

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITE DES FRERES MENTOURI, CONSTANTINE1
FACULTE DES SCIENCES DE LA TECHNOLOGIE
DEPARTEMENT D'ELECTRONIQUE

N° d'ordre :230/DS/2018

Série :23/Elec/2018

Thèse

présentée par

Mr. Azzeddine MENASRIA

pour l'obtention du

Doctorat en Sciences

Option : Traitement du Signal

Thème

**Développement des techniques de segmentation et de
reconnaissance d'écriture arabe manuscrite**

Soutenue le : 27/11/2018

Devant le jury :

Présidente :	N. MANSOURI	Professeur	Université de Constantine
Rapporteur :	A. BENNIA	Professeur	Université de Constantine
Examineur :	H. BOUROUBA	MCA	Université de Guelma
Examineur :	K. MESSAOUDI	MCA	Université de Souk Ahras
Invité :	M. SEDRAOUI	Professeur	Université de Guelma

Année 2018

REMERCIEMENTS

Au terme de ce travail, je voudrai exprimer ma profonde gratitude envers dieu tout puissant qui, grâce à son aide, j'ai pu finir mon travail.

Tout d'abord je tiens à exprimer ma reconnaissance et mon profond respect à mon directeur de thèse Professeur **Abdelhak BENNIA** de m'avoir propose ce sujet et diriger constamment mon travail. J'apprécie la confiance qu'il m'a témoignée.

Je souhaite adresser mes sincères remerciements à M^{elle} **N. MANSOURI** professeur à l'universite de Constantine, d'avoir accepté de présider le jury.

Je tiens à remercier également les membres du jury : Mr. **H. BOUROUBA** Maître de conférences à l'université de Guelma et Mr. **K. MESSAOUDI** Maître de conférences à l'université de Souk Ahras qui ont voulu accepter de valoriser cette thèse.

Je tiens à remercier également Professeurs **M. SEDRAOUI** et **A. YOUNSSI** pour son aide.

ملخص:

نقترح في هذا العمل إستراتيجية جديدة للتعرف على الكتابة الخطية المحررة باللغة العربية ونخص في هذا المجال التعرف على الكتابة الحرفية للصوصك البنكية ، حيث قمنا باقتراح ثلاث مساهمات رئيسية لتطوير نموذج التعرف . تعتمد هذه المساهمات على الأجزاء الهيكلية للكلمة (PAWS) ميزة تختص بها الكتابة العربية وتظهر بشكل متباين يسهل استخراجها دون تجزئة الكلمة هذه المساهمات يمكنها تعزيز وتحسين مراحل التعرف على الكلمة ، وتكمن المساهمة الأولى في تحسين مرحلة استخراج السيمات المناسبة وكيفية اختيارها حيث في الأول يتم استخراج السيمات الهيكلية من أجزاء الكلمة وهذا الاختيار أحسن من إستخراجها من الكلمة ذاتها ، إلا أنه في بعض الأحيان نجد كلمتين مختلفتين لهما نفس الخصائص الهيكلية لهذا إقترحنا إضافة السيمات الإحصائية من الكلمة نفسها وتضاف إليها الخصائص الهيكلية لتكون في الأخير جملة من السيمات المشتركة تسهل بشكل كبير في مرحلة التعرف ، في هذه المرحلة ادرجنا المساهمة الثانية حيث قمنا بتقسيم جميع الكلمات المكونة لمعجم الصوك الى أربعة فئات وذلك حسب عدد الأجزاء الهيكلية المكونة للكلمة من واحد الى أربعة. بعد استخراج السيمات لكل مجموعة على حدى تأتي مرحلة التصنيف باستخدام أجهزة المتجهات الاعتمادية (SVM) ، إلا أن هذه المرحلة تتأثر بمشكلة التصاق الأجزاء الهيكلية للكلمة مما يؤدي إلى خطأ في التصنيف لهذا اقترحنا المرحلة الثالثة ما قبل التصنيف لإعطاء العدد الصحيح مع جميع احتمالات الإلتصاق ليتم في الأخير تصنيف نهائي في الكلمة بإجتياز اكبر احتمال لظهور الكلمة الصحيحة وقد تم تقييم الطريقة المقترحة باستخدام قاعدة بيانات AHDB حيث حققت نتائج متميزة مقارنة مع أعمال أخرى.

