représente le nombre des pixels noirs de la composante connexe dans la zone inférieure du mot. On la désigne par ACI.
- ♣ La hauteur du composante : C'est la hauteur en pixels entre la ligne basse et la ligne haute de la composante connexe, voir Figure (III.18). On la désigne par HC.
- ♣ La hauteur de composante supérieure : C'est la hauteur en pixels de la composante connexe dans la zone supérieure du mot. Elle est définie entre la ligne haute de la composante et la ligne base haute du mot. On la désigne par HCS.
- ♣ La Hauteur de composante inférieure : C'est la hauteur en pixels de la composante connexe dans la zone inférieure du mot. Elle est définie entre la ligne basse de la composante et la ligne de base du mot. On la désigne par HCI.

La figure suivante récapitule tout ce qu'on a défini jusqu'à présent :

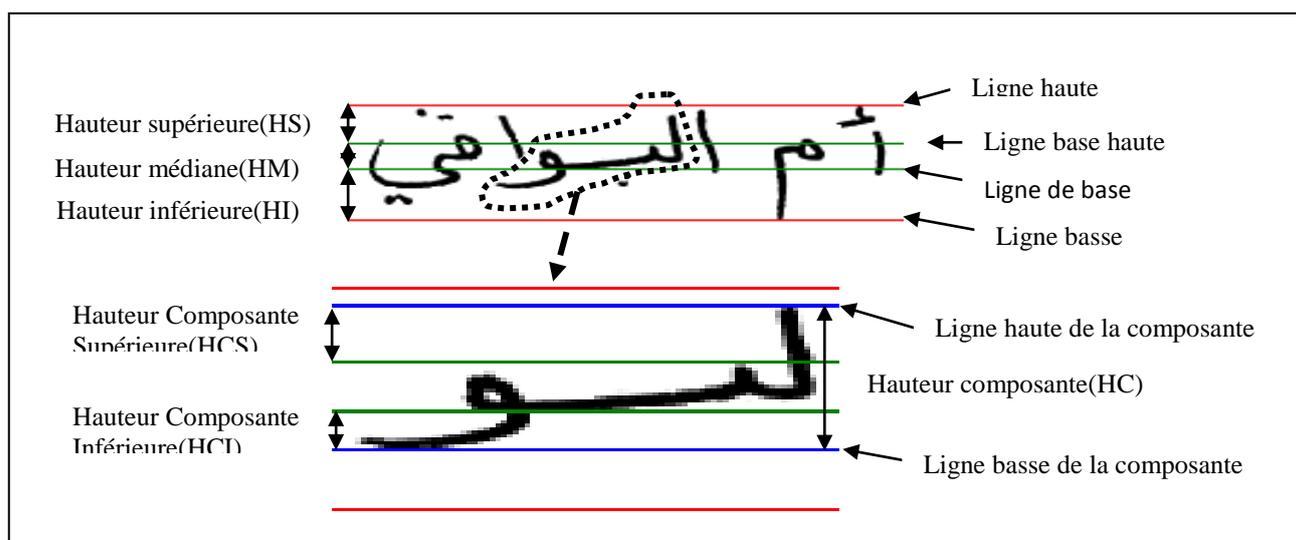


Figure IV.18. Illustrations des différentes informations nécessaires pour l'analyse d'une composante connexe.

En ce qui concerne les différentes informations de la composante connexe (AC, ACS, ACI, ligne haute de la composante et la ligne basse de la composante), elles sont extraites à partir d'une analyse de l'histogramme horizontale de cette dernière. On fait une projection horizontale de tous les pixels de la composante connexe pour obtenir cet histogramme, ensuite on obtient :

- Ligne haute de la composante : correspond au début de l'histogramme.
- Ligne basse de la composante : correspond à la fin de l'histogramme.
- AC : est la densité des pixels dans l'histogramme.
- ACS : est la densité des pixels du début de l'histogramme jusqu'à Ligne base haute du mot.
- ACI : est la densité des pixels du ligne base du mot jusqu'à la fin de l'histogramme.

Comme on a indiqué pour les mots, certaines composantes connexes aussi peuvent ne pas avoir des zones inférieures et/ou des zones supérieures. Alors, pour chaque composante connexe détecter lors de l'étape de prétraitement, on a effectué un certain nombre de traitement pour extraire leurs caractéristiques structurelles, c'est ce qu'on va discuter dans ce qui suit.

Dans l'écriture arabe, un mot se compose généralement d'une ou plusieurs composantes connexes, chacune contient un ou plusieurs caractères, et contient différentes caractéristiques structurelles comme :

- ❖ Les hampes qui sont les dépassements hauts (ascendants).
- ❖ Les jambages qui sont les dépassements bas (descendants).
- ❖ Les boucles (occlusions).
- ❖ Les points diacritiques qui sont :
  - ❖ le point unique haut (PSH) ;
  - ❖ les deux points liés hauts (PDH) ;
  - ❖ les trois points liés hauts (PTH) ;
  - ❖ le point unique bas (PSB) ;
  - ❖ les deux points liés bas (PDB) ;

Les figures (IV.19.a et IV.19.b) suivante récapitulent ces caractéristiques :

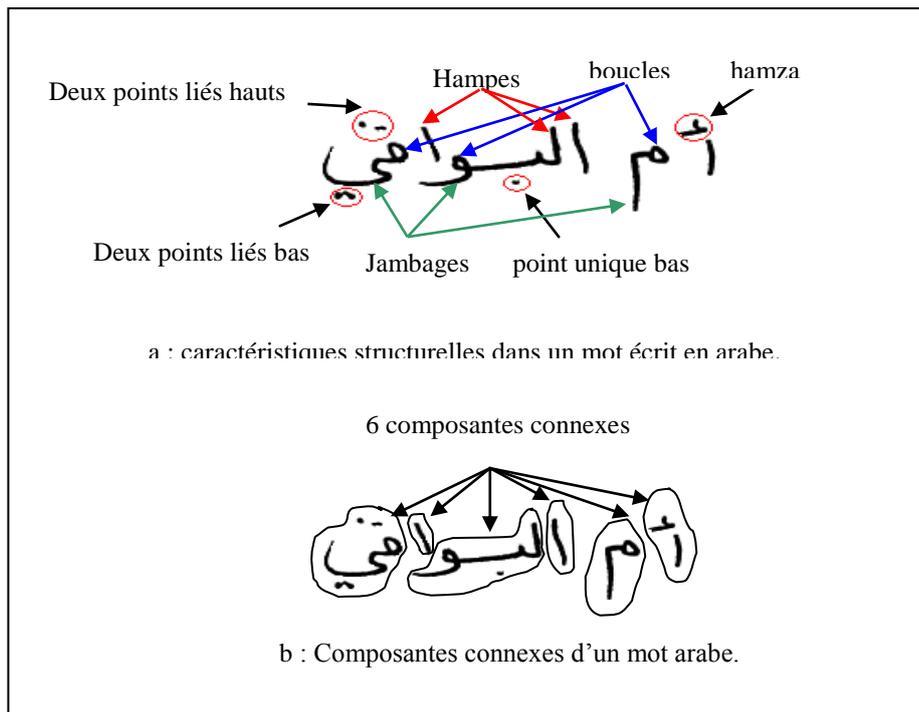


Figure IV.19. Topologie De L'écriture Arabe Illustrée Dans Le Mot « Oum-El-Bouaghi ».

Dans notre travail, l'extraction de toutes ces caractéristiques, consiste à prendre chaque composante connexe détectée dans la phase de prétraitement et de lui faire subir un ensemble de traitements. La détection de ces composantes connexes est effectuée par un balayage de l'image étiquetée de droite à gauche dans le sens de l'orientation de l'écriture arabe.

#### IV.7.5.1. Détection des différents types des points diacritiques

On a procédé notre analyse par la vérification si la composante connexe détectée est un point diacritique ou non, cette vérification est faite par le test suivant :

##### **Test Vérification\_Point\_Diacritique ;**

Si ((AC = ACS ) et ( HC <= HM )) Alors c'est un point diacritique haut

Sinon Si ((AC = ACI ) et ( HC <= HM )) Alors c'est un point diacritique bas

Sinon composante connexe n'est pas un point diacritique ;

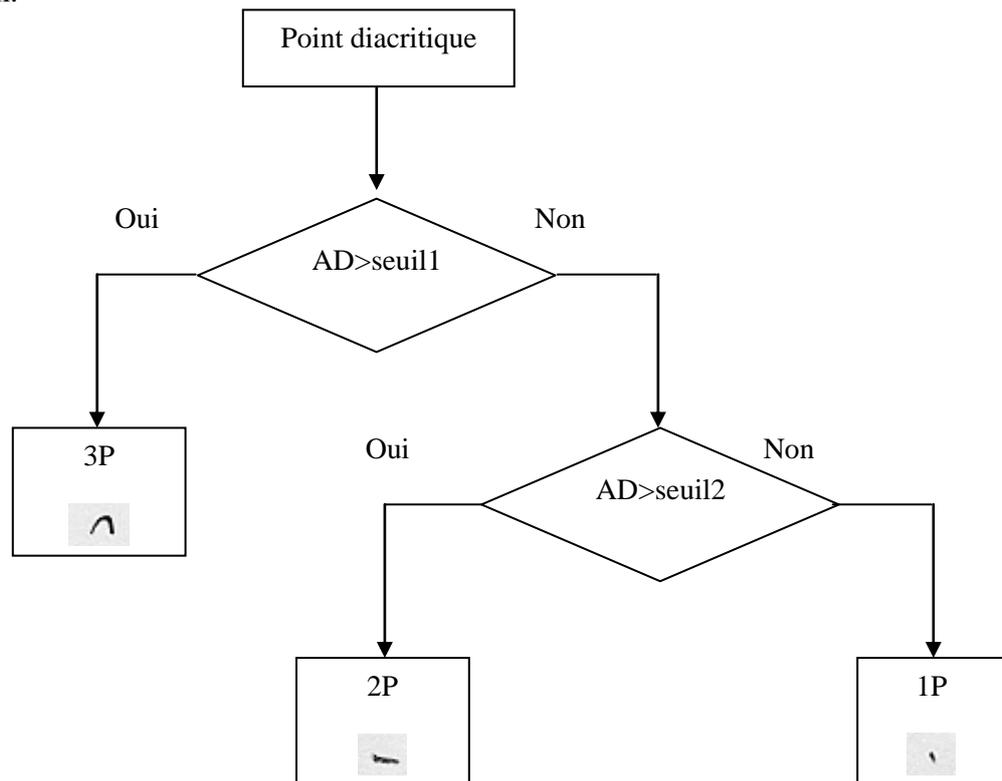
En effet dans un point diacritique haut, tout les pixels qui le constituent se situent dans la zone supérieure du mot et sa hauteur ne dépasse pas la hauteur médiane, cette dernière était démontrée expérimentalement. De la même façon on détecte les points diacritiques bas dans la zone inférieure.

Si le résultat du test assure que la composante connexe est un point diacritique. Nous devons classer ce point selon leur types (point unique, deux point liés...), comme le montre le tableau ci-dessous :

catégories types	Catégorie1	Catégorie2	Catégorie3
Un point		-	-
Deux points liés			-
Trois points liés			

Tableau IV.3. Les Différents Type Des Points Diacritiques.

Dans un premier temps, on va déterminer la catégorie1 pour les trois types des points. Après une étude statistique sur la densité des pixels de chaque type on a constaté les résultats dans l'organigramme dans la figure IV.20, tous les seuils utilisés ultérieurement dépendent du tracé de l'écriture et la résolution du capteur utilisé pour l'acquisition.



AD : aire du point diacritique.

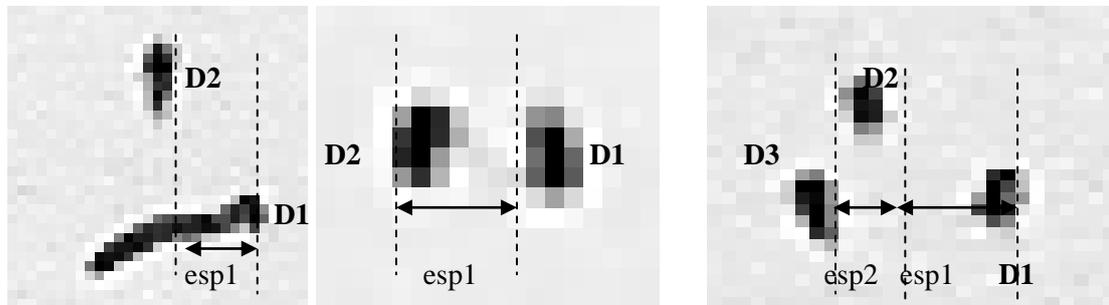
1P : point unique.

2P : deux points liés.

3P : trois points liés.

Figure IV.20. Organigramme de détection des diacritiques (Catégorie 1).

Pour la catégorie2 et la catégorie3 on aperçoit que ce sont un mélange entre les points diacritiques de la catégorie1, donc après avoir déterminé chaque type des points diacritiques dans tout le mot (l'organigramme précédent), on effectue une analyse basée sur l'espace entre les points diacritiques dans la zone supérieure et la zone inférieure séparément.



[a] : trois points liés.

[b] : deux points liés.

[c] : trois points liés.

Figure IV.21. Types de points diacritiques (catégorie2 et catégorie3).

La détection de chaque catégorie de cette figure est résumée dans l'organigramme suivant :

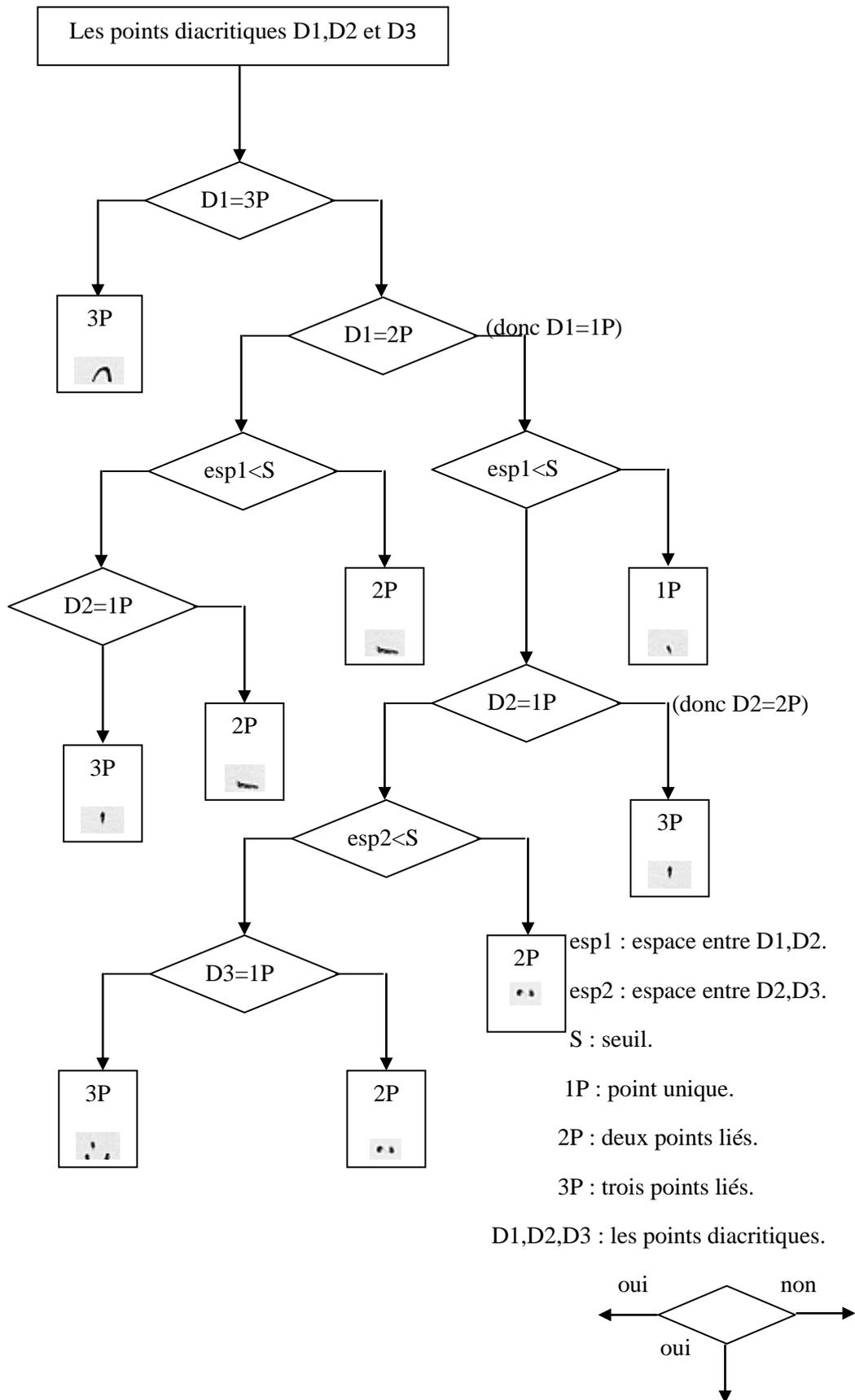


Figure IV.22. Organigramme de détection des diacritiques (les 3 catégories).

Nous passons maintenant à la détection du reste des caractéristiques structurelles si le résultat du test assure que la composante connexe n'est pas un point diacritique.

#### IV.7.5.2. Détection des boucles (occlusions)

Elles appartiennent aux composantes primaires à proximité de la ligne de base. L'idée qu'on a utilisée pour la détection des boucles dans une composante connexe est l'étiquetage du fond (le blanc de l'image).

Le total des boucles dans une composante connexe sera le nombre des fonds détectés moins un (moins le fond de l'image), Figure (IV.23) :

*Algorithmes de détection des nombres de boucles:*

Copier l'image du mot dans une nouvelle image, et inverser les couleurs de l'image binarisée.

Détecter le nombre de groupes de pixels blancs.

Nombre de boucles ← nombre de groupes de pixels blancs – 1.

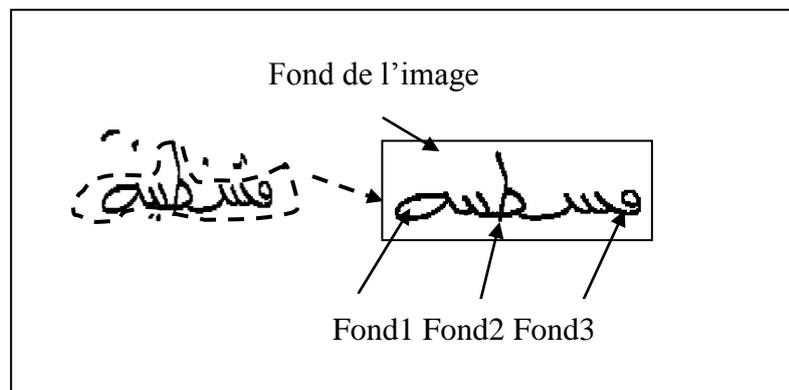


Figure IV.23. Trois Fonds Correspondent Aux Trois Boucles.

#### IV.7.5.3. Détecter le nombre d'ascendants (hampes) et de descendants (jambages)

Les ascendants et les descendants sont les primitives les plus utilisées dans la reconnaissance de l'écriture arabe. Elles sont détectées par une montée ou une descente qui se termine en dehors de la zone médiane déterminée à partir de la ligne de base.

##### **Algorithmes de détection des nombres des ascendants et des descendants**

Détecter le bottom de l'image par rapport à la ligne de base.

Si le bottom est inférieur à la ligne de base

Alors incrémenter le nombre de descendants

Détecter le top de l'image par rapport à la ligne de base.

Si le top est supérieur à la ligne de base

Alors incrémenter le nombre d'ascendants.

#### *IV.7.5.3.1. Détection des jambages*

Une composante connexe d'un mot arabe contient au maximum un seul jambage, le test suivant vérifie l'existence de ce jambage ou non :

<p><b>Test Vérification_Jambage ;</b> Seuil_J <math>\leftarrow</math> (HM / 2) ; Si (HCI <math>\geq</math> Seuil_J ) Alors il existe un jambage Sinon il n'existe pas un jambage ;</p>
--

Le Seuil\_J était calculé expérimentalement par l'analyse suivante, on a remarqué dans la plupart des mots manuscrits que la hauteur du jambage dépasse un certain seuil, on a essayé de calculer sa valeur minimale. On a trouvé que le plus court jambage est celui qui correspond à la lettre « و » et son hauteur dépasse la moitié de la zone médiane dans la plupart des mots de la base.

#### *IV.7.5.3.2. Détection des hampes*

Par opposition des jambages, une composante connexe d'un mot arabe peut contenir plusieurs hampes, alors le test sur l'existence ne suffit pas, mais un autre traitement est ajouté pour compter le nombre de ces hampes s'ils existent.

<p><b>Test Vérification_hampe ;</b> Seuil_H <math>\leftarrow</math> HM ; Si (HCS <math>\geq</math> Seuil_H ) Alors il existe des hampes Sinon il n'existe pas des hampes ;</p>
--

Le test précédent nous permet de vérifier l'existence des hampes dans une composante connexe, le Seuil\_H était obtenu expérimentalement et il est égal à la hauteur de la zone médiane.

Si le test précédent affirme qu'il existe des hampes dans la composante on fait appelle à l'algorithme suivant pour calculer leurs nombre :

<p><b>Algorithme Calcul_Nombre_Hampe ;</b> Seuil <math>\leftarrow</math> (HCS / 3) ; Niveau <math>\leftarrow</math> Ligne haute de la composante + Seuil ; On fait une projection verticale de la ligne d'ordre Niveau de l'image ; On calcule le nombre des hampes dans l'histogramme vertical résultant ;</p>
---

Nous pouvons expliquer cet algorithme schématiquement (Seuil est obtenu par expérience) :

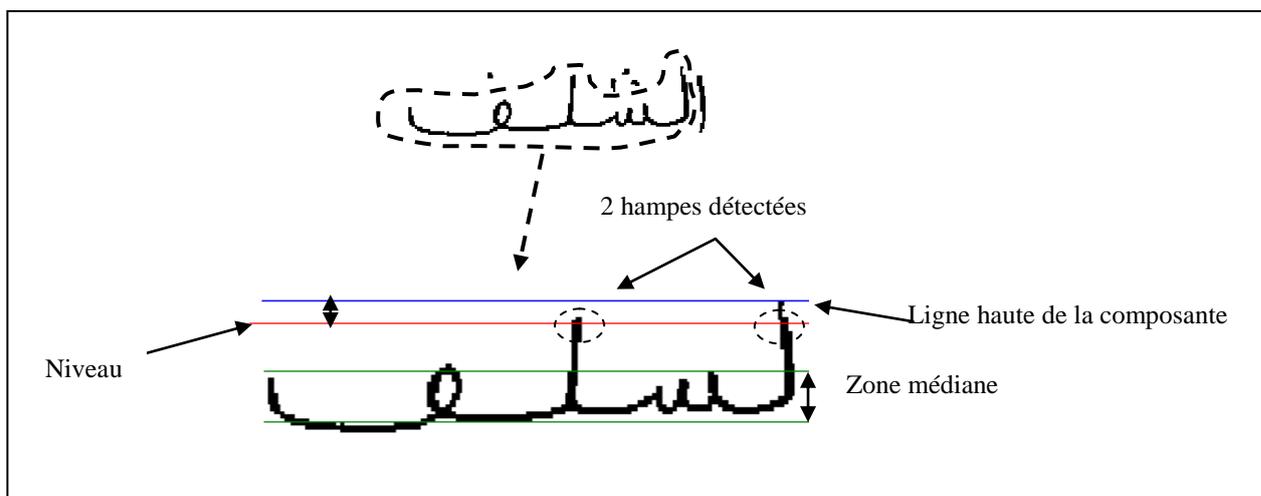


Figure IV.24. Illustration du niveau de projection verticale pour obtenir le nombre de hampes.

## IV.8. décomposition du vocabulaire considéré

Les mots arabes et plus précisément manuscrits comportent des différentes façons d'écriture entre les scribes, pour cette raison lors de la construction de la base des images de notre système, nous avons pris pour chaque mot un nombre d'exemples.

Pour la construction de la base d'images, nous avons obtenu 300 pages de formulaires remplis, les documents ont été scannés et les mots qui y sont contenus ont été extraits et étiquetés de manière automatique (d'après un algorithme réalisé par [BEN 00]). La phase d'étiquetage consiste à classer toutes les images, sachant que le vocabulaire des noms de wilayas algériennes contient 48 mots, nous avons conçu un modèle de formulaire pré-imprimé constitué de trois pages (16 mots/page), chaque mot devant être recopiés trois fois par chaque scribe (l'image contient un seul mot). La base contient donc 14400 images de mots manuscrits (les 48 mots du vocabulaire, écrits 3 fois, par chacun des 100 scribes). Dans ce qui suit, nous présentons en échantillon des trois pages des formulaires (voir annexe1).

La limitation du lexique est un avantage de la reconnaissance de noms de wilayas algériennes, chacun des 48 mots de notre lexique est composé d'un, deux, ..., ou six sous-mots SM (voir tableau IV.4). Nous avons pris l'idée de partitionner la base selon le nombre de sous-mots d'après le travail réalisé par [SOU 06].

N°	Description	Mot	N°	Description	Mot
1	1- Tebessa00	تبسة	25	3- Tiaret00	تيارت
2	1- Jijel00	جيجل	26	3- Tamanrasset00	تمنراست
3	1- Tissemsilt00	تسمسلت	27	3- Tippaza00	تيبازة
4	1- Khenchela00	خنشلة	28	3- Ouargla00	ورقلة
5	1- Mila00	ميلة	29	3- Illizi00	اليزى
6	1- MSILA00	مسيلة	30	3- Relizane00	غليزان
7	1- Mascara00	معسكر	31	3- Medea00	المدية
8	1- Setif00	سطيف	32	3- Blida00	البايدة
9	1- Constantine00	قسنطينة	33	3- Naama00	النعاما
10	2- Saida00	سعيدة	34	4- Ghardaia00	غرداية
11	2- Bechar00	بشار	35	4- Oran00	وهران
12	2- Bejaia00	بجاية	36	4- SidiBELABES00	سيدي بلعباس
13	2- Annaba00	عنابة	37	4- Bouira00	البويرة
14	2- Batna00	باتنة	38	4- Alger00	الجزائر
15	2- Biskra00	بiskرة	39	4- AinDefla00	عين الدفلة
16	2- Mostaganem00	مستغانم	40	4- Eltaref00	الطارف
17	2- Skikda00	سكيكدة	41	5- TiziOuzou00	تيزى وزو
18	2- Tlemcen00	تلمسان	42	5- Boumerdes00	بومرداس
19	2- ELbayad00	البيض	43	5- Adrar00	ادرار
20	2- Ghuelma00	قالمة	44	5- Eloued00	الوادي
21	2- Chlef00	الشلف	45	5- Laghouat00	الاغواط
22	2- Djelfa00	الجلفة	46	6- BBArrieridj00	برج بوعريديج
23	3- Tindouf00	تندوف	47	6- SoukAhras00	سوق اهراس
24	3- AinTemouchent00	عين تموشنت	48	6- OumElBouaghi00	ام البواقي

Tableau IV.4. Description globale des noms de wilayas en mots.

## IV.9. Reconnaissance

Tous les types de caractéristiques (statistiques, structurelles, géométriques, transformations globales) et les méthodes de classification (statistique, structurelle, stochastique, connexionniste, logique flou,..) qui existent dans la littérature sont pratiquement utilisées dans la reconnaissance de l'écriture arabe (voir section III.3).

Il y a trois façons d'aborder cette problématique [VIN 02]. Soit le système reconnaît le mot comme une entité entière et indivisible, il s'agit d'une approche *globale* ou *holistique* [ARR 02]. Soit il reconnaît le mot à partir de ses caractères préalablement segmentés, il s'agit d'une approche *analytique* [SAY 73, ARR 02]. Soit il n'utilise que certaines propriétés et raffine sa description du mot par rebouclage, nous parlerons alors de *systèmes basés sur la lecture humaine* [ARR 02].

A partir ce que nous avons vue dans le paragraphe ci-dessus, nous proposons une combinaison entre quatre types de classifieurs en focalisant à la première approche (globale). Dans notre système nous allons aborder la reconnaissance de mots arabes en appliquant deux méthodes statistiques KPPV ; K-MEANS et FCM, et une connexionniste en utilisant le réseau PNN qui est spécifique au problème de classification. Nous obtenons un système qui combine parallèlement les sorties de ces classifieurs qui sont décrit en détails dans le chapitre précédent.

Notre méthode comporte deux phases principales : la sélection des attributs par Les trois types de descripteurs cités ci-dessus (moment de Zernike, code de Freeman et Zonage) ; et la classification des mots à partir des ces attributs. Nous allons dans cette partie présenter chacune de ces deux phases ainsi que les problèmes rencontrés et les choix effectués.

#### **IV.9.1. Sélection des attributs**

Dans notre système, nous avons retenu les trois ensembles de caractéristiques, d'où chacun est extrait à partir d'un type d'image différents : le premier ensemble est extrait à partir d'une image binaire utilisant les moments de Zernike, le deuxième est extrait a partir de contour de l'image utilisant le code de Freeman et le troisième est extrait à partir de squelette utilisant la méthode de zonage. Enfin toutes ces caractéristiques sont regroupées en un seul vecteur identifiant le mot.

##### *Algorithme résumant l'extraction de caractéristiques*

Pour chaque mot faire :

- 1) Détecter les caractéristiques structurelles du mot (9 caractéristiques).
- 2) Détecter le vecteur de caractéristiques par le moment de Zernike (100 caractéristiques).
- 3) Détecter le code de Freeman (8 caractéristique).
- 4) Détecter le nombre de pixels noir pour chaque zone du mot (16 caractéristiques)..
- 5) Combiner ces trois types de caractéristique en un seul vecteur dite vecteur de caractéristiques.

Il faut noter que, comme beaucoup de techniques sont utilisées en REM, l'approche retenue pour notre système dans ce domaine en tant que reconnaissance de mots est celle basée sur le calcul des vecteurs de caractéristiques. En effet, les caractéristiques retenues pour notre démarche sont celles basées sur le zonage avec le calcul de densité des pixels, les moments de Zernike et code de Freeman. Les 133 caractéristiques retenues (16 pour le zonage (D'après la décomposition choisie, nous obtenons 16 caractéristiques, chacune représente la concentration des points blancs. Où chaque valeur représente la densité de points blancs présents dans chaque zone (lire le tableau ligne par ligne de la zone 1 à la zone 16), 100 pour les moments de Zernike et 8 pour le code de Freeman), et les 9 caractéristiques structurelles. Toutes ces caractéristiques constituent les composantes du vecteur de caractéristiques utilisé pour la reconnaissance. De ce fait, les vecteurs de caractéristiques sont calculés pour chaque mot et stockés dans trois bases : une pour les caractéristiques de zonage, une pour les moments de Zernike et l'autre pour la chaîne de Freeman.

La difficulté du choix des caractéristiques réside dans le compromis à établir entre les trois contraintes suivantes:

- ◆ rapidité d'extraction des caractéristiques.
- ◆ facilité de mise en œuvre.
- ◆ insensibilité aux distorsions du caractère (styles d'écriture différents, bruits dans l'image, taille du caractère, ...)

## IV.9.2. Classification par les méthodes implémentées

"Comme chaque méthode de classification a ses avantages et ses défauts, nous pouvons déduire que la performance d'un système peut être augmenté de manière significative en combinant plusieurs classifieurs." [SUE 98].

Après avoir sélectionné les caractéristiques discriminantes de chaque image, il faut ensuite séparer distinctement les différentes classes des mots. Pour cela, nous avons appliqués trois méthodes en comparant les résultats de chacune afin d'améliorer le taux de reconnaissance : la méthode KPPV ; KMEANS ; FCM, KPPV et un réseau de neurone de type PNN.

### IV.9.2.1. Classification par la méthode KPPV

*Un mauvais voisin est une calamité, un bon voisin un vrai trésor.  
Hésiode (poète grec)*

Un classifieur KPPV nécessite avant tout la création d'une base de référence, qui est constituée d'un nombre d'échantillons (mots), Chaque entrée sera comparée avec chaque élément de la base de référence (images des mots ayant déjà été classées). La comparaison s'effectue en calculant la distance entre les deux (distance euclidienne), Puis attribuer le mot à la classe la plus représentée parmi les K plus proche d'elle (auxquels la distance est minimale). En se basant sur ses caractéristiques extraites.

**Algorithme du KPPV est comme suit:**

**Etape1** : pour chaque échantillon  $y$  à identifier, on calcule la distance Euclidienne donnée par cette formule:

$$d_1(X, Y) = \left( \sum_{i=1}^k (X_i - Y_i)^2 \right)^{1/2} \quad 4.7$$

**Etape2** : consiste à prendre en considération les k mots les plus proches voisins du candidat inconnu  $y$ , afin de prendre la décision le concernant. On sélectionne la classe la plus représentée parmi les k votes où le mot  $y$  appartient à l'ensemble  $\{X_1, X_2, \dots, X_k\}$  dont la distance est minimale.