الكلمات المفتاحية:

التعرف على الكتابة الخطية ، التصنيف ، إستخراج السيمات ، تقطع الكلمات

Résumé :

Ce travail propose une nouvelle stratégie de reconnaissance assurant le développement d'un nouveau système de reconnaissance très efficace pour les écritures manuscrites arabes en particuliers les montants littéraux utilisés dans les chèques bancaires.

La construction du modèle proposé est fondée sur la proposition de trois contributions principales avec lesquelles le taux de reconnaissance est amélioré de façon très remarquable. Ces trois contributions sont proposées afin d'améliorer les résultats des deux phases qui sont : la phase d'extraction des caractéristiques et la phase de classification. Ces améliorations sont basées sur la segmentation naturelle d'un seul mot candidat qu'à partir duquel un ensemble des pseudo-mots (PAWs) est généré. La première contribution est proposée dans la phase d'extraction des primitives dans laquelle les caractéristiques structurelles d'un mot candidat sont extraites à partir de ses PAWs. Sachant que certains mots peuvent produire les mêmes caractéristiques précédentes. A cet effet, d'autres caractéristiques statistiques sont extraites à partir du mot puis les combinées avec les caractéristique structurelles fournies par ses PAWs. De plus, la deuxième contribution est proposée dans la phase de classification dans laquelle le multi-classifier SVMs est utilisé pour la classification au lieu de mono-classifier SVM. Par conséquent, le problème multi-classe est divisé, selon le nombre des PAWs, en quatre sous-problèmes où chaque sous-problème est résolu indépendamment de l'autre en utilisant quatre ensembles séparés de classifier. La troisième contribution est proposée dans le but de corriger les PAWs connectés entre eux qui sont apparus à cause de la mauvaise écriture. Ce problème est donc résolu par l'introduction d'une étape additionnelle de pré-classification dans laquelle le nombre erroné des PAWs est corrigé convenablement. Finalement, le système proposé est validée sur la base de données AHDB (*Arabic Handwriting Database*) et la comparaison des résultats obtenus avec ceux fournis par d'autres systèmes conventionnels de reconnaissance confirment ses efficacités en termes de précision de classification ainsi que le taux de reconnaissance.

Mots clés : Caractéristiques structurelle et statistique, reconnaissance d'écriture arabe, classification, prétraitement, pré-classification.

Abstract :

This work proposes a new recognition strategy involving an efficient recognition system for Arabic manuscript. Especially, for the handwritten Arabic literal amounts that are used in the check bank. Three main contributions are proposed in the development of our recognition model. These contributions are based on the natural segmentation of the candidate word in which a set of the Part of Arabic Words PAWs is generated. Accordingly, these proposed contributions have the ability to enhance both recognitions steps, which are: the feature extraction and the classification steps. The first contribution is proposed in the feature extraction step where the structural features are extracted from the PAWs of the candidate word. Notice that, some candidate words may be provided the same structural features. To avoid this problem, the statistical features are extracted from the candidate word and then combined with the above-mentioned structural ones. The second contribution is proposed in the classification step, in which the multi-classifiers based SVMs is used instead the single classifier SMV. Therefore, the multi-classes problem may be divided according the number of the PAWs where each one is independently solved using four separate set of the classifier. The main goal of the third contribution is to solve the touching problem of certain PAWs, which is appeared when a candidate manuscript is ill-written. This problem is solved by including an additional pre-classification step where the erroneous number of the PAWs is well corrected. Finally, the proposed recognition system is validated on the AHDB data base and the obtained results are compared by those given by the conventional recognition systems. Accordingly, the proposed recognition system provides better results in terms of classification accuracy and recognition rate.

Key-words: Structural and statistical features, recognition of the handwritten Arabic word, classification, preprocessing, pre-classification.