On note deux aspects importants de l'algorithme KPPV : d'une part, à chaque nouvelle classification il est nécessaire de parcourir l'ensemble de la base d'apprentissage, ce qui en fait un algorithmes qui n'est pas nécessairement très efficaces (surtout que, habituellement, on cherche à avoir la base d'apprentissage la plus grande possible afin d'avoir un meilleur classifieur), et d'autre part, un point crucial de cet algorithme est la fonction de distance utilisée pour mesurer la proximité des objets. Il n'existe pas de distance/similarité universellement optimale et une bonne connaissance du problème traité guide généralement le choix de cette distance/similarité [BEL 01].

#### IV.9.2.2. Classification par la méthode C-Moyenne Floue

Une des caractéristiques du raisonnement humain est qu'il est basé sur des données imprécises ou incomplètes, un ordinateur, lui, est basé sur des données exactes d'où la nécessité de développer une nouvelle logique dite logique floue cette dernière est une extension de la logique booléenne dans laquelle les niveaux de vérités, au lieu d'être vrais ou faux peuvent prendre des valeurs entre 0 et 1. Il n'y a en fait pas de domaine où l'on ne puisse appliquer les raisonnements propres à la Logique Floue car elle a été conçue pour s'adapter aux techniques de pensée humaine. Nous présentons dans la suite une description de la méthode floue de Fuzzy C-Means clustering « FCM ».

La méthode C-moyenne floue est une méthode de classification itérative qui permet de classer les individus selon C classes [CHU 00, GUI 01]. Elle calcule à chaque fois les centres des classes et génère la matrice d'appartenance U des individus à ces classes.

Soient  $V_i$  le centroïde ou prototype de la classe i, U la matrice des coefficients  $\mu_{ik}$  et  $X_C$  celle des coordonnées des centres. Etant donné le nombre de classe C, le nombre d'individus n et l'exposant flou m ( $m > 1$ ), l'objectif de la méthode est de trouver U et  $X_C$  qui minimisent la fonction coût donnée par la relation (4.8 et 4.9).

$$J(U, V, m) = \sum_{i=1}^C \sum_{k=1}^n \mu_{ik}^m D_{ki}^2 \quad 4.8$$

Avec

$$\sum_{i=1}^C \mu_{ik} = 1 \quad \forall k = 1, \dots, n \quad 4.9$$

$D_{ki}$  Est une métrique choisie au sens d'une norme. Généralement, il s'agit de la norme euclidienne. Ainsi  $D_{ki} = \|X_k - V_i\|_2$  : distance entre le vecteur  $X_k$  et le prototype  $V_i$ . La technique de classification par FCM repose sur l'algorithme présenté dans la figure IV.25.

### Algorithme de FCM

1. Soit  $X^i = (x_1, x_2, \dots)$  les vecteurs représentant les individus à classer

2. Fixer les paramètres

- m : coefficient flou
- C : le nombre de classes
- $\varepsilon$  : critère d'arrêt

3. initialiser le vecteur  $V$  par  $C$  centres aléatoirement choisis

4. calculer la matrice  $U$  de taille  $(C \times n)$  par les équations :

$$\mu_{ik} = \left( \sum_{j=1}^C (D_{jk})^{\frac{2}{m-1}} \right) / \left( \sum_{k=1}^n (D_{ik})^{\frac{2}{m-1}} \right)$$

$D_{jk}$  est la distance entre l'individu  $k$  et le centre  $V_j$

5. Calculer le nouveau centre de chaque classe à l'aide de l'équation :

$$V_i = \left( \sum_{k=1}^n (\mu_{ik})^m X_k \right) / \left( \sum_{k=1}^n (\mu_{ik})^m \right)$$

6. Mettre à jour la matrice  $U$  et incrémenter le compteur  $t$

7. Calculer la distance entre les nouveaux et les anciens centres par :

$$h = \|V^{t-1} - V^t\|$$

8. Répéter les étapes de 3 à 6 tant que  $h > \varepsilon$ .

Figure IV.25. Algorithme de classification par C-moyenne floue.

Le coefficient flou  $m$  qui est important dans la classification, fera l'objet d'un ensemble de test pour choisir la valeur optimale. L'exposant flou  $m$  est un coefficient  $\geq 1$ . Le poids de cet exposant peut être interprété comme un paramètre de distorsion. Le « flou » de la partition augmente avec ce coefficient. Un tel coefficient accentue les faibles niveaux d'appartenance et contribue donc à mieux séparer les classes. Dans cette partie, nous nous proposons d'étudier l'influence de ce paramètre sur le résultat de la classification FCM. Nous donnons un exemple pour classer les mots ayant 6 sous-mots (la dernière partition contenant 3 mots : سوق اهراس, برجوع عريبيج, البواقي ام), en variant ce paramètre. Nous avons utilisé les trois ensembles de caractéristiques (voir section IV.7). Le taux de reconnaissance sera calculé comme suit :

Taux de reconnaissance = (nombre d'images reconnues \* 100) / nombre totale d'images.

Dans cet exemple le nombre totale = 30\*3=90 images (30 images pour chaque mots).

Nom \ m	2	3	4
BBArredj00	25	24	25
SoukAhras00	24	25	28
OumElBouaghi00	25	27	25
totale	82.22	84.44	86.66

Tableau IV.5. Taux de reconnaissance obtenue en faisant varier le coefficient flou  $m$  et en utilisant les caractéristiques de zonage.

Nom \ m	2	3	4
BBArredj00	27	26	25
SoukAhras00	28	27	28
OumElBouaghi00	25	24	25
totale	90	85.55	86.66

Tableau IV.6. Taux de reconnaissance obtenue en faisant varier le coefficient flou  $m$  et en utilisant les caractéristiques de freeman.

Nom \ m	2	3	4
BBArredj00	26	24	25
SoukAhras00	24	25	27
OumElBouaghi00	25	28	26
totale	83.33	84.44	86.66

Tableau IV.7. Taux de reconnaissance obtenue en faisant varier le coefficient flou  $m$  et en utilisant les caractéristiques de moment de Zernike.

#### IV.9.2.3. Classification par méthode K-Means

La méthode K-Means, est une méthode classique de partitionnement [BOI 87]. Le but de cette méthode est de trouver la partition et les centroides de façon à minimiser la distorsion totale  $D$ . Une procédure itérative peut être basée sur les deux observations suivantes:

- ◆ Pour un ensemble donné de centroides, la partition qui minimise  $D$  (la distorsion totale pour l'ensemble des vecteurs) est celle pour laquelle chaque vecteur  $x_i$  (vecteur de caractéristiques) est affecté à la classe dont le centroïde est le plus proche.
- ◆ Pour une partition donnée, il existe pour chaque classe  $i$  un vecteur  $J_i$  qui minimise la distorsion totale  $D_i$  de la classe  $i$ .

Le déroulement typique de cet algorithme est résumé de la manière suivante :

- a. Initialisation consistant à faire le choix.
  - ♣ Des centres initiaux des classes (initialisation de code book).
  - ♣ Du seuil ( $10 \text{ e-}6$ )
  - ♣ De la valeur initiale de la distorsion totale  $D$ .
- b. Une boucle faisant alterner 2 étapes jusqu'à stabilité des résultats ou pour un nombre fixé d'itérations :

- ♣ Affectation des objets restants au groupe le plus proche.
- ♣ Redéfinition des représentants des groupes.

### **Algorithme de résolution**

Le principe de fonctionnement de l'algorithme K-means est décrit comme suit :

#### **Algorithme de k-means**

**Entrée:** vecteur d'apprentissage  $x_i$

*/\** Après l'analyse des images, nous obtenons des mots ( $X$ ) dont chacun d'eux est composé d'un ensemble de vecteurs (zonage ; Zernike ; code de Freeman). *\*/*

#### **1 Initialisation**

Choisir un code-book initial  $C = \{y_i ; 1 \leq i \leq M$ .

*/\** le choix des  $M$  centroides ( $y_i$ ) initiaux est arbitraire, pratiquement, ils peuvent être choisis comme étant les  $M$  ( $M = 9$ ) premières paramètres des mots de la base c à d : on prend pour chaque mot du vocabulaire le vecteur de caractéristiques correspondant au premier groupe contenant 9 mots *\*/*

Mettre  $m = 0$ ;  $D_m = D_0$  (valeur max).

#### **2 Construction des classes**

Partitionner les vecteurs d'apprentissage  $x_j$  en  $M$  classes selon l'hypothèse suivante :  $x_j \in C_i$  si  $d(x_j, y_i) \leq d(x_j, y_k)$  pour tout  $1 \leq k \leq M$ , où  $d$  représente la norme euclidienne.

Mettre  $m = m + 1$ ; et calculer :

$D = \sum_{i=1}^M \sum d(x_j, y_i)$ ; avec  $x_j \in C_i$

*/\** Une fenêtre  $x_j^i$  est affectée à une classe ( $i$ ) si la distance euclidienne  $d(x_j^i, y_i)$  est minimale par rapport aux autres classes *\*/*

#### **3 Mise à jour des centroides**

Mettre à jour chaque centroïde  $y_i$  par:  $Y_i = (1/|C_i|) \sum x_j$ ; avec  $x_j \in C_i$

*/\**  $|C_i|$  est le nombre de vecteurs dans la classe  $C_i$  *\*/*

#### **4 Test d'arrêt**

Si  $(D_m - D_{m-1})/D_m > \zeta$  alors aller à **2** */\**  $\zeta = 0.01$  *\*/*

Le code-book est constitué de tous les vecteurs  $y_i ; 1 \leq i \leq M$

Fin de l'algorithme

Figure IV.26. L'algorithme de résolution pour la méthode K-Means [LLO 57]

#### IV.9.2.4. Classification par réseaux de neurones probabilistes (PNN)

Ces réseaux sont généralement utilisés pour des problèmes de classification. La première couche qui est un réseau à base radiale, donne une information sur la ressemblance entre la donnée d'entrée et le jeu de données utilisé lors de l'apprentissage. La deuxième couche produit comme sortie un vecteur de probabilité. Finalement, une fonction de transfert compétitive produit 1 ou 0.

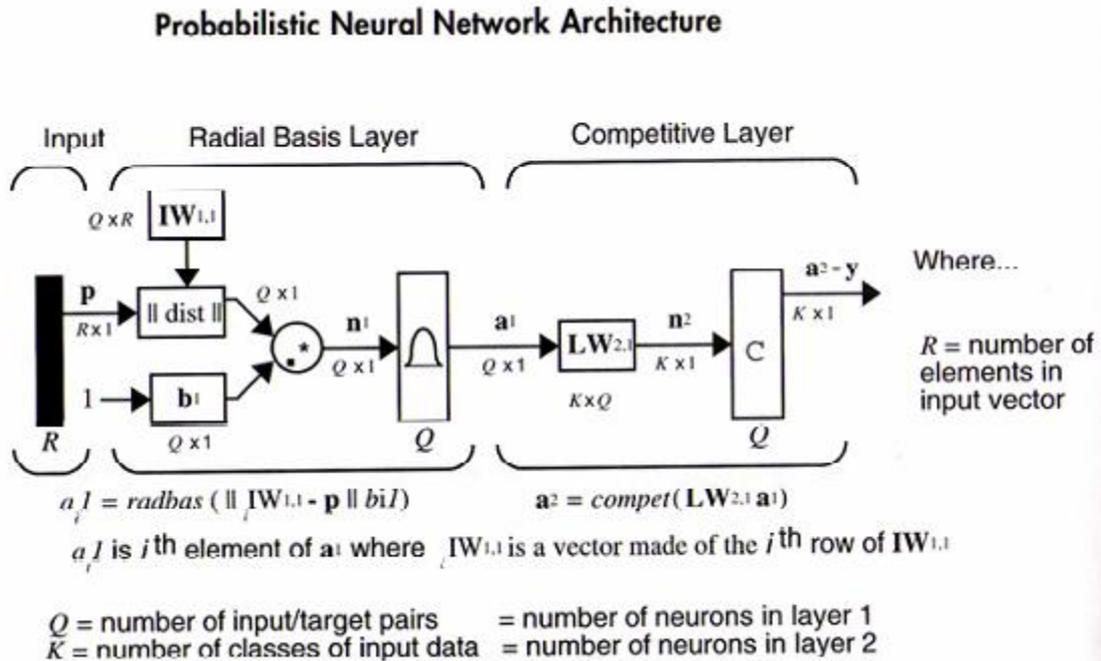


Figure IV.27. Architecture du réseau PNN.

#### III.9.2.4.1. Propriétés générales d'un PNN

- ❖ Classificateur neuronique 'feedforward'
- ❖ Structure: réseau à 3 couches :
  1. couche d'entrée: unités de distribution des caractéristiques
  2. couche cachée: noyaux Gaussiennes sphériques centrés sur chaque patron de TRAIN. Elle calcule les distances entre le vecteur d'entrée et tous les vecteurs d'apprentissage et produit un vecteur dont les éléments indiquent comment ce vecteur d'entrée est proche de chaque vecteur d'apprentissage.
  3. couche de sommation: sommation des neurones cachés par classe. Elle additionne ces contributions pour chaque classe des entrées pour produire à la sortie du réseau un vecteur des probabilités. Enfin, une fonction de transfert de concurrence sur la sortie de la deuxième couche sélectionne le maximum de ces probabilités, et produit un 1 pour cette classe et un 0 pour les autres classes.

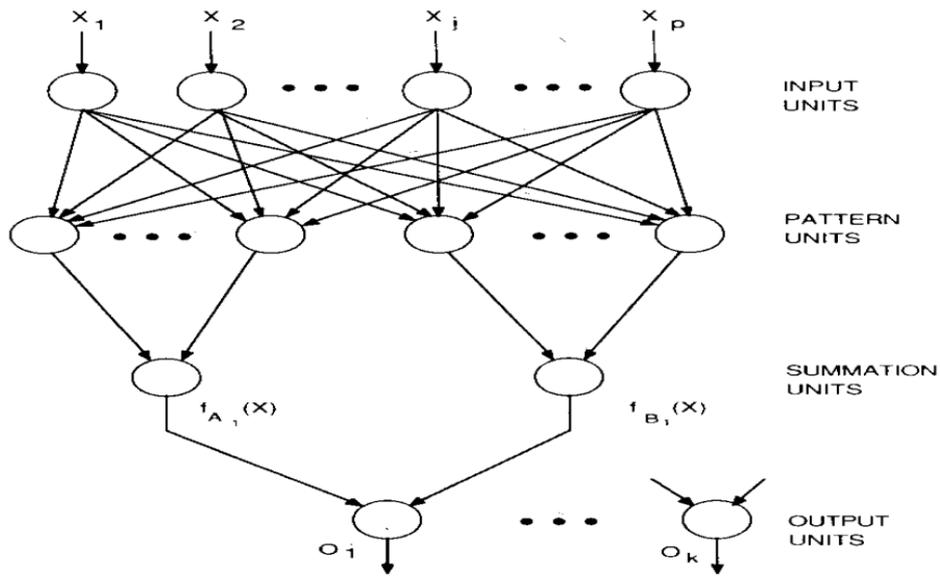


Figure IV.28. Structure d'un réseau PNN.

#### IV.9.2.4.1. Avantages et inconvénients de PNN

##### Avantages:

- ✓ apprentissage rapide (*entraînement* en une passe)
- ✓ converge asymptotiquement vers la borne de décision optimale quand  $|\text{TRAIN}|$  augmente
- ✓ peut donner des résultats comparables au MLP.

##### Inconvénients:

- ✓ ressources: un neurone caché est requis pour chaque patron de TRAIN
- ✓ alors, on limite aux problèmes avec petites bases.

### IV.9.3. Résultats et discussion

Dans ce paragraphe nous donnons quelques résultats de classification avec différents couples de descripteurs, classifieurs. L'objectif est de trouver les couples intéressants pour obtenir des combinaisons entre les méthodes de classifications présentées ci-dessus.

Pour pouvoir évaluer les performances du système réalisé on doit passer certainement par l'étape test qui consiste bien évidemment à démontrer l'efficacité des méthodes appliquées à savoir :

- ◆ Test de La reconnaissance effectué par les classifieurs adoptés dans notre système (KPPV, FCM, K-Means et PNN).
- ◆ Test de la combinaison de ces classifieurs cités précédemment.

Nous avons décomposé le vocabulaire selon le nombre de sous-mots constituant le mot afin de choisir le FCM adéquat (nous avons obtenu 6 FCMs), nous avons fait ça parce que nous constatons que ce type de classifieur donne des résultats mieux lorsque le nombre de classes ainsi que le nombre de données diminuent ; le même principe sera suivi pour les autres.