TABLE DES MATIÈRES

LISTE DES FIGURES	i
LISTE DES TABLEAUX	ii
INTRODUCTION GENERALE.....	iii

CHAPITRE I: LA RECONNAISSANCE DE L'ECRITURE MANUSCRITE

I.1 INTRODUCTION	1
I.2 DIFFERENTS ASPECTS DE RECONNAISSANCE DE L'ECRITURE.....	1
I.2.1 TYPE D'ACQUISITION.....	2
I.2.1.1 La reconnaissance en-ligne (On line).....	2
I.2.1.2 La reconnaissance hors-ligne (Off line)	2
I.2.2 STYLE DE L'ECRITURE.....	3
I.2.2.1 Reconnaissance de caractères ou l'analyse de documents:.....	4
I.3 LES DIFFICULTES DE RECONNAISSANCE DE L'ECRITURE MANUSCRITE	5
I.3.1 NOMBRE DE SCRIPTEURS	5
I.3.2 TAILLE DU VOCABULAIRE TRAITE	6
I.3.3 CONTRAINTES MORPHOLOGIQUES	6
I.3.4 LA DISPOSITION SPATIALE DU MANUSCRIT	6
I.3.5 STRATEGIES DE RECONNAISSANCE DE L'ECRITURE.....	7
I.3.5.1 L'approche globale:.....	8
I.3.5.2 L'approche analytique:.....	8
I.4 PRINCIPALES PHASES D'UN SYSTEME DE RECONNAISSANCE D'ECRITURE HORS LIGNE.....	8
I.4.1 ACQUISITION.....	9
I.4.2 PRETRAITEMENT ET NORMALISATION	10
I.4.2.1 Filtrage	10
I.4.2.2 Binarisation	10
I.4.2.3 Lissage.....	12
I.4.2.4 Normalisation.....	13
I.4.2.5 Squelettisation.....	13
I.4.2.6 Détection des lignes de référence.....	14
I.4.3 SEGMENTATION.....	14
I.4.3.1 Segmentation explicite.....	15
I.4.3.2 Segmentation implicite	16
I.4.4 EXTRACTION DE PRIMITIVES.....	16
I.4.4.1 Caractéristiques statistiques	17
I.4.4.2 Caractéristiques structurelles.....	18
I.4.4.3 Les transformations globales et les moments.....	18
I.4.5 CLASSIFICATION	21
I.4.5.1 Les approches statistiques.....	22
I.4.5.2 Les approches structurelles	23
I.4.5.3 Les approches stochastiques	26

I.4.6	LA DECISION.....	27
I.4.6.1	<i>Evaluation des performances d'un classifieur</i>	28
I.5	DOMAINES D'APPLICATIONS DE LA RECONNAISSANCE DE L'ECRITURE	29
I.6	CONCLUSION	29

CHAPITRE II:LA RECONNAISSANCE DE L'ECRITURE ARABE MANUSCRITE

II.1	INTRODUCTION	30
II.2	PRESENTATION ET PROBLEMATIQUE DE L'ECRITURE ARABE	30
II.3	PRINCIPALES BASES DE DONNEES EXISTANTES	35
II.3.1	CENPARMI 1 DATABASE	35
II.3.2	IFN/ENIT DATABASE	36
II.3.3	AHDB DATABASE	36
II.3.4	CEDARABIC DATABASE.....	36
II.3.5	IFN/FARSI DATABASE	37
II.3.6	KHATT DATABASE.....	37
II.3.7	AHTID/MW DATABASE 2012	37
II.3.8	QUWI DATABASE.....	38
II.4	SURVEYS EXISTANTS	38
II.4.1	B. AL-BADR ET S. A. MAHMOUD, 1995	38
II.4.2	A. AMIN, 1998.....	39
II.4.3	M. S. KHORSHEED, 2002.....	39
II.4.4	N. ESSOUKRI BEN AMARA ET F. BOUSLAMA, 2003	39
II.4.5	L. M. LORIGO ET V. GOVINDARAJU, 2006	40
II.4.6	M. T. PARVEZ ET S.A. MAHMOUD, 2013.....	40
II.4.7	A. LAWGALI , 2015	41
II.5	PROCESSUS DE RECONNAISSANCE D'ECRITURES ARABES MANUSCRITS	41
II.5.1	LE PRETRAITEMENT.....	42
II.5.1.1	<i>Squelettisation</i>	42
II.5.1.2	<i>Correction de l'inclinaison des lignes</i>	43
II.5.1.3	<i>Estimation de la ligne de base</i>	43
II.5.2	ESTIMATION DE LA BANDE DE BASE	48
II.5.3	SEGMENTATION DES MOTS EN CARACTERES	48
II.5.4	EXTRACTION DE PRIMITIVES.....	48
II.6	DIFFERENTES APPROCHES ET SYSTEMES EXISTANTS	50
II.6.1	APPROCHE HOLISTIQUE PAR HMM	51
II.6.2	APPROCHE HOLISTIQUE PAR RNN	51
II.6.3	APPROCHE HOLISTIQUE PAR (SVMs)	51
II.7	CONCLUSION	51

CHAPITRE III: SEGMENTATION DE L'ECRITURE ARABE MANUSCRITE

III.1	INTRODUCTION	59
III.2	SEGMENTATION DE L'ECRITURE CURSIVE	59

III.3	DIFFICULTES DANS LA SEGMENTATION DES CARACTERES	60
III.4	LES METHODES DE SEGMENTATION DES CARACTERES	61
III.4.1	LA SEGMENTATION PAR TECHNIQUES DE DISSECTION	61
III .4.1.1	<i>L'analyse des composants connexes</i>	<i>62</i>
III .4.1.2	<i>L'espace blanc et l'estimation du pitch</i>	<i>62</i>
III .4.1.3	<i>Segmentation à partir Les profils de projection verticales</i>	<i>63</i>
III .4.1.4	<i>Détection des bornes limites</i>	<i>64</i>
III.4.2	LA SEGMENTATION PAR RECONNAISSANCE	64
III .4.2.1	<i>Les approches de fenêtrage</i>	<i>65</i>
III .4.2.2	<i>Les approches basées sur les attributs caractéristiques.....</i>	<i>65</i>
III.4.3	LA SEGMENTATION HYBRIDE.....	66
III.4.4	SEGMENTATION HOLISTIQUE (IMPLICITE)	68
III.5	LA TECHNIQUE DETECTION DES POINTS DE COUPURES	68
III.5.1	ANALYSE DU CONTOUR	69
III.5.2	SEGMENTATION A PARTIR DU SQUELETTE.....	69
III.5.3	SEGMENTATION A PARTIR DU CONTOUR	69
III .5.3.1	<i>Minima des profils horizontaux et verticaux.....</i>	<i>69</i>
III .5.3.2	<i>L'algorithme Hit-and-deflect</i>	<i>70</i>
III .5.3.3	<i>L'algorithme Drop-and-Falling</i>	<i>70</i>
III .5.3.4	<i>L'algorithme des points de coupure critique</i>	<i>70</i>
III.5.4	SEGMENTATION A BASE DE FENETRES GLISSANTES	71
III.6	EXEMPLES DE METHODES DE LA SEGMENTATION DE CARACTERES.....	71
III.6.1	LA SEGMENTATION PAR HMMs	71
III.6.2	LA SEGMENTATION BASEE SUR LA CORRESPONDANCE DE MODELES	72
III.7	CONCLUSION	73
IV.1	INTRODUCTION	67
IV.2	PRINCIPALES ETAPES DU SYSTEME DE RECONNAISSANCE DE L'ECRITURE	70
IV.2.1	PRETRAITEMENT.....	70
IV.2.2	EXTRACTION DES CARACTERISTIQUES	70
IV.2.2.1	<i>Extraction des caractéristiques structurelles.....</i>	<i>72</i>
IV.2.2.2	<i>Extraction de caractéristiques statistiques.....</i>	<i>76</i>
IV.2.3	ÉTAPE DE CLASSIFICATION.....	79
IV.3	LA METHODE PROPOSEE:	80
IV.4	ÉTAPE DE RECONNAISSANCE BASEE SUR LA RESOLUTION DU PROBLEME DE PAWS CONNECTEES.....	83
IV.5	PHASE DE TESTS	86
IV.6	CONCLUSION	88
	CONCLUSION GENERALE.....	90
	BIBLIOGRAPHIE.....	92