Nous avons divisé notre base en deux parties, une partie pour classer les mots (270 images pour chaque nom de wilaya) et une partie pour le test (30 image pour chaque nom de wilaya).

Une mauvaise reconnaissance se produit quand un mot testé est affecté à une autre classe que celle à laquelle il devait appartenir. Nous avons constaté que ce problème est généralement causé lors du calcul de sous-mots (nombre de composantes connexes), de plus certaines composantes connexes secondaires (les points diacritiques) seront considérées comme composantes primaires à cause de la variation de l'écriture et les heuristiques utilisées dans notre algorithme d'évaluation d'existence ou du type de diacritiques (voir section IV.7.5.1).

Nous avons utilisé des classifications simples dont les résultats sont combinés de deux manières ; par vote de majorité puis par priorité. Les résultats obtenus par le système sont résumés dans les tableaux (IV.8.a...d, IV.9.a, IV.9.b) :

Avec :

**Groupe1(G1) :** { قسنطينة, سطيف, معسكر, مسيلة, ميلة, حنشلة, تسمسيلات, جيجل, تبسة }

**Groupe2 (G2) :** { مستغانم, قالمة, عنابة, سكيكدة, سعيدة, الجلفة, تلمسان, بشار, بسكرة, بجاية, باتنة, الشلف, البيض }

**Groupe3 (G3) :** { تموشنت, النعامة, تيبازة, تندوف, اليزى, ورقلة, المدية, تيارت, تمنراست, البليدة, غليزان, عين }

**Groupe4 (G4) :** { عين الدفلة, غرداية, الطارف, وهران, سيدى بلعباس, الجزائر, البويرة }

**Groupe5 (G5) :** { الوادى, بومرداس, تيزى وزو, الاغواط, ادرار }

**Groupe6 (G6) :** { ام البواقي, اهراس سوق, بوعريريج برج }

#### IV.9.3.1. Classification simple

Voici les résultats obtenus utilisant les quatre types de classification avec une décision sans rejet :

- ♣ Classification paramétrique gaussienne par un réseau PNN
- ♣ Classification floue par FCM.
- ♣ Classification statistique par K-Means.
- ♣ Classification par la règle des k-PPV avec une distance euclidienne.

Nous avons comparé les résultats de classification simple de trois types de caractéristiques :

- ♣ Les descripteurs de Zernike.
- ♣ Le descripteur externe de code de Freeman.
- ♣ Par le Zonage, plus les 9 caractéristiques structurelles du mot arabe.

Les résultats obtenus par chacun des classifieurs indépendamment :

<b>KPPV</b>							
	Zonage (%)		Zernike (%)		Freeman (%)		Taux par groupe (%)
G1	<b>67.03</b>	k=6	56.30	k=7	64.44	k=6	<b>62.60</b>
G2	<b>61.02</b>	k=6	56.67	k=8	<b>58.46</b>	<b>k=8</b>	<b>58.72</b>
G3	<b>73.93</b>	k=10	<b>67.90</b>	<b>k=8</b>	73.33	k=11	<b>71.72</b>
G4	75.71	k=6	<b>78.09</b>	k=5	71.90	k=4	<b>75.23</b>
G5	<b>87.33</b>	k=3	85.33	k=6	86.66	k=6	<b>86.44</b>
G6	<b>95.55</b>	<b>k=3</b>	94.44	k=3	94.44	k=6	<b>94.81</b>
Taux par type de primitives (%)	<b>76.76</b>		<b>61.80</b>		<b>74.87</b>		

(a)

<b>FCM</b>				
	Zonage (%)	Zernike (%)	Freeman (%)	Taux par groupe (%)
G1	60	41.48	<b>61.48</b>	<b>54.32</b>
G2	43.33	38.72	<b>55.13</b>	<b>45.73</b>
G3	66.06	52.12	<b>67.27</b>	<b>61.82</b>
G4	64.29	<b>67.62</b>	66.19	<b>66.03</b>
G5	76.67	72	<b>78</b>	<b>75.56</b>
G6	81.11	87.78	<b>88.89</b>	<b>85.93</b>
Taux par type de primitives (%)	<b>65.24</b>	<b>59.95</b>	<b>69.49</b>	

(b)

<b>PNN</b>				
	Zonage (%)	Zernike (%)	Freeman (%)	Taux par groupe (%)
G1	<b>73.33</b>	62.96	72.22	<b>74.20</b>
G2	63.85	57.70	<b>65.13</b>	<b>62.05</b>
G3	<b>78.18</b>	72.73	77.88	<b>72.12</b>
G4	71.43	<b>77.62</b>	70.48	<b>76.68</b>
G5	<b>86.67</b>	82.67	85.33	<b>84.44</b>
G6	<b>94.45</b>	93.33	92.22	<b>93.71</b>
Taux par type de primitives (%)	<b>77.98</b>	<b>74.50</b>	<b>77.21</b>	

(c)

<b>K-means</b>				
	Zonage (%)	Zernike (%)	Freeman (%)	Taux par groupe (%)
G1	<b>60</b>	41.85	58.89	<b>53.58</b>
G2	<b>56.41</b>	37.95	<b>56.41</b>	<b>50.26</b>
G3	<b>66.67</b>	49.09	66.36	<b>60.71</b>
G4	62.86	<b>64.76</b>	61.90	<b>63.17</b>
G5	71.33	<b>72</b>	70.67	<b>71.33</b>
G6	86.67	<b>87.79</b>	85.55	<b>86.67</b>
Taux par type de primitives (%)	<b>67.32</b>	<b>58.91</b>	<b>66.63</b>	

(d)

Tableau IV.8. Résultats de classification simple (a, ..., d)

D'après les résultats obtenus, nous déduisons que le type de caractéristiques les plus intéressantes sont les caractéristiques structurelles avec les caractéristiques de zonage. Du côté des classifieurs, nous notons que l'évaluation paramétrique gaussienne (PNN) donne de bons résultats avec le zonage. De l'autre côté, nous voyons que le taux de reconnaissance augmente lorsque le nombre de classe soit plus petit dans le groupe.

Globalement les deux classifieurs KPPV et PNN atteignent des résultats comparables, et plus intéressants par rapport aux deux autres.

#### IV.9.3.2. Combinaison de classifieurs

À l'issue des résultats précédents, nous avons combiné les quatre classifieurs. Cette combinaison est faite soit de manière démocratique, dans le sens où elle ne favorise aucun classifieur par rapport à un autre, où bien de manière dirigée et, dans ce cas, on attribue à la réponse de chaque classifieur un poids en fonction de ses performances et ceci est en fonction généralement du taux obtenu lors de la phase d'apprentissage ; selon les résultats obtenus précédemment, nous avons donné la priorité au PNN puis au KPPV, et les deux autres possédant la même priorité. Les résultats obtenus se résument dans les deux tableaux suivants :

<b>Combinaison (vote de majorité)</b>				
	Toutes les caractéristiques (%)	Zonage (%)	Zernike (%)	Freeman (%)
G1	<b>67.8</b>	62.59	42.59	60.74
G2	<b>63.07</b>	56.67	40.77	56.15
G3	<b>74.85</b>	69.09	53.33	69.09
G4	<b>78.58</b>	70.48	66.67	65.24
G5	<b>87.33</b>	79.33	72.67	80
G6	<b>93.33</b>	86.67	90	86.67
totale	<b>77.50</b>	70.80	61	69.65

(a)

<b>Combinaison avec priorité entre les classifieurs</b>				
	Toutes les caractéristiques (%)	Zonage (%)	Zernike (%)	Freeman (%)
G1	<b>68.14</b>	62.59	42.96	61.85
G2	<b>63.33</b>	58.46	42.82	57.43
G3	<b>75.15</b>	70.61	55.76	68.79
G4	<b>88</b>	72.86	66.67	65.24
G5	<b>90.67</b>	81.33	74	80
G6	<b>93.33%</b>	86.67	90	87.78
totale	<b>79.80%</b>	72.09	62.03	70.18

(b)

Tableaux IV.9. Résultats de combinaison des classifieurs

### IV.9.3.3. Influence de paramètre K sur les résultats de la classification KPPV

Nous donnons un exemple pour classer les mots en variant le paramètre K. Nous avons utilisé l'ensemble de caractéristiques de zonage. Le tableau suivant présente les résultats obtenus :

zonage	K=1 (%)	K=3 (%)	K=6 (%)
G1	63.70	64.07	67.03
G2	58.46	59	61.02
G3	66.67	69.70	71.51
G4	71.90	74.28	74.76
G5	85.33	87.33	84
G6	85.33	87.87	95.55

Tableau IV.10. L'influence de paramètre K sur le résultat de la classification KPPV.

Nous constatons que le paramètre K influence directement le taux de reconnaissance. Dans l'exemple présenté ci-dessus ; le meilleur score est obtenu lorsque K=6.

### IV.9.3.4. Utilité de décomposition du vocabulaire

<b>K-means sans découpage de vocabulaire</b>				
	Zonage (%)	Zernike (%)	Freeman (%)	Taux par groupe (%)
G1	<b>51.11</b>	25.55	49.26	<b>41.97</b>
G2	37.95	24.87	<b>38.20</b>	<b>33.67</b>
G3	<b>46.06</b>	28.79	<b>46.06</b>	<b>40.30</b>
G4	35.24	26.67	<b>42.38</b>	<b>34.76</b>
G5	<b>46.67</b>	34.67	<b>46.67</b>	<b>42.67</b>
G6	<b>66.67</b>	51.11	65.55	<b>61.11</b>
Taux par type de primitives (%)	<b>47.28</b>	<b>31.94</b>	<b>48.02</b>	

Tableau IV.11. Résultats de K-means sans découpage de vocabulaire.

Le tableau précédent montre que le taux de reconnaissance a diminué lorsqu'on a utilisé toute la base, cela met en évidence l'importance de découper la base selon le nombre de sous-mots qui le constituent.

Enfin, Nous avons évalué les performances de notre système, nous l'avons testée sur une base de 1440 images des mots. Environ 80% de ces mots ont été bien affectés à la classe.

## **IV.10. Conclusion**

Dans ce chapitre, nous avons présenté le travail réalisé au cours du développement de notre système, où les différentes phases du processus de reconnaissance de l'écriture manuscrite qui le constituent ont été détaillées.

Grace à la spécificité de l'écriture arabe et afin d'augmenter le taux de reconnaissance, nous avons effectué une décomposition du vocabulaire selon le nombre de sous-mots constituant le mot. Nous avons extrait les caractéristiques du mot selon plusieurs visions (section IV.7)

Dans le système proposé, nous avons essayé d'adapter les concepts de la logique floue à la classification des noms des wilayas algériennes. Pour cela, nous avons choisi d'appliquer l'algorithme FCM afin de réaliser une classification supervisée des 48 noms des wilayas algériennes en les comparants avec les autres classifieurs implémentés. Nous terminons ce chapitre par l'exposition des différents résultats obtenus par notre système et qui sont encourageants.





## *Conclusion générale*

Comme nous l'avons vu en introduction de ce mémoire, une question reste ouverte : est ce qu'on peut concevoir un système de reconnaissance permettant de déchiffrer un texte manuscrit ? La réponse à cette question, que malgré les efforts et les travaux intensifs réalisés dans ce domaine, aucun système OCR n'est jugé fiable à 100%. Mais au fur et à mesure les auteurs essaient d'améliorer les scores pour de meilleurs résultats.

Cependant les problèmes majeurs influençant la recherche en AOCR dû principalement à la complexité de la morphologie de l'écriture arabe, le manque d'échange entre les chercheurs du domaine et celui d'infrastructures adéquates dans le sens budgétaire, de bases de données, de dictionnaires et d'outils de test et de validation. Mais dernièrement plusieurs travaux ont récemment vu le jour. La résolution de ces problèmes serait d'un apport considérable, tant au niveau simplification de la tâche de l'AOCR, qu'aux niveaux validation et portabilité des produits réalisés.

L'écriture arabe et l'écriture latine cursive ont de nombreux points communs, qui permettent le transfert vers l'arabe de techniques déjà éprouvées sur le latin. Telles que la segmentation en graphèmes, l'extraction de primitives et le moteur de reconnaissance sont les mêmes que ceux utilisés pour la reconnaissance de l'écriture latine. En revanche, certaines adaptations sont indispensables pour mettre en œuvre un système de reconnaissance de l'écriture arabe : L'extraction de la bande de base en particulier doit être adaptée. Notamment en raison de la présence de nombreux signes diacritiques, et également à cause de la forme des ascendants/descendants qui, contrairement à l'usage dans l'écriture latine, peuvent être étendus horizontalement sous la bande de base. Le type de classifieur est également le même que celui utilisé dans le cadre de la reconnaissance latine.

L'objectif principal de toute recherche est d'aboutir à des bons résultats, ce qui conduit à des systèmes performants, mais en focalisant sur le choix de la méthode celle qui donne au mieux que possible la bonne décision, dans le domaine de la reconnaissance de formes on dispose de différentes méthodes telles que les méthodes statistiques, structurelles, les réseaux de neurones,..., avec possibilité de les combinées nous proposons un système multi-classifieurs qui combine parallèlement quatre méthodes de classification: KPPV ; KMEANS; FCM et le réseau PNN.

Ce manuscrit a exposé le travail orienté vers la reconnaissance des mots manuscrits arabe hors ligne (les 48 noms des wilayas algériennes). Deux problématiques ont été abordées, l'extraction des attributs statistiques et morphologiques en analysant l'image du mot par un descripteur globale les moments de Zernike, un descripteur externe le code de Freeman et la méthode de zonage d'une part, et d'autre part, une classification dans l'espace des attributs par quatre méthodes de classification : KPPV, FCM, KMEANS, et PNN. L'objectif principal de ce travail, est de faire une comparaison entre ces différentes méthodes par rapport aux différents attributs extraits.

Toutes les phases classiques d'un processus de reconnaissance ont été réalisées commençant par les prétraitements, l'extraction des caractéristiques et enfin la reconnaissance. Nous avons démontré la procédure d'extraction d'informations fondamentales qui nous aide à la description globale du mot à traiter. Nous avons proposé une nouvelle idée au niveau de cette dernière qui comporte : L'analyse par les moments de Zernike au niveau de l'image binaire (taille de vecteur=100), l'analyse par Freeman faite au niveau du contour (taille de vecteur=8), l'analyse par zonage au niveau du squelette (taille de vecteur=16), de plus les caractéristiques structurelles du mot (taille de vecteur=9). Ainsi nous obtenons un vecteur des attributs composé de 133 caractéristiques identifiant le mot.

Les résultats obtenus par notre approche sont encourageants, notons que nous avons focalisé notre travail sur l'approche parallèle. Cette dernière nécessite l'activation de tous les classifieurs du système qui doivent participer de manière concurrente et indépendante. Par contre, la décision finale est prise avec le maximum de connaissances mises à disposition par chaque classifieur. Dès lors se posent les problèmes de précision des informations fournies par les classifieurs et de la confiance qu'on peut accorder à chacun d'eux. Donc ce travail ne représente qu'un point de départ, nous lui prévoyons des possibilités d'évolution : Nous voyons qu'on peut augmenter les performances de système en appliquant une approche analytique, nous pensons que cette dernière va donner de bons résultats puisque elle prend en compte la lettre arabe comme entité de traitement. Donc le nombre des classes à identifier va diminuer (29 lettres arabe). Mais cette approche nécessite une étape de segmentation qui est de même une tâche plus ardue surtout pour l'écriture arabe manuscrite qui est un peu plus complexe par rapport à la latine, nous espérons dans l'avenir pouvoir intégrer cette approche dans notre système.