LISTE DES FIGURES

Fig.I.1	Écriture hors ligne (ensemble de pixels) et écriture en ligne (ensemble de coordonnées).....	03
Fig.I.2	Styles d'écritures (a)Écriture Arabe imprimé (b) Écriture Arabe manuscrite (c)Écriture latine manuscrite.....	05
Fig.I.3	Structure Générale d'un système de reconnaissance d'écriture	09
Fig.I.4	Choix de seuil sur l'histogramme de l'image à niveaux	11
Fig.I.5	Les différents masques pour les opérations de nettoyage un mot.....	13
Fig.I.6	Exemple de lignes de base d'écriture (a) cas de d'écriture latine, (b) cas de d'écriture arabe	14
Fig.I.7	Différentes segmentation pour le mot quatorze.....	16
Fig.I.8	SVM : Trouver une frontière de séparation linéaire dans un espace de grande dimension revient à trouver une frontière non-linéaire dans l'espace de départ.....	25
Fig.II.1	Exemple de pseudo-mots constituant les mots arabes.....	31
Fig.II.2	Cursivité de l'écriture arabe.....	31
Fig.II.3	Difficultés liées aux points diacritiques. (a) différents styles d'écriture des points diacritiques ; (b) problème d'association des diacritiques à leur lettre originale.....	34
Fig.II.4	Les problèmes liés à d'extraction des pseudo-mots. (a) un mot arabe correctement écrit (3PAW) ; (b) liaison indésirable entre caractères et sous segmentation en PAWs (2 seul PAW) ; (c) coupure indésirable et sur-segmentation en PAWs (4 PAWs).....	34
Fig.II.5	Ligatures verticales et inversion de l'ordre du tracé	35
Fig.II.6	Estimation de la ligne de base par la projection horizontale.....	43
Fig.II.7	lignes de base	44
Fig.II.8	Extraction de la ligne de base par la combinaison de la méthode de projection et la transformée de Hough.....	45
Fig.II.9	Problèmes liés à la méthode de projection horizontale : pics parasites dû à la succession des descendants avec des portions horizontales suffisamment longues.....	45
Fig.III.1	Exemple de segmentation en graphèmes.....	53
Fig III.2	Classification des méthodes de segmentation de caractères	54
Fig.III.3	Profils de projection horizontal et vertical	57
Fig.III.4	Segmentation à partir du squelette	62
Fig.IV.1	Déférentes étape de prétraitements	78
Fig.IV.2	Aperçu sur la base de données AHDB	79
Fig.IV.3	Nombre de PAWs extraits du mot « اربعون »	81

Fig.IV.4	Extraction des boucles; (a) Méthode basée sur le contours intérieurs. (b) Méthode basée sur l'algorithme CCL	81
Fig.IV.5	Positionnement des caractères par rapport à les lignes	82
Fig.IV.6	Nombre de points diacritiques et leurs positions du mot «ثمانية - Huit»	83
Fig.IV.7	Extraction des points diacritiques	84
Fig.IV.8	Histogrammes locaux de chaine de code du mot «سبع - sept »	85
Fig.IV.9	Topologies des méthodes de zonage: (a) Méthode de zonage rectangulaire régulier (b) Méthode de zonage chevauchée basée sur une grille régulière 3x3. (c) Méthode de zonage triangulaire	86
Fig.IV.10	a) Les huit projections des profils b) Histogrammes de la densité des profils	87
Fig.IV.11	Exemple de deux mots inappropriés «عشرون-Vingt» composés de deux PAWs connectés	91
Fig.IV.12	Système de reconnaissance proposé	93
Fig.IV.13	Taux de reconnaissance pour les 1er, 2eme, 3eme et 4eme sous-groupes	97

LISTE DES TABLEAUX

Tab II.1	Les caractères arabes et leurs différentes formes, (b) les caractères additionnels « ة » et « ل », (c) et (d) Hamza et Madda et les positions qu'elles occupent en association avec Alif, Waw et Ya	29
Tab II.2	Différents hambages et jambages dans les lettres arabes	32
Tab II.3	Lettres arabes ayant des points diacritiques.....	32
Tab II.4	Primitives utilisées dans les systèmes de reconnaissance de mots arabes	52
Tab IV.1	Vocabulaire des montants littéraux de cheque arabe (AHDB database)	80
Tab IV.2	Différentes significations utilisées pour coder les primitives existantes.....	90
Tab IV.3	Différents sous-groupes obtenus en fonction du nombre de PAWs isolés	90
Tab IV.4	Différents vecteurs de caractéristiques globales utilisés pour alimenter les classificateurs proposés	
Tab IV.5	PAWs corrects et incorrects possibles dans chaque sous-groupe	91
Tab IV.6	Résultats complets de la classification.....	92
Tab IV.7	Comparaison des résultats de reconnaissance sur l'ensemble de données DHDB	95
Tab IV.8	La moyenne du temps d'exécution pour l'étape de reconnaissance	95