## ***BIBLIOGRAPHIE***

---

- [ABU 93] I. S. I. Abuhaiba, P. Ahmed: "Restoration of temporal information in off-line arabic recognition", Pattern Recognition, Vol. 26, N° 7, pp. 1009-1017, July 1993.
- [ABU 94] I.S.I. Abuhaiba, S.A. Mahmoud, R.J. Green: "Recognition of handwritten cursive Arabic characters", IEEE Transactions on PAMI, Vol. 16, N° 6, pp. 664-672, June 1994.
- [ABU 95] I. S. I. Abuhaiba, S. Datta, M.J.J. Holt: "Fuzzy state machines to recognize totally unconstructed handwritten strokes", Image and Vision Computing, Vol. 13, N° 10, pp. 755-769, December 1995.
- [ABU 96] I. S. I. Abuhaiba, M.J. J. Holt, S. Datta: "Processing of binary images of handwritten text documents", Pattern Recognition, Vol. 29, N° 7, pp. 1161-1177, July 1996.
- [ABU 98] I. S. I. Abuhaiba, Holt M. J. J., Datta S., "Recognition of off-line cursive handwriting", Computer Vision and Image Understanding, Vol. 71, N° 1, pp. 19-38, July 1998.
- [ALA 05] Pr. B. ALAIN, Anh. Tuan. NGHIEM : "Reconnaissance d'écriture manuscrite", 15 juillet 2005.
- [AIB 94] B. Al-Badr , R.M. Haralick : « Symbol recognition without prior segmentation ». Conference SPIE-EI 1994.
- [ALB 95a] B. Al-Badr, R. M.Haralick: "Segmentation-free word recognition with application to Arabic". IEEE proceedings of ICDAR'95 Montreal, Canada, 1995.pp 355-359.
- [ALB 95b] B. Al-Badr, S.A. Mahmoud: "Survey and bibliography of Arabic optical text recognition". Signal processing, vol. 41, pp. 49-77, 1995.
- [ALE 90] S.Al-Emami, M. Usher: "On-line recognition of handwritten Arabic characters". IEEE Transactions on Pattern analysis and machine intelligence. Vol. 12. N°7.1990.
- [ALI 97] M. A. Alimi : "An evolutionary neuro-fuzzy approach to recognize on-line Arabic handwriting", The Proceedings of ICDAR'97, 4th International Conference on Document Analysis and Recognition, Vol. 1, pp. 382-386, Ulm, Germany, August 1997.
- [ALI 06] A. M. Alimi, O.A. Ghorbel : "Etude de l'influence du nombre de prototypes dans la reconnaissance en ligne de lettres arabes moulées". CNED' 94. pp 493-499. Rouen 6-7-8 Juillet 1996.
- [ALM 87] H. Almuallim, S. Yamaguchi: "A method of recognition of arabic cursive handwriting", IEEE Transactions on PAMI, Vol. 9, N° 5, pp. 715-722, September 1987.
- [ALM 02] S. Al-Ma'adeed, C. Higgins, D. Elliman: "A database for Arabic handwritten text recognition research, Proceedings of the English IWFHR, Ontario, Canada, pp. 485-489, 2002.
- [ALM 04] S. Al-Ma'adeed, D. Elliman, C. Higgins: "Off-line recognition of handwritten Arabic words using multiple hidden Markov models", Knowledge-Based Systems, Vol. 17, N° 2-4, pp. 75-79, May 2004.
- [ALM 06] S.Al-Ma'adeed: "Recognition of off-line handwritten Arabic words using neural network ", Proceeding of GMAI'06, International Conference on Geometric Modeling and Imaging, pp. 141-114, London, England, July 2006.

- [ALO 02] Y. Al-Ohali: "Handwritten Word Recognition – Application to Arabic Cheque Processing", PhD Thesis, Concordia University, Montreal, Quebec, Canada, February 2002.
- [ALP 97] A. Alper Atici, F. T. Yarman-Vural/ "A heuristic algorithm for optical character recognition of Arabic script", Signal Processing, Vol. 62, N°. 1, pp. 87-99, October 1997.
- [ALR 06] H. Al-Rashaideh/, "Preprocessing phase for Arabic word handwritten recognition", Information Transmissions in Computer Networks, Vol. 6, N°.1, pp. 11-19, 2006. (Disponible sur la toile : [www.jip.ru/2006/11-19-2006.pdf](http://www.jip.ru/2006/11-19-2006.pdf))
- [ALS 06] B. Alsallakh, H. Safadi : "AraPen: an Arabic online handwriting recognition system", Proceeding of ICTTA'06, 2nd IEEE International Conference on Information & Communication Technologies: from Theory to Applications, Vol. 1, pp. 1844-1849, Damascus, Syria, April 2006.
- [ALT 95] M. Altuwaijri, M. Bayoumi: "A new thinning algorithm for Arabic characters using self optimizing neural network", Proceedings of ISCAS'95, IEEE Symposium on Circuits and Systems, Vol. 3, pp. 1824-1827, Seattle, WA, USA, April-May 1995.
- [AMA 96] J.L. Amat, G. Yahiaoui : "Techniques avancées pour le traitement de l'information". Edition CEPADUES 1996.
- [AME 94] A. Ameer, K. Romeo-Pakker, H. Miled, M. Cheriet : "Approche globale pour la reconnaissance de mots manuscrits Arabes", Actes CNED'94, 3ème Colloque National sur l'Écrit et le Document, pp. 151-156, Juillet 1994.
- [AMI 82] A. Amin: "Machine recognition of hand written arabic words by the IRAC II system", Proceedings ICPR'82, 6th International Conference on Pattern Recognition, Vol. 1, pp. 34-36, 1982.
- [AMI 85] A. Amin : "IRAC : un système pour la reconnaissance et pour la compréhension de l'Arabe écrit et imprimé". Thèse d'état, Université de Nancy I. 1985.
- [AMI 96] A. Amin, H.B. Al-Sadoun, and S. Fisher: "Hand-printed Arabic character recognition system using an artificial network ". Pattern recognition, vol. 29, N° 4, pp. 663-675, 1996.
- [AMI 97] A. Amin: "Off line Arabic character recognition - a survey", Proceedings of ICDAR'97, 4th International Conference on Document Analysis and Recognition, Vol. 2, pp. 596-599, Ulm, Germany, August 1997.
- [AMI 00a] A. Amin : "Recognition of printed Arabic text based on global features and decision tree learning techniques", Pattern Recognition, Vol. 33, pp. 1309-1323, 2000.
- [AMI 00b] A. Amin, N. Murshed : "off-line recognition of printed Arabic words through global features and neural networks", 4<sup>th</sup> International workshop on Document Analysis Systems, DAS'2000, Rio de Janeiro, Brazil, pp. 267-277, 2000.
- [AMI 03] A. Amin : "Recognition of hand-printed characters based on structural description and inductive logic programming", Pattern Recognition Letters, Vol. 24, pp. 3187-3196, 2003.
- [ANI 92] J. Anigbogu : " Reconnaissance de textes imprimés mutifontes à l'aide de modèles stochastiques et métriques ". Thèse de doctorat, Université de Nancy I, 1992.
- [ARR 02] D. Arrivault : "Apport des Graphes dans la Reconnaissance Non-Contrainte de Caractères Manuscrits Anciens", (Université de Poitiers)(Diplôme National - Arrêté du 25 Avril 2002).

- [AVI 96] M.Avila : "Optimisation des modèles Markoviens pour la reconnaissance de l'écrit", Thèse de Doctorat, Université de Rouen, 1996.
- [AYA 04] N. Ayat : "Sélection De Modèle Automatique Des Machines À Vecteurs De Support: Application À La Reconnaissance D'images De Chiffres Manuscrit", thèse présentée à l'École de Technologie Supérieure comme Exigence Partielle à l'obtention du Doctorat en génie P.H.D. Montréal, Le 20 Janvier 2004.
- [AZI 02a] N. Azizi : "Combinaison de classifieurs neuronaux basée sur la logique floue : application à la reconnaissance des mots arabes manuscrits", Mémoire de magister, Laboratoire LRI, Département d'informatique, Université d'Annaba, Algérie, 2002.
- [AZI 02b] N. Azizi, T. Sari, L. Souici-Meslati, M. Sellami : "Une architecture de combinaison floue de classifieurs neuronaux pour la reconnaissance de mots arabes manuscrits", CIFED'02, 7ème Colloque International Francophone sur l'Écrit et le Document, pp. 89-96, Hammamet, Tunisie, Octobre 2002.
- [BEL 91] S.O. BELKASIM, M. SHRIDHAR et M. AHMADI: "Pattern recognition with moment invariants: A comparative study and new results". Pattern Recognition, Vol. 24, N°12, pp. 1117–1138, 1991.
- [BEL 92] A. Belaid, Y. Belaid : "Reconnaissance des formes: Méthodes et applications ", InterEditions, jan-vier 1992.
- [BEL 97] A. Belaid et G. SAON : "Utilisation des processus markoviens en reconnaissance de l'écriture". Revue Traitement du Signal, Vol.14, N°2, pp. 161–177, 1997.
- [BEL 01] A. Belaid : "La reconnaissance automatique de l'écriture et du document", pour la Science, octobre 2001.
- [BEN 98] N. Benamara, A. Belaid, N. Ellouze : "Modélisation pseudo bidimensionnelle pour la reconnaissance des chaînes de caractères arabes imprimées". Proc. 1er Colloque International francophone sur l'écrit et le document (CIFED'98), pp. 131-140, Québec, Canada, 1998.
- [BEN 99] N. Benamara : "Utilisation des modèles de Markov cachés planaires en reconnaissance de l'écriture arabe imprimée ". Thèse de doctorat, spécialité Génie Electrique, Université des sciences, des Techniques et de médecine de Tunis II, 1999.
- [BEN 00] A. Benouareth : " Reconnaissance de l'écriture arabe manuscrite par une approche hybride ". Mémoire de magister, Labo. LRI, Département d'informatique, Université d'Annaba, Algérie, Février 2000.
- [BEN 01a] N. Ben Amara, A. Belaid : "Une Methode Stochastique pour la Reconnaissance de L'écriture Arabe Imprimée ". Ecole Nationale d'Ingénieurs de Monastir 5019 Monastir – Tunisie, Centre de Recherche en Informatique de Nancy Bâtiment Loria, Campus Scientifique- B.P. 239 54506 Vandoeuvre-Lès-Nancy- France. Soumis en 2001.
- [BEN 01b] N. Ben Amara, A. Belaid et N. Ellouze : " Utilisation des modèles markoviens en reconnaissance de l'écriture arabe : Etat de l'art". Tunisie. Soumis en 2001.
- [BEN 02] N.E. Ben Amara : "Sur la problématique et les orientations en reconnaissance de l'écriture arabe", Proceedings of the Colloque International Francophone sur l'Écrit et le Document, CIFED'2002, Hammamet, Tunisia, October 2002, pp. 2-10.
- [BEN 04] A. Bensefia : "Analyse des documents manuscrits : identification et vérification du scribe". Thèse de Doctorat, Laboratoire LITIS, Université de Rouen, Mai 2004.

- [BEN 05] A. Benouareth, A. Ennaji, M. Sellami : " Utilisation des HMMs de Durée d'Etat Explicite pour la Reconnaissance des Mots Arabes Manuscrits". Article soumis en 2005.
- [BEN 07] A. Benouareth : "Reconnaissance de Mots Arabes Manuscrits par Modèles de Markov Cachés à Durée d'Etat Explicite". Thèse Présentée en vue de l'obtention du diplôme de Doctorat, 2007/2008.
- [BIS 95] C. Bishop: "Neural Networks for Pattern Recognition", University Press, Oxford, Great Britain, 1995.
- [BON 96] N. Bonnert, M. Herbin : "Classification automatique", Séminaire LERI – Université de Reims – Février 1996.
- [BOR 97] A. Boris: "A Thalamocortical Algorithm That Performs Handwritten Character Recognition", dans A.C. Downton & S. Impedovo (éditeurs): « Progress in Handwriting Recognition », publié par World Scientific. 1997.
- [BOZ 89] R.M. Bozinovic and S.N. Srihari : " Off-line cursive script word recognition ", IEEE Transaction on Pattern Analysis and Machine Recognition PAMI, Vol 11, NO. 1, pp: 68-83, January, 1989.
- [BRE 03] S. Bres, J.M. Jolion, F. Lebourgeois : "Traitement et analyse des images numériques". Hermes, 2003
- [BRO 83] M.K. Bown: «Preprocessing techniques for cursive script word recognition", Pattern Recognition, Vol. 16, N°. 5, pp. 447-458, 1983.
- [BUN 01] H. Bunke: "Recent advances in structural pattern recognition with applications to visual form analysis". In Workshop on Visual Form, page 11 ff., 2001.
- [BUR 04] P. Burrow: "Arabic handwriting recognition". Master of Science thesis. School of Informatics, university of Edinburg, England, 2004.
- [CAS 96] R. G. Casey, E. Lecolinet: "A survey of methods and strategies in character segmentation", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 18, No. 7, pp. 690–706, July 1996.
- [CHE 94] M.Y. Chen, A. Kundu, J. Zhou: "off-line handwritten word recognition using a hidden markov model tupe stochastic network", IEEE Transactions on PAMI, Vol. 16, No. 5, pp. 4
- [CHU 00] S. Chuai-Aree, C. Lursinsap, P. Sophatsathit and S. Siripant: "Fuzzy C-Mean: A statistical feature classification of text and image segmentation method", Proc. of Intern. Conf. on Intelligent Technology 2000, December 13-15, Assumption University Bangkok, Thailand, pp. 279-284, 2000.
- [COH 94] E. Cohen, J.J. Hull, S.N. Srihari: "Control-structure for interpreting hand-written addresses". IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 16, N°.10, pp. 1049–1055, octobre 1994.
- [DAR 94] P. Dargenton : "Contribution à la segmentation et à la reconnaissance de l'écriture manuscrite", Thèse de Doctorat, Lyon, 1994.
- [DUI 00] R.P.W. Duin and D.M.J. Tax: "Experiments with Classifier Combining Rules". Multiple Classifier Systems (Proc. First International Workshop, MCS 2000, Cagliari, Italy, June 2000), Lecture Notes in Computer Science, vol. 1857, Springer, Berlin, 2000, 16-29, 2000.
- [DUP 04] X. Dupré : "Contribution à la reconnaissance de l'écriture cursive à l'aide de Modèles de Markov Cachés". Thèse de Doctorat, université Réni Descartes – Paris V, 2004.

- [ELD 90] S. S. El-Dabi, R. Ramsis, A. Kamel: "Arabic character recognition system: a statistical approach for recognizing cursive typewritten text". Pat. recog., Vol. 23, No. 5, pp. 485-495 1990.
- [ELH 05] R. El-Hadj, L. Likeforman-Seulem, C. Mokbel: "Arabic handwriting recognition using baseline dependant features and hidden Markov models", Proceedings of ICDAR'05, International Conference on Document Analysis and Recognition, Vol 2., pp. 893-897, Seoul, Korea, August-September 2005.
- [FAH 01] M.M.M. Fahmy, S.Al Ali: "Automatic recognition of handwritten Arabic characters using their geometrical features ". Studies in informatics and control journal (SIC journal), vol. 10, No 2, 2001.
- [FAN 95] M. FANTON : "Ligatures et informatique ". Cahiers Gutenberg, no 22, September 1995, p. 61-85.
- [FAR 05] N. Farah, A. Ennaji, T. Kadir, M. Sellami: "Benefit of multiclassifier systems for Arabic handwri-tten words recognition", ICDAR'05, 8<sup>th</sup> International Conference on Document Analysis and Recognition, Vol. 1, pp. 222-226, Seoul, Korea, August-September 2005.
- [FAR 06] N.Farah, L. Souici, M. Sellami : "Classifiers combination and syntax analysis for Arabic literal amount recognition", Engineering Applications of Artificial Intelligence, Vol. 19, N°. 1, pp. 29-39, February 2006.
- [FUK 90] K.Fukunaga: «Introduction to Statistical Pattern Recognition Academic Press". San Diego , 1990.
- [GAA 01] N.Gorski, V. Anisimov, E. Augustin, O. Baret ET S. Maximov: "Industrial bank check processing: the A2iA CheckReader™". International Journal on Document Analysis and Recognition, 3(4):196–206, 2001.
- [GAS 05] I. Gasmi, H. Merouani, L. Souici-Meslati: "Combinaison de classifieurs". SETIT 2005 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications March 27-31, 2005 – TUNISIA
- [GON 78] R.C. Gonzalez ET M.G. Thomason: "Syntactic Pattern Recognition: An Introduction". Addison-Wesley, 1978.
- [GOR 92] H .Goraine, S. Al Emami: "Off-line arabic character recognition", IEEE Computer Journal , pp. 71-74, July 1992.
- [GOS 96] B. Gosselin : " Application de réseaux de neurones artificiels a la reconnaissance automatique de caractères manuscrits", 95/96.
- [GOW 90] K.Gowely,O. El Dessouki, A. Nazif/ "Multi-phase recognition of multifont photocopied Arabi text", Proceedings of ICPR'90, 10<sup>th</sup> International Conference on Pattern Recognition, Vol. 1, pp. 700-702, Atlantic City, NJ, USA, June 1990.
- [GRA 03] F. Grandidier : "Un nouvel algorithme de sélection de caractéristiques-Application à la lecture automatique de l'écriture manuscrite ". Thèse de Doctorat, Ecole de Technologie Supérieure, Université du Québec, Canada, Janvier 2003.
- [GUI 95] D. Guillevic: "Unconstrained Handwriting Recognition Applied to the Processing of Bank Cheque", Doctorat thesis, Computer Science Department, Concordia University, Montreal, September 1995.
- [GUI 01] S.Guillaume : "Induction de règles floues interprétables", Thèse de Doctorat Laboratoire d'analyse et d'architecture des systèmes du CNRS Toulouse – Novembre 2001

- [GUO 92] Z. Guo,R. Hall:"Fast fully parallel thinning algorithm", Computer Vision, Graphical and Image Processing, Vol. 55, N°. 3, 317-328, May 1992.
- [HAI 07] S.Haitaamar : "Segmentation de Textes en Caracteres pour la Reconnaissance Optique de l'écriture Arabe", Université El-Hadj Lakhdhar Batna, Pour l'obtention du diplôme de Magister en informatique Option Informatique Industrielle. Soutenue le 08 Juillet 2007.
- [HAS 06] A. Hassaine : "Codage De Graphemes Et Compression Sans Perte D'images De Manuscrits Anciens". Master Recherche II IGI : Spécialité Informatique Graphique et Image.
- [HA 96] T.M. Ha, G. Kaufmann, H. Bunke: " Text localization and handwriting recognition", Technical report, university of Berne, 1996.
- [HEU 03] L. Heutte : "Analyse et Reconnaissance de l'écriture : de Nouvelles Perspectives en Traitement Automatique de Documents Manuscrits". Habilitation à Diriger les Recherches, Université de Rouen, 2003.
- [HEU 05] L.Heutte : "Combinaison de Classifieurs Pourquoi et comment les combiner? " Université de Rouen, 2005.
- [HO 92] T. K. Ho: "A theory of multiple classifier systems and its application to visual word recognition". PhD, Faculty of the Graduate School of State University of New-York at Buffalo, May, 1992.
- [HUS 93] D. R. Hush, B. G. Horne: «Progress in Supervised Neural Networks". IEEE Signal Processing Magazine, pp 8-39, Janvier 1993
- [IMP 97] S. Impedovo, P.S.P. Wang, H. Bunke (Editors): "Automatic bankcheck processing", Séries in Machine Perception and Artificial Intelligence, Vol. 28, World Scientific, 1997.
- [JAI 00] A. K. Jain, R.P.W. Duin, J. Mao: "Statistical pattern recognition ". A Review, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No 1, January 2000.
- [JAN 92] B. Jang, R. Chin/:"One-pass parallel thinning analysis, properties, and quantitative evaluation", IEEE Transactions on PAMI, Vol. 14, N°. 11, pp. 1129-b1140, November 1992.
- [KAR 06] H. Karim : "Une étude de l'évolutivité des modèles pour la reconnaissance de documents arabes dans un contexte interactif ". Thèse de doctorat soumise à la Faculté des Sciences de l'Université de Fribourg (Suisse) pour l'obtention du grade de Doctor Scientiarum Informaticarum, Thèse n°1513 Imprimerie St Paul, Fribourg 2006
- [KAU 00] G. Kaufmann ET H. Bunke: "Automated reading of cheque amounts". Pattern Analysis & Applications, 3(2):132–141, 2000.
- [KB 00] G. Kaufmann et H. Bunke: "Automated reading of cheque amounts", Pattern Analysis & Applications, 3(2):132–141, 2000.
- [KER 00] S. Kermi : "Classifieur neuronal base connaissances, application à la reconnaissance des caractères arabes isolés manuscrits ". Mémoire de Magister, Université Badji Mokhtar, Annaba, Algérie 2000.
- [KHP 05] G. Koch, L. Heutte ET T. Paquet: "Automatic extraction of numerical sequences in handwritten incoming mail documents". Pattern Recognition Letters, 26(8):1118–1127, juin 2005.
- [KHA 99] N. Kharma, R. Ward: "A New Comprehensive Database of Hand6Written Arabic Words , Numbers, and Signatures used for OCR Testing". Proceeding of the IEEE Canadian

- Conference on Electrical & Computer Engineering, Edmonton, Canada, pp. 766-768, 1999.
- [KHO 90a] A.Khotanzad ET Y.H. Hong: "Invariant image recognition by zernike moments". IEEE Trans. Pattern Analysis and Machine Intelligence, 12(5):489–497, mai 1990.
- [KHO 90b] A.Khotanzad ET Y.H. Hong: "Rotation invariant image recognition using features selected via a systematic method". Pattern Recognition, 23(10):1089–1101, 1990.
- [KHO 00] M. S. Khorsheed: "Automatic recognition of words in Arabic manuscripts". PHD Theis, Churchill College, University of Cambridge, UK, Also available as University of Cambridge, Computer Laboratory Technical Report No. 495, 200.
- [KHO 03] M. S. Khorsheed:"Recognising handwritten Arabic manuscripts using a single hidden Markov model", Pattern Recognition Letters, Vol. 24, N°.14, pp. 2235- 2242, October 2003.
- [KIM 00] H.K. Kim ET J.D. Kim: "Region-based shape descriptor invariant to rotation, scale and translation ". Signal Processing: Image Communication, 16(1-2):87–93, septembre 2000.
- [KIT 98] J. Kittler, M. Hatef, R. Duin et J. Matas : "On Combining Classifiers", IEEE Trans. On Pattern Analysis and Machine Intelligence, 226-238, 1998.
- [KUN 01] L. I. Kuncheva, C. J. Bezdek, R.P.W. Duin: "Decision templates for multiple classifier fusion: an experimental comparison". Pattern Recognition 34, pp. 299-314, 2001.
- [LAA 97] J.Laaksonen: "Subspace Classifiers in Recognition of Handwritten Digits". Computer science, Helsinki University of Technology, Finland, May 1997.
- [LAR 98] J. Larmagnac: "Thinning and line segmentation by line following technique", Proceedings of SPIE'98, the International Society of Photo-Optical Instrumentation Engineers, Vol. 3305, pp. 210-219, San Jose, CA, USA, 1998.
- [LEC 93] Y. Lecourtier, B. Dorizzi, P. Sebire, A. Ennaji: "MLP Modular Versus YPR--EL Classifiers". IWANN 1993: 569-574.
- [LEC 94] E.Lecolinet, O. Baret: "Cursive word recognition: Methods and strategies ". in Fundamentals in handwriting recognition, S. Impedovo Ed., pp. 235-263, NATO ASI Series F: Computer and Systems Sciences, Vol. 124, Springer Verlag, 1994.
- [LET 95] E. Lethelier, M. Leroux ET M. Gilloux: "An automatic reading system for handwritten numeral amounts on frenches checks. In 3th International Conference on Document Analysis and Recognition", Montreal, Canada, pages I :92–97, 1995.
- [LEY 04] Y. Leydier, F. Le Bourgeois et H. Emptoz : "Sérialisation du k-means pour la segmentation des images en couleur". Dans Actes du 8ième colloque international francophone sur l'écrit et le document, CIFED, 21-25 juin 2004, La Rochelle. pp. 191-196. 2004.
- [LIP 87] R.P Lippmann: «An introduction to computing with neural nets ". IEEE, ASSP magazine, April 1987.
- [LLG 95] E. LETHELIER, M. LEROUX et M. GILLOUX : An automatic reading system for handwritten numeral amounts on french checks. In *3th International Conference on Document Analysis and Recognition, Montreal, Canada*, pages I :92–97, 1995.
- [LOA 99] H. Loaiza : "Capteur de Vision Stéréoscopique Multiconfiguration Pour la Localisation D'un RobotMobile", Thèse Université d'Evry Val d'Essonne, 1999.
- [LON 98] S. Loncaric: "A survey of shape analysis techniques". Pattern Recognition, 31(8):983–1001, août 1998.

- [MAD 97] S. Madhvanath et V. Krpasundar: "Pruning large lexicons using generalized word shape descriptors ". In 4th International Conference on Document Analysis and Recognition, Ulm, Allemagne, page Poste, 1997.
- [MAH 91] S.A. Mahmoud, I. Abuhaiba,R.J. Green/ "Skeletonization of Arabic characters using clustering based skeletonization algorithm (CBSA)", Pattern Recognition, Vol. 24, N°. 5, pp. 453-464, 1991.
- [MAH 94] S.A. Mahmoud: "Arabic character recognition using Fourier descriptors and character contour encoding ". Pattern recognition, vol. 27, No 6, pp. 815-824, 1994.
- [MAS 00] S. Masmoudi, H. Amiri/"Reconnaissance de mots arabes manuscrits par modélisation markovienne", Actes CIFED'00, 6ème Colloque International Francophone sur l'Écrit et le Document, pp. 473-482, Lyon, France, 2000.
- [MER 04] D. Merad : "Reconnaissance 2D/2D et 2D/3D d'objets à partir de leurs squelettes", thèse de Doctorat Spécialité : Robotique, Soutenue le 13 Décembre 2004.
- [MEZ 05] N. Mezghani : "Densités de probabilité d'entropie maximale et mémoires associatives pour la reconnaissance en ligne de caractères Arabes", Thèse de Ph.d, Institut national de la recherche scientifique (INRS), centre Énergie, Matériaux et Télécommunications, Canada, Juin 2005.
- [MIL 98] H. Miled : "Stratégies de résolution en reconnaissance de l'écriture semi-cursive : application aux mots manuscrits arabes", Thèse de doctorat, Université de Rouen, 1998.
- [MOH 96] M. Mohamed and P. Gader: «Handwritten word recognition using segmentation-free hidden Markov modeling and segmentation-based dynamic programming techniques», IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 5, pp. 548-554, 1996.
- [MON 95] L. Montoliu : "Architecture multiagents et réseaux connexionnistes : Application à la lecture de chèques manuscrits". Thèse de Doctorat, Laboratoire LIX, Ecole polytechnique, Palaiseau, France, 1995.
- [MOO 96] B. Moobed: "Combinaison de classifieurs: une nouvelle approche", Thèse de Doctorat, Laboratoire LIX, Ecole Polytechnique, Palaiseau, 1996.
- [MOR 03] M. E. Morita : " Automatic recognition of handwritten dates on brazilian bank cheques", Thèse de Doctorat, Ecole de Technologie Supérieure, Université du Québec, Canada, 2003.
- [NEM 09] S. Nemouchi, N. Farah : " Reconnaissance de l'Écriture Arabe par systèmes Flous", Proceedings de la Conférence nationale en Informatique, JEESI'09, Alger, Algérie, Vol.6, N°.76, Mai 2009.
- [NEM 10] S. Nemouchi, N. Farah:"The Recognition of Arabic Writing by Fuzzy System" , Proceeding IEEE International Conference on Computer and Communication Engineering ICCCE 2010, Malaysia, Vol.6, N°.45, May, 2010.
- [OLL 99] B.Ollivier : "Une approche économisant les traitements pour reconnaître l'écriture manuscrite : application à la de montants littéraux de chèques bancaires", Thèse en Informatique, Laboratoire LIX, Paris, 1999.
- [OSO 98] F. S. Osorio : "INSS: un système hybride neuro-symbolique pour l'apprentissage automatique constructif", Thèse en Informatique, Laboratoire LEIBNIZ\_IMAG, INPG-Grenoble, 1998.

- [OTS 79] N. Otsu: "Athreshold selection method from gray-level histogram ", IEEE Transactions on systems, man and cybernetics, Vol. 9, N°. 1, pp. 62-66, 1979.
- [PAQ 92] T. Paquet : "Segmentation et classification de mots en reconnaissance optique de textes manuscrits", Thèse de Doctorat, Université de Rouen, 1992.
- [PAR 89] J. Park, V. Govindaraju, S. Srihari : "Efficient word segmentation driven by unconstrained handwritten phrase recognition", in : Proceedings of 5<sup>th</sup> International Conference on Document Analysis and Recognition, Vol. 1, Bangalore, pp. 605-608, 1989.
- [PAV 78] T. Pavlidis: "A review of algorithms for shape analysis". Computer Graphics Image Processing, Vol. 7, N°.2, pp. 243-258, Avril 1978.
- [PAV 81] T. Pavlidis: "A Flexible Parallel Thinning Algorithm". Proceeding IEEE Comput. Soc. Conference. on Pattern Recognition and Image Processing. Août. 1981, pp. 162-167.
- [PAV 82] T. Pavlidis: "Algorithms for Graphic and Image Processing", Rockville, MD: Computer science press, 1982.
- [PEC 03] M. Pechwitz, V. Märgner, "HMM based approach for handwritten Arabic word recognition using the IFN/ENIT-database", Proceeding of ICDAR'03, 7<sup>th</sup> International Conference on Document Analysis and Recognition, Vol. 2, pp. 890-894, Edinburgh, Scotland, 2003.
- [PEC 06] M. Pechwitz, V. Märgner, H. El Abed: "Comparison of two different feature set for off-line recognition of handwritten Arabic words", Proceedings of IWFHR'06, 10<sup>th</sup> International Workshop on Frontiers in Handwriting Recognition, pp. 109-114, La Baule, France, October 2006.
- [QUI 86] J. R. Quinlan: "Induction of Decision Trees ", dans Machine Learning, 1:81-106, 19.
- [RAB 89] L. Rabiner : "tutorial on Hidden Markov Models and selected applications in speech recognition", in A. Waibel, L. Kai-Fu (Eds), Readings in Speech Recognition , Morgan Kaufmann, Palo Alto, CA, pp. 267-296, 1989.
- [REI 69] G.M. Reicher: "Perceptual recognition as a function of meaningfulness of stimulus material". Journal of Experimental Psychology, Vol.81, pp. 275-280, 1969.
- [ROM 95] K Romeo-Pakker, H. Miled, Y. Lecouturier: "A new approach for Latin/Arabic character segmentation". Proc. IEEE. ICDAR'95 Montreal, Canada, pp 874-877, 1995.
- [SAA 85] S. Saadallah, S. G. Yacu : "Design of an Arabic character processing and transmission of the Arabic language". Kuwait 1985.
- [SAO 98] G. Saon : "Modèles Markoviens uni-bidimensionnels pour la reconnaissance de l'écriture manuscrite hors-ligne". Thèse de doctorat, Université Henri Poincaré Nancy 1, 1997.
- [SAR 00] T. Sari : "un système de reconnaissance de mots arabes manuscrits basé segmentation ", Mémoire de Magistère, Laboratoire LRI, Département d'Informatique, Université Badji Mokhtar, Annaba, Algérie, 2000.
- [SAR 05] Sari T., Sellami M., "Cursive Arabic script segmentation and recognition system", International Journal of Computers and Applications, Vol. 27, N°. 3, 2005.
- [SAU 97] J. Sauvola, T. Seppänen, S. Haapakoski, ET M. Pietikainen: "Adaptive Document binarization", 4<sup>th</sup> International Conference on Document Analysis and Recognition, Ulm, Allemagne, Août, 1997.
- [SAY 73] K.M. Sayre: "Machine recognition of handwritten words: a project report ". Pattern recognition, vol.5, pp. 213-228, 1973.

- [SEN 98] A. W. Senior, J. Rbinson: "AN off-line cursive handwriting recognition system", IEEE Transac-tion on Pattern Analysis and Machine Intelligence 20(3) (1998) 309-327.
- [SEY 99] K. Seymore, A. McCallum, R. Rosenfeld: "Learning Hidden Markov model structure for informa-tion extraction ". AAAI. Workshop on machine learning for information extraction, pp. 37-42, 1999.
- [SIM 84] J.C. Simon : "La Reconnaissance des Formes par Algorithmes". Masson, 1984.
- [SIM 92] J.C. Simon: "Off-line cursive word recognition", Proceedings of IEEE, Vol. 80, N°. 7, pp. 1150-1161, July 1992.
- [SIN 99] S. Singh, A. Amin : "Fuzzy Recognition of Chinese Characters ", Proc. Irish Machine Vision and Image Processing Conference (IMVIP'99), Dublin, 8-9 September, 1999.
- [SNO 02a] S. Snoussi-Maddouri,H. Amiri,A. Belaid,C. Choisy: "Combination of Local and Global Vision Modeling for Arabic Handwritten Word Recognition", International Workshop Frontier in Handwriting IWFHR'02, Canada, pp. 128-135, 2002.
- [SNO 02b] S.Snoussi-Maddouri,A. Belaid,C. Choisy,H. Amiri : " Modèle perceptif neuronal à vision globale-locale pour la reconnaissance de mots manuscrits arabes", CIFED'02, Colloque International Francophone sur l'Ecrit et le Document, pp. 11-20, Hammamet, Tunisie, Octobre 2002.
- [SOU 97] L. Souici, Z. Zmirli, M. Sellami : " Système connexionniste pour la reconnaissance de l'arabe manuscrit ". 1ères journées scientifiques et techniques (JST FRANCIL), pp. 383-388, Avignon, France, 1997.
- [SOU 06a] L. Souici-Meslati : "Reconnaissance des mots arabes manuscrits par intégration neuro-symbo-lique", Thèse de Doctorat d'Etat, Labo. LRI, Département d'informatique, Université d'Annaba, Algérie, Février 2006.
- [SOU 06b] L. Souici-Meslati, M. Sellami: "Toward a generalization of neuro-symbolic recognition: An application to arabic words". KES, International Journal of Knowledge-Based and Intelligent Engineering Systems, Vol.10, N°.5, pp. 347-361, 2006.
- [SRI 97] S. Srihari: «Recent Advances in Off-line Handwriting Recognition" at CEDAR, dans A .C. Down-ton & S. Impedovo (éditeurs) : "Progress in Handwriting Recognition ", publié par World Scientific, 1997.
- [STE 99] T.Steinherz, E. Rivlin, N. Intrator: "Off-line cursive word recognition: a survey ". IJDAR, International journal on document analysis and recognition, Vol 2, pp. 90-110, 1999.
- [SUE 98] C. Y. Suen : "Réflexions sur la reconnaissance de l'écriture cursive", CIFED'98, premier Congrès International Francophone sur l'Ecrit et le document, Vol. 14, pp : 1-8, Québec, Canada, Mai 1998.
- [TEA 80] M. R. Teague: "Image analysis via the general theory of moments". Journal of the Optical Society of America, Vol.70, N°.8, pp. 920-930, Août 1980.
- [THE 88] C.H. The, R.T. Chin: "On image analysis by the methods of moments". IEEE Transactions Pattern Analysis and Machine Intelligence, Vol.10, N°.4, pp.496-513,Jjuillet 1988.
- [TOL 90] M. F. Tolba, E. Shaddad: "On the automatic reading of printed Arabic characters", IEEE International Conference on Systems, Man and Cybernetics Conference Proceedings, pp. 496 498, 1990.