INTRODUCTION GENERALE

L'écriture est, a été et restera indubitablement un moyen essentiel de communication et d'échange d'informations entre les personnes. A cet effet, de nombreux logiciels ont été développés pour faciliter cet échange, en particulier, ceux visant à automatiser des tâches assez difficiles à réaliser par l'être humain. Toutefois, en dépit des avancées remarquables réalisées dans les domaines de l'informatique et de l'automatique, notamment le développement des machines puissantes et des logiciels très performants, le problème de la reconnaissance automatique de l'écriture reste toujours un défi majeur pour de nombreux systèmes de reconnaissance développés au cours de ces dernières décennies. A cet égard, la reconnaissance automatique de l'écriture a figurée comme l'un des objectifs principaux de recherche de plusieurs équipes de recherche, de par le monde, durant ces dernières décennies et cela dans le but de transformer les images de textes en leurs transcriptions numériques codées soit en ASCII soit en UNICODE. Cette transformation assurera alors une lecture facile à comprendre par les machines d'une part, ainsi qu'une manipulation aisée par les logiciels de traitement de texte, d'autre part.

Dans la plupart des applications quotidiennes, réelles, le développement d'un système de reconnaissance efficace peut, notamment, dépendre de la nature de l'écriture à acquérir. En général, il existe deux types de reconnaissance: la reconnaissance en ligne et la reconnaissance hors ligne. Le premier type de reconnaissance concerne les écritures fournies par les stylos électroniques, les agendas numériques et les *Personal Digital Assistant* (PDA). Parallèlement à cela, la reconnaissance hors ligne a trait aux opérations de tri automatique des courriers, au traitement des formulaires et à la lecture automatique des chèques bancaires. Dans ce travail, nous verrons qu'il reste encore de nombreux problèmes à résoudre avant d'arriver à développer un système de reconnaissance efficace capable d'automatiser la reconnaissance de l'écriture arabe dans les chèques bancaires. Cet objectif représente l'objectif principal de ce travail.

Plusieurs systèmes de reconnaissance automatique ont été développés durant ces dernières années. Ces systèmes sont assurés par des outils commerciaux présents dans la plupart des établissements. A ce titre, les services postaux utilisent, depuis assez longtemps et de manière systématique, des machines spécifiques pour la lecture des adresses des colis postaux et pour le tri automatique des courriers. Dans les pays développés, les banques font largement appel aux systèmes de reconnaissance pour le traitement des chèques dans les distributeurs automatiques. A ce stade, il faut noter que contrairement à l'écriture latine ou encore l'écriture chinoise, qui ont enregistré des avancées très remarquables dans ce domaine, le

problème de la reconnaissance automatique de l'écriture arabe reste encore d'actualité. Ce retard est la conséquence à de nombreux obstacles inhérents à la langue arabe et nécessitant plus de règles de reconnaissance spéciales qui sont spécifiques à cette langue. Parmi ces obstacles, on peut citer: l'indisponibilité de bases de données réelles, grande diversité et variabilité de l'écriture arabe (qui peut elle-même dépendre de la forme morphologique du scripte arabe) ainsi que l'existence d'une grande variation des styles d'écriture effectués par le scripteur, ... etc.