- [TRE 97] J.Trenkle, A. Gillies, S. Schlosser: "An off-line Arabic recognition 136 system for machine printed documents ". Proceeding Of the symposium on document image understanding technology (SDIUT'97), pp. 155-161 1997.
- [TRE 04] A. Trémeau, C. Fernandez-Maloigne et P. Bonton, Image numérique couleur, de l'acquisition au traitement, DUNOD, pp. 300.
- [TRI 96] O.D. Trier, A.K. Jain et T. Taxt: "Feature-extraction methods for character recognition ": A survey. Pattern Recognition, Vol. 29, N°. 4, pp. 641–662, Avril 1996.
- [VIN 02] A. Vinciarelli: "A survey on off-line cursive word recognition". Pattern Recognition, Vol. 35, N°.7, pp. 1433–1446, juillet 2002.
- [XU 92] L. Xu, A. Krzyzak, C. Y. Seun: "Methods of combining multiple classifiers and their applications to handwriting recognition", IEEE Transactions on Systems, Man, And Cybernetics, Vol. 22, N°. 3, pp.418-485, may/june 1992.
- [YU 90] S.Yu S, W. Tsai: "A new thinning algorithm gray-scale images by the relaxation technique " , Pattern Recognition, Vol. 23, N°. 10, pp. 1067-1076, 1990
- [ZAD 00] L.A. Zadeh: "From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions". In Intelligent Systems and Soft Computing: Prospects, Tools and Applications, pp. 3–40, London, UK, 2000. Springer-Verlag.
- [ZAH 90] A.Zahour : "Une méthode de reconnaissance de l'écriture arabe cursive", Thèse de Doctorat, Université du Havre, France, 1990.
- [ZER 34] F. Zernike: "Diffraction theory of the cut procedure and its improved form, the phase contrast method". Physica, 1:689–704, 1934.
- [ZER 07] N. Zermi, M. Ramdani, M. Bedda: "Arabic handwriting word recognition based on hybride HMM/ANN approach", International Journal of Soft Computing, Vol. 2, N°. 1, pp. 5-10, 2007.
- [ZGH 02] R. Zeghibi : "Le codage informatique de l'écriture arabe", d'ASMO 449 à Unicode et ISO/CEI 10646, Document numérique, Vol. 6, N°. 3, pp. 155-182,2002.
- [ZOU 02] H. Zouari, L. Heutte, Y. Lecourtier, A. Alimi : "Un panorama des méthodes de combinaison de classifieurs en reconnaissance de formes". RFIA'2002, 11<sup>ème</sup> Congrès francophone AFRIFAFIA de Reconnaissance des formes et Intelligence Artificielle, Vol. 2, pp. 499-508, Angers, France, janvier, 2002.
- [ZOU 04] H. Zouari, L. Heutte, Y. Lecourtier, and A. Alimi: "Building diverse classifier outputs to evaluate the behavior of combination methods: the case of two classifiers. Multiple Classifier Systems", Lecture Notes in Computer Science, Springer, LNCS 3077, pp. 237-282, 2004.

***FORMULAIRES DE LA BASE DE DONNEES  
DE MOTS ARABES MANUSCRITS***

## ***DESCRIPTION DE SYSTEME REALISE***

### **B.1. Introduction**

Après une description détaillée de notre système et les différents algorithmes utilisés pour le modéliser, on parlera de l'environnement de développement choisi pour réaliser notre application (reconnaissance de mots arabes manuscrits).

Le système réalisé est un système de reconnaissance des 48 noms des wilayas algériennes basé sur une approche globale « mulicalssifieurs », il comporte la plupart des phases d'un processus de la reconnaissance de l'écriture manuscrite, voire Figure (IV.1) de chapitre IV.

Dans cette annexe, nous allons présenter l'environnement de développement de notre système, ensuite nous allons démontrer le fonctionnement du système et les différentes étapes qui permettent leur mise en œuvre.

### **B.2. Environnement MATLAB VERSION 7.7**

L'application a été développée sous l'environnement Windows XP avec le langage de programmation MATLAB 7.7.

Le nom de MATLAB dérive du terme Matrix Laboratory. Il s'agit d'un outil de développement pour des problèmes scientifiques et, plus généralement pour tous les domaines où des calculs numériques importants doivent être faits.

Le langage de programmation Matlab consiste en un langage interprété qui s'exécute dans une fenêtre dite d'exécution. L'intérêt de Matlab tient, d'une part, à sa simplicité d'utilisation : pas de compilation, déclaration implicite des variables utilisées et, d'autre part, à sa richesse fonctionnelle : arithmétique matricielle et nombreuses fonctions de haut niveau dans de nombreux domaines (analyse numérique, graphique, ...). La programmation sous Matlab consiste à écrire des scripts de commandes Matlab, exécutables dans la fenêtre d'exécution. En outre, grâce aux diverses Toolboxes spécialisés. Dans notre système nous intéressons au traitement d'image : Filtrage et conception de filtres 2-D - Amélioration et restauration d'images, opérations morphologiques, géométriques, et coloration, transformations 2-D, Analyse d'images et statistiques.

### **B.3. Fonctionnement du système**

Notre système se divise en quatre grands modules distincts :

- ♣ Module de prétraitement.
- ♣ Module d'extraction de caractéristiques.
- ♣ Module de classification :
  - ◆ Les K Plus Proches Voisins (KPPV).
  - ◆ K-MEANS
  - ◆ Fuzzy-C-Means (FCM).
  - ◆ Réseau de neurones (PNN).

- ❖ Module de combinaison :
  - ◆ Par vote de majorité.
  - ◆ Par priorité.

### B.3.1. Page d'accueil

Cette fenêtre permet une première interaction de l'utilisateur avec notre système. Elle contient cinq menus **Fichier**, **Prétraitement**, **Analyse par**, **Classifier par**, **combinaison**. Plus d'un bouton **Close** pour quitter le système

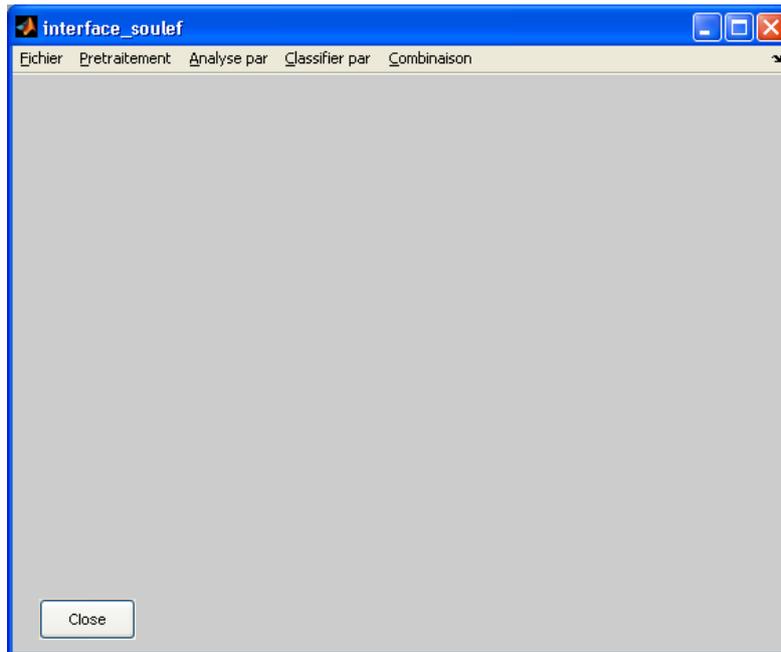


Figure B.1. Page d'accueil de système.

Les autres interfaces réalisées sont utilisées principalement de manière interactive. Cet ensemble de menus mis à la disposition de l'utilisateur afin de mieux guider ses choix. Ils fonctionnent de la même manière que les menus des programmes Windows. Chaque menu est une liste de commandes parmi lesquelles, l'utilisateur réalise son choix à l'aide de la souris. Il commence d'abord par choisir le menu qu'il souhaite ouvrir et sélectionne ensuite, la commande qu'il veut dans ce menu.

- ❖ **Fichier** : contient deux sous menus :
  1. **Ouvrir** : pour choisir la partition la où il y a le dossier des images sauvegardées.
  2. **Quit** : pour fermer la fenêtre et sortir du logiciel.

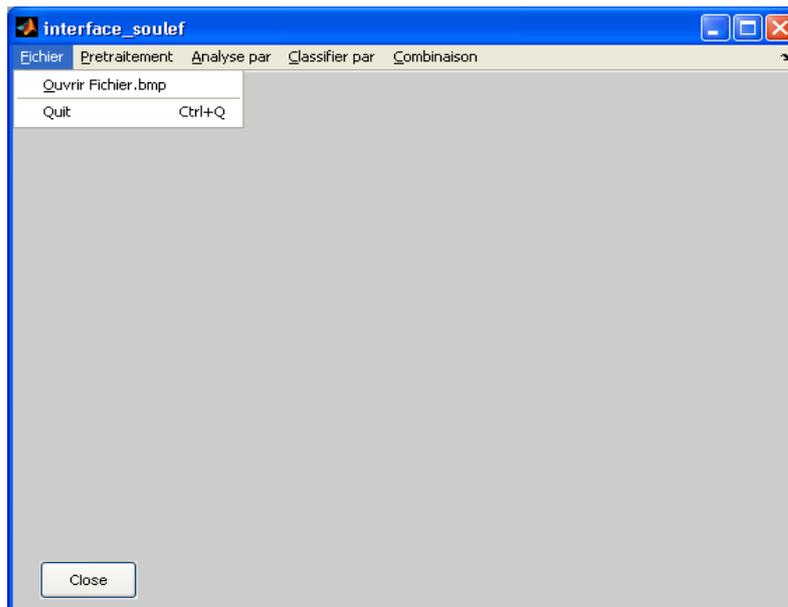


Figure B.2. Interface de menu fichier.

- ❖ **Prétraitement** : Après avoir sélectionné l'image du mot, l'utilisateur peut lancer la procédure de traitement d'image (Binarisation, Lissage, Squelette, Extraction Contour).

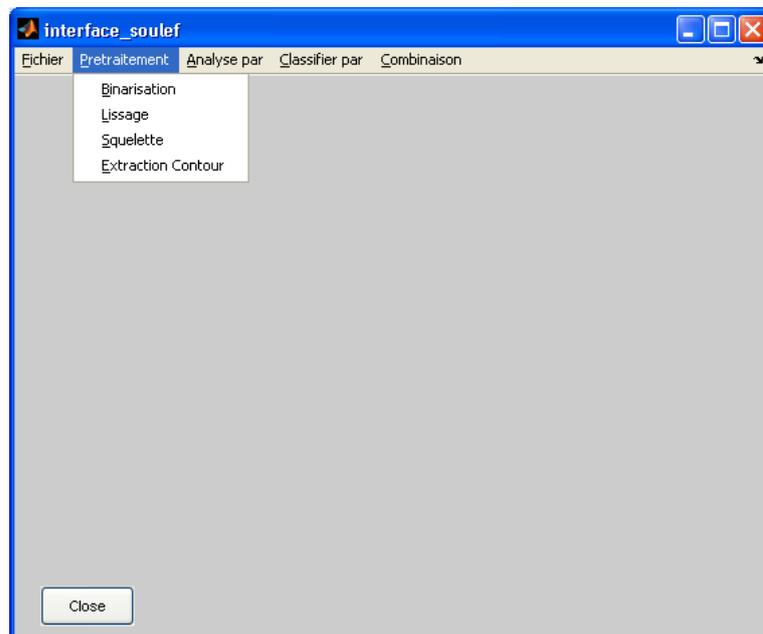


Figure B.3. Interface de menu prétraitement.

- ❖ **Analyse par** : permet d'extraire les caractéristiques d'une image prétraitée, en choisissant le type de caractéristiques voulu d'extraire (Extraction des caractéristiques : structurelles, Zonage, Zernike ou Code de Freeman).

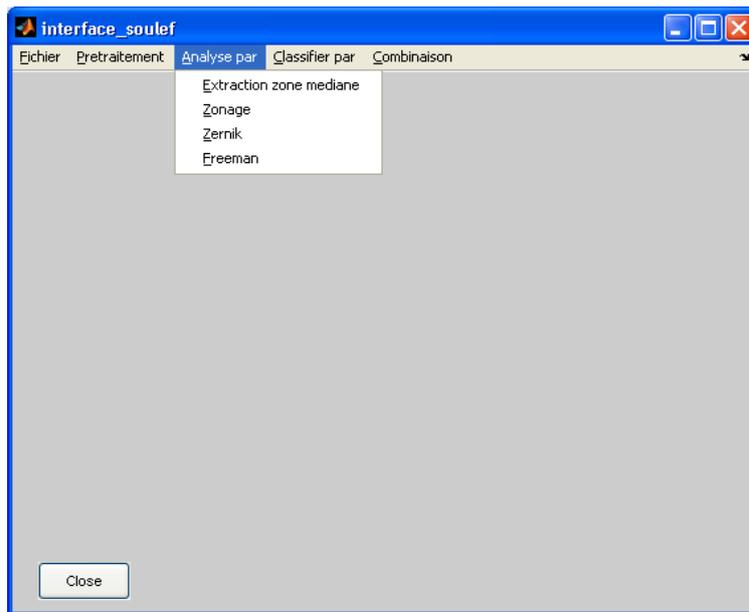


Figure B.4. Interface de menu d'analyse de mot.

- ❖ **Classifier par** : permet de choisir la méthode de classification parmi les quatre.



Figure B.5. Interface de menu de classification.

- ❖ **Combinaison** : nous avons réalisé la combinaison entre les quatre types de classifieurs de deux manières :
  1. Combinaison par vote de majorité : où tous les classifieurs ont la même priorité.
  2. Combinaison avec priorité : en attribuant 2 points aux PNN et KPPV, et un seul point pour FCM et K-MEANS.

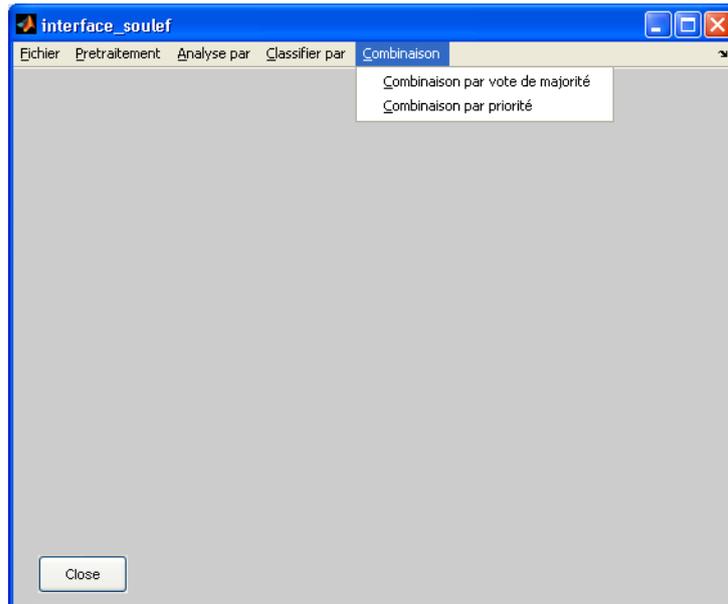


Figure B.6. Interface de menu de combinaison.

Après la description générale de différents modules de notre système, nous détaillons chacun de ces modules :

### B.3.2. Module de Prétraitement

1. pour obtenir l'image binaire on sélectionne le sous menu : Binarisation.
2. Pour éliminer le bruit, on sélectionne le sous menu : Lissage.

La figure suivante illustre le résultat de prétraitement.

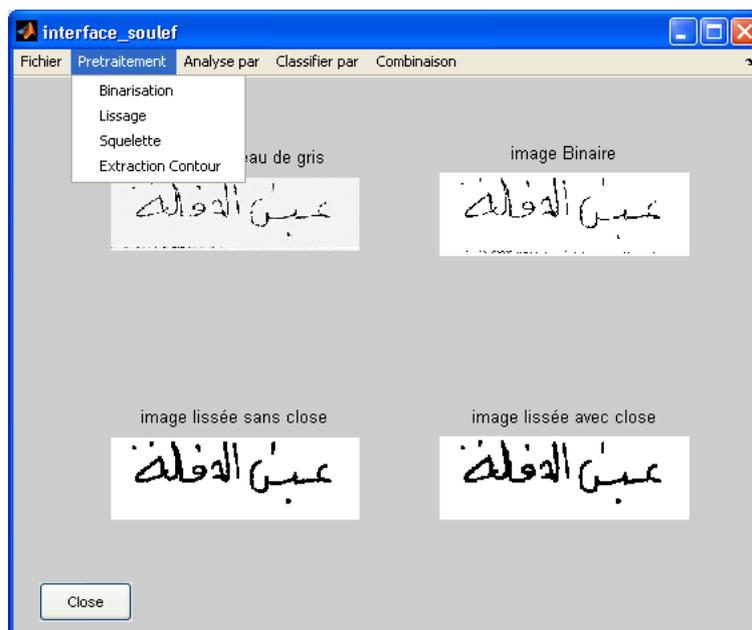


Figure B.7. Résultat de prétraitement.

- Après le nettoyage de l'image, nous détectons les PAWs constituant le mot. Et nous extrairons le squelette du mot.

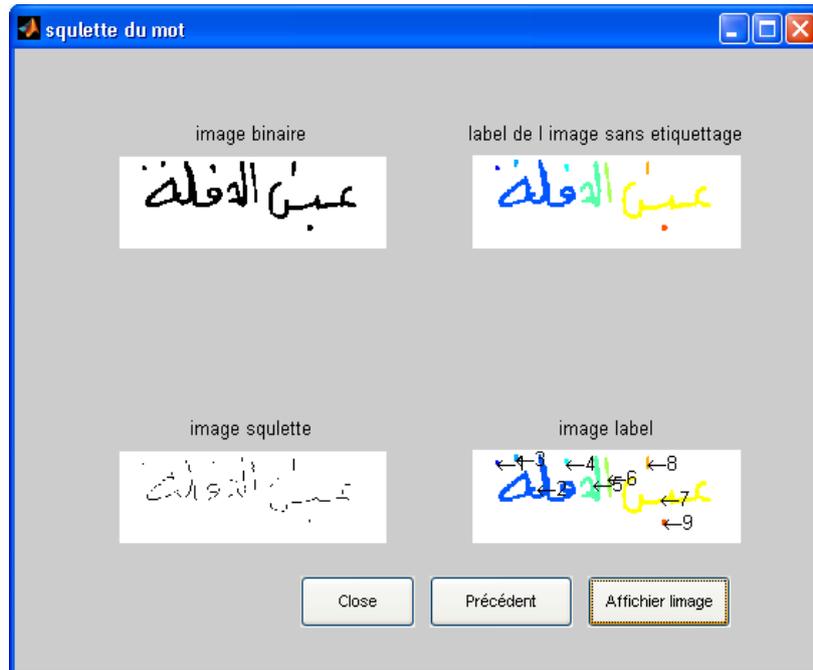


Figure B.8. Détection des composantes connexes et le squelette

- Pour obtenir l'image contour du mot, on sélectionne le sous menu : Extraction Contour.

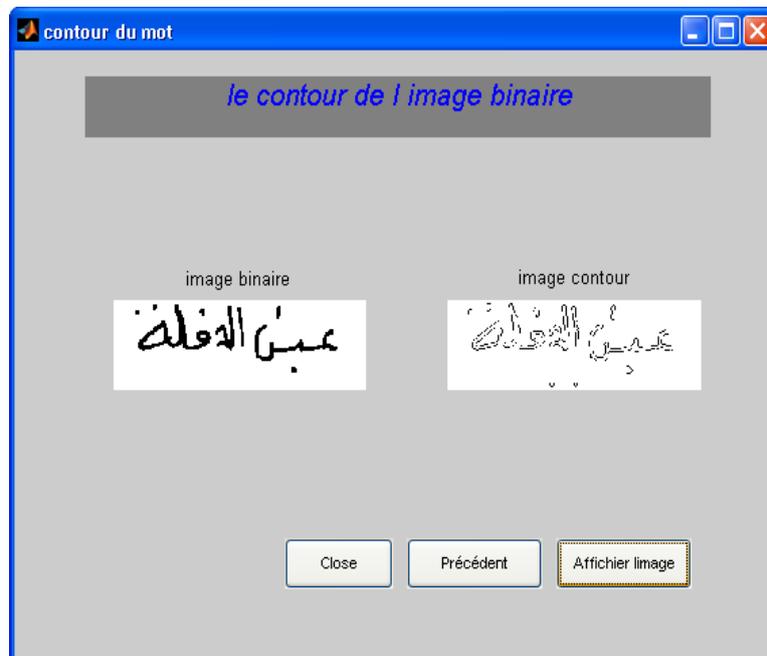


Figure B.9. Détection de contour.

### B.3.3. Module d'Analyse

Comme nous avons vu dans le dernier chapitre, nous avons suivi trois méthodes pour l'extraction des caractéristiques. Pour démarrer l'analyse, nous cliquons sur le bouton à droite «Démarrer l'analyse ». Et pour sauvegarder le résultat d'analyse, nous cliquons sur le bouton à gauche « save ».

1. Les caractéristiques structurelles du mot :

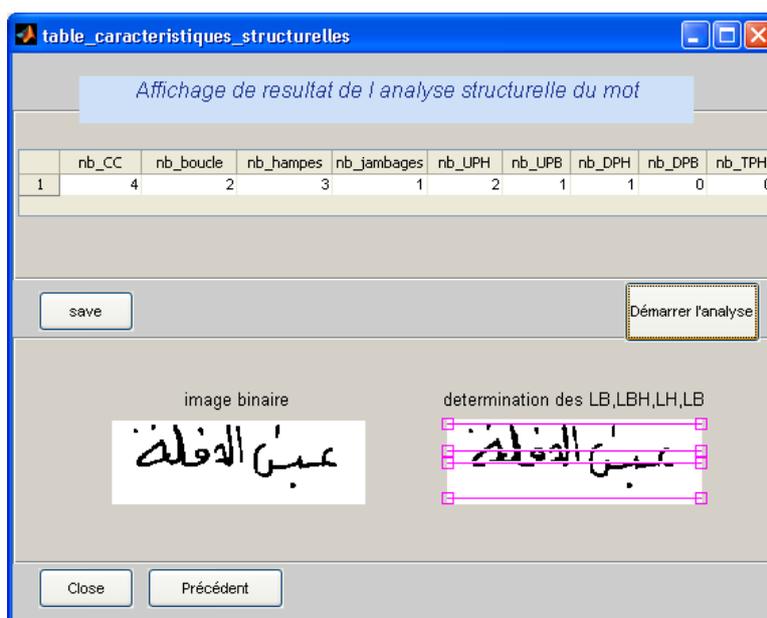


Figure B.10. Extraction de caractéristiques structurelles.

2. Les caractéristiques de zonage :

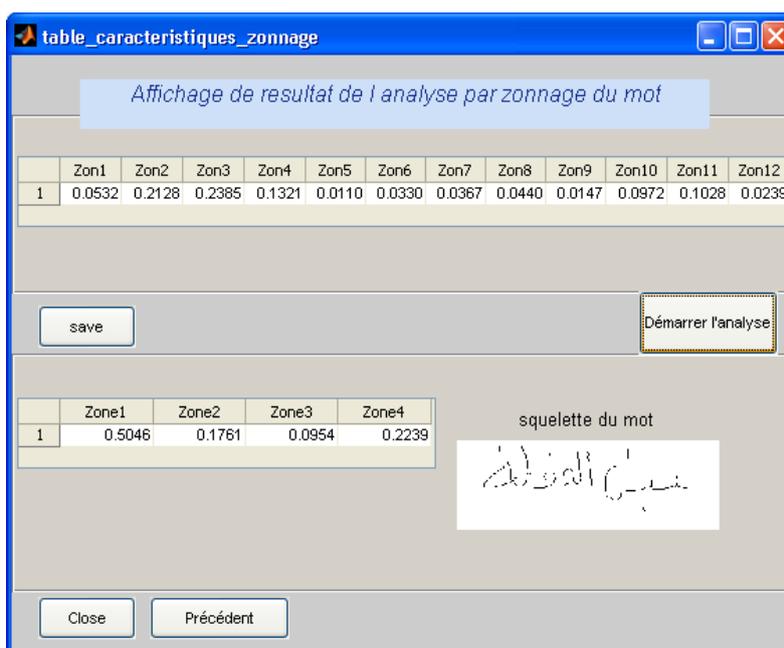


Figure B.11. Extraction de caractéristiques de zonage.

### 3. Extraction de code de Freeman :

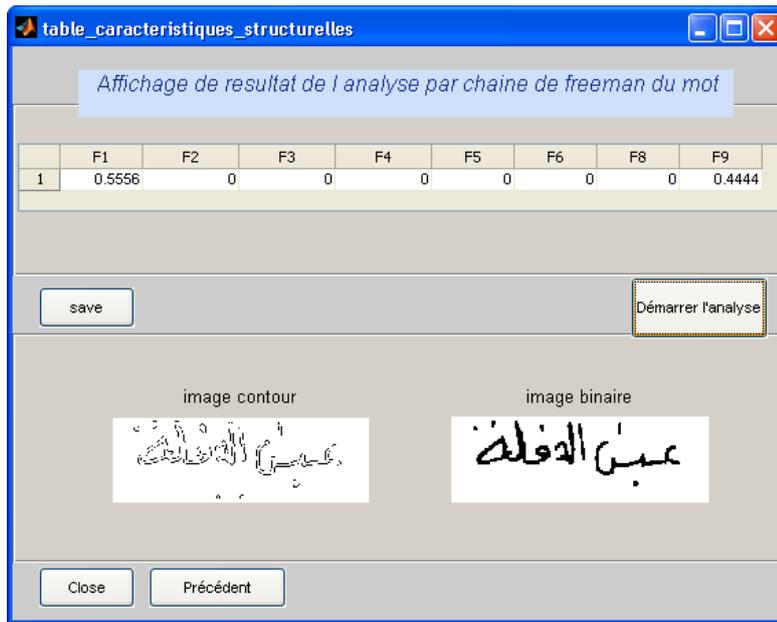


Figure B.12. Extraction de code de Freeman

#### B.3.4. Module de Classification

A partir de ce module nous pouvons classier le mot par les quatre classifieurs implémentés selon les trois ensembles de caractéristiques extraits dans le module précédent. Chacun va donner la classe où appartient le mot, comme le montre la figure suivante :



Figure B.13. Résultat de classification.

#### B.3.4. Module de Combinaison

Donne une décision finale concernant l'image d'entrée en combinant les différentes décisions données par chacun des classifieurs.

### B.4. Résultats

Pour évaluer la performance de la méthode proposée, des expériences ont été réalisées sur un sous ensemble de notre base de référence LRI relative au vocabulaire des noms de wilayas algériennes. Ainsi, les expériences d'évaluation ont été dérivées sur un ensemble de 12960 images pour l'apprentissage, et un ensemble de 1440 images pour le test. Le tableau B.1 présente les résultats du système proposé

Noms des wilayas	FCM(%)			KMEANS(%)			PNN(%)		
Tebessa	96.66	40	96.66	96.66	86.66	96.66	73.33	80	76.66
Jijel	70	23.33	63.33	63.33	56.66	63.33	73.33	50	76.66
Tissemsilt	40	63.33	33.33	33.33	30	26.66	53.33	93.33	53.33
Khenchela	70	63.33	70	60	30	63.33	76.66	60	80
Mila	53.33	56.66	33.33	56.66	53.33	53.33	50	66.66	43.33
MSILA	26.66	20	30	13.33	13.33	13.33	30	46.66	13.33
Mascara	63.33	33.33	73.33	73.33	63.33	73.33	73.33	40	73.33
Setif	53.33	20	73.33	70	63.33	70	66.66	53.33	66.66
Constantine	66.66	53.33	80	73.33	66.66	73.33	86.66	16.66	86.66
Saida	40	16.66	76.66	70	43.33	70	73.33	56.66	63.33
Bechar	60	73.33	73.33	66.66	56.66	63.33	83.33	93.33	76.66
Bejaia	60	13.33	66.66	63.33	46.66	63.33	46.66	53.33	60
Annaba	33.33	26.66	30	36.66	33.33	36.66	26.66	50	43.33
Batna	20	60	46.66	60	30	60	23.33	63.33	53.33
Biskra	43.33	40	53.33	50	26.66	50	70	33.33	50
Mostaganem	53.33	26.66	56.66	56.66	16.66	56.66	80	70	66.66
Skikda	33.33	33.33	33.33	30	2	30	63.33	36.66	33.33
Tlemcen	60	60	56.66	56.66	36.66	56.66	50	80	73.33
ELbayad	53.33	26.66	53.33	46.66	30	50	80	40	73.33
Ghuelma	36.66	50	70	70	53.33	70	66.66	33.33	73.33
Chlef	50	50	66.66	80	73.33	80	56.66	66.66	70
Djelfa	26.66	50	43.33	46.66	40	46.66	36.66	40	36.66
Tindouf	70	46.66	66.66	66.66	56.66	66.66	73.33	83.33	73.33
AinTemouchent	70	83.33	66.66	66.66	46.66	63.33	60	86.66	60
Tiaret	80	70	76.66	76.66	66.66	76.66	83.33	83.33	83.33
Tamanrasset	40	50	46.66	46.66	16.66	46.66	50	86.66	46.66
Tippaza	73.33	46.66	66.66	60	43.33	60	73.33	56.66	73.33
Ouargla	76.66	56.66	73.33	76.66	66.66	76.66	80	60	80
Illizi	76.66	43.33	66.66	73.33	46.66	73.33	86.66	50	86.66
Relizane	60	46.66	76.66	73.33	66.66	73.33	73.33	70	73.33
Medea	46.66	23.33	43.33	46.66	23.33	46.66	53.33	50	53.33
Blida	73.33	33.33	83.33	83.33	40	80	80	63.33	80
Naama	60	73.33	73.33	66.66	33.33	66.66	73.33	50	73.33
Ghardaia	83.33	70	73.33	76.66	40	76.66	86.66	66.66	83.33
Oran	73.33	50	73.33	73.33	60	73.33	70	53.33	80
SidiBELABES	73.33	83.33	76.66	73.33	60	73.33	70	80	73.33
Bouira	50	66.66	60	56.66	43.33	50	63.33	93.33	66.66
Alger	73.33	56.66	80	76.66	43.33	76.66	90	73.33	96.66
AinDefla	46.66	76.66	60	50	46.66	50	53.33	90	40
Eltaref	50	70	40	33.33	13.33	33.33	70	83.33	53.33
TiziOuzou	70	56.66	73.33	73.33	36.66	70	83.33	76.66	80
Boumerdes	83.33	70	80	76.66	40	76.66	83.33	83.33	80
Adrar	100	100	90	63.33	50	63.33	100	100	100
Eloued	46.66	63.33	56.66	56.66	33.33	56.66	70	83.33	76.66
Laghouat	83.33	70	90	90	73.33	86.66	96.66	70	90
BBAreridj	86.66	90	90	90	80	90	96.66	93.33	100
SoukAhras	96.66	80	93.33	96.66	73.33	93.33	100	93.33	80
OumElBouaghi	60	93.33	83.33	73.33	46.66	73.33	86.66	90	100

Tableau B.1. Résultats de reconnaissance