Dans ce travail, nous nous intéresserons à la reconnaissance hors ligne de l'écriture arabe manuscrite, en particulier, celle qui est caractérisée par un vocabulaire limité utilisé par exemple pour la rédaction des montants littéraux des chèques arabes. A cet effet, l'objectif principal de ce travail consiste à améliorer le taux de reconnaissance en termes de précision et robustesse face aux changements d'écriture effectuée par le scripteur lui-même et en présence des problèmes causés par la nature de l'écriture du scripteur. L'amélioration souhaitée, susceptible d'assurer, autant que possible, une reconnaissance précise, est souvent effectuée lors de l'une des trois étapes de base suivantes: étape de prétraitement, étape d'extraction de caractéristiques structurelle et géométrique et finalement étape de classification. Dans notre travail, nous nous intéresserons plus précisément aux deux dernières étapes dans lesquelles nous proposerons un nouveau modèle d'extraction des caractéristiques structurelle et statistique qui est basé sur l'analyse d'une partie du pseudo-mot (ou PAW, de l'anglais *Part of Arabic Words*). Une deuxième amélioration sera également proposée au cours de l'étape de classification. Cette dernière consistera à utiliser un multi-classificateur au lieu d'un seul. Ceci sera effectué par une subdivision du problème multi-classe en quatre sous-problèmes. Chaque sous-problème est indépendamment résolu en utilisant quatre ensembles séparés de classificateur. Une troisième proposition visera à résoudre le problème des PAWs attachés entre eux. Ce problème peut être le résultat d'une écriture erronée effectuée par le scripteur lui-même. Cette problématique sera résolue par la proposition d'une étape auxiliaire de pré-classification. La validation de notre système de reconnaissance proposé a été confirmée via une application expérimentale sur la base de données (AHDB). Les résultats obtenus sont meilleurs que ceux fournis par les systèmes de reconnaissance conventionnels existants dans la littérature.

Le plan de cette thèse s'articule autour des quatre chapitres suivants :

- **Le premier chapitre** portera sur l'état de l'art de la reconnaissance de l'écriture manuscrite. Il comportera un apport principal de quelques travaux de recherche existants dans la littérature, en l'occurrence, la reconnaissance automatique de l'écriture. Il présentera

également les différents problèmes rencontrés lors du développement du système de reconnaissance automatique de l'écriture manuscrite ainsi que sur les différentes étapes qui constitue un système de reconnaissance telles que: l'acquisition, le prétraitement, l'extraction des primitives et la classification. Ce chapitre sera finalement achevé par la présentation de quelques domaines d'applications existant dans la vie quotidienne.

- **Le deuxième chapitre** portera sur les détails de la reconnaissance de l'écriture arabe manuscrite ainsi que sur ses différentes propriétés morphologiques. Il présentera également un apport sur les principales bases de données existantes dans la littérature, en l'occurrence, CENPARMI1, IFN/ENIT, CEDARBIC, IFN/FARSI, KHATT, LAHTID/MW, QUWI et la base de données AHDB qui a été utilisée dans cette étude pour la validation de notre système proposé. Les étapes requises pour développer un processus de reconnaissance de l'écriture arabe manuscrite y ont été également détaillées. Finalement, ce chapitre regroupera quelques travaux de recherche utilisant l'approche globale basée sur le type de classificateur employé, à savoir, le modèle de Markov caché (HMM), les réseaux de neurones (RNN), *Support Vector Machine* (SVM), ... etc.

- **Le troisième chapitre** a pour but de détailler les différentes méthodes de segmentation d'écriture arabe ainsi que les difficultés engendrées lors de l'étape de segmentation d'une écriture arabe manuscrite.

- **Le quatrième chapitre** développera la contribution apportée par ce travail. A ce propos, nous présenterons les différentes étapes permettant de développer notre nouveau système de reconnaissance pour les montants littéraux arabes. Ces derniers permettront de diviser la base AHDB en quatre sous-groupes principaux en ce, eu égard au nombre des PAWs constituant les mots utilisés dans le chèque bancaire. Ensuite, les caractéristiques structurelles (extraites à partir d'un PAW isolé) ainsi que les caractéristiques statistiques (extraites à partir d'un mot global correspondant) seront combinées. Nous ferons ressortir le vecteur caractéristique concaténé qui alimente le classificateur SVMs. En outre, une étape de pré-classification sera ajoutée afin de corriger les PAWs erronés engendrés suite à une écriture manuscrite mal effectuée. Finalement, la classification de chaque mot sera assurée simultanément avec et sans correction des PAWs et la décision finale sera donnée en fonction du meilleur score.

Finalement, ce travail sera terminé par une conclusion générale et une proposition de quelques perspectives pour les futurs candidats intéressés par cet axe de recherche.

Chapitre **1**

*La Reconnaissance de
L'écriture Manuscrite*

I.1 Introduction

L'une des innovations les plus extraordinaires de l'Homme à travers tous les âges, sinon la plus prodigieuse, a été l'invention de l'écriture. Sans l'écriture, beaucoup d'œuvres des plus grands maîtres à penser, physiciens, philosophes, historiens, aventuriers, ... auraient été significativement altérées lors de leurs transmissions orales, voire complètement perdues à jamais. Depuis l'invention de l'écriture par les Sumériens, l'Homme a su communiquer aux générations postérieures ses désirs, ses craintes, ses découvertes et ses pensées. A cet égard, l'écriture a été, et restera, indubitablement l'un des grands fondements de toutes les civilisations de par le monde et le mode, par excellence, de conservation et de transmission des connaissances. Ainsi, l'écriture a permis le cumul des connaissances et a permis à l'Homme de voler dans l'espace, à faire reculer les frontières de l'univers, à le faire plonger au cœur de la matière, à faire parler et danser des machines, ...

Il n'y a qu'à regarder autour de nous pour mesurer l'importance des écrits dans notre vie quotidienne: les affiches publicitaires, les enseignes des magasins, les pancartes de la route, les guides utilisateurs des machines, les magazines, les livres, les obélisques égyptiennes, ... etc.

Permettre à une machine d'appréhender l'ensemble de ces informations environnantes, les analyser et les restituer à l'utilisateur ferait gagner énormément en termes de fiabilité, rapidité et prise de décision. Dans ce contexte, la reconnaissance de l'écriture a déjà aujourd'hui investi de nombreux secteurs d'activité qui s'étendent du tri postal jusqu'à la saisie automatique de formulaires administratifs en passant par la vérification et la lecture des chèques bancaires.

Après avoir brossé un tableau très succinct sur le thème de la reconnaissance de l'écriture manuscrite, nous traiterons, dans ce chapitre, des difficultés rencontrées dans cette discipline, des diverses approches de reconnaissance ainsi que des différentes étapes du processus de reconnaissance.

I.2 Différents aspects de reconnaissance de l'écriture

La reconnaissance automatique de l'écriture est aujourd'hui un domaine très vaste qui s'intéresse principalement à l'élaboration de systèmes à même d'assurer la reconnaissance de textes manuscrits ou imprimés. De nombreux éléments participent dans la mise au point de ces systèmes, en l'occurrence, le nombre de scripteurs, le type ou le style de l'écriture, la taille et le type du vocabulaire. Ces éléments, qui sont à la base de toute conception d'un

système de reconnaissance, sont incontournables pour la mise en pratique de ces systèmes de reconnaissance. Ils se distinguent entre eux selon de nombreux critères. Citons à titre d'exemple, la stratégie adoptée pour la localisation des unités de base de l'écriture des mots ainsi que la nature des informations disponibles.

I.2.1 Type d'acquisition

Quand il s'agit de reconnaître une écriture manuscrite, le type d'acquisition de l'écriture est un paramètre fondamental qui permet la distinction des systèmes de reconnaissance entre eux. En effet, la reconnaissance d'une écriture en ligne ou bien en hors-ligne (**Figure I.1**) constitue un processus différent pour les systèmes de reconnaissance et concourt donc à une différenciation essentielle entre eux.

I.2.1.1 La reconnaissance en-ligne (On line)

Appelée également reconnaissance dynamique, cette méthode reconnaît l'écriture en temps réel. En ce sens qu'au fur et à mesure que les mots sont écrits à la main, les symboles sont immédiatement reconnus. Ce mode de reconnaissance, généralement réservé aux textes manuscrits, présente plusieurs avantages dans la mesure où l'approche de la reconnaissance est du type signal. En effet, elle est effectuée sur des données selon une seule dimension. L'écriture se présente sous la forme d'une séquence de points dont les coordonnées sont fonction du temps. Ainsi, l'ordre de l'écriture, l'information temporelle, les levers du stylo sont systématiquement exploités dans la reconnaissance. Par ailleurs, cette méthode se prête bien à une pédagogie interactive dans la mesure où l'utilisateur est autorisé à corriger et à modifier son écriture au fur et à mesure de son apprentissage vu que le système de reconnaissance répond immédiatement et affiche instantanément sur l'écran les résultats de l'écriture. Le système de reconnaissance est généralement constitué d'une tablette graphique et d'un stylo électronique qui indique les coordonnées du point cursif.

I.2.1.2 La reconnaissance hors-ligne (Off line)

La procédure utilisée pour la reconnaissance de l'écriture hors-ligne est similaire à celle utilisée en traitement d'images. En effet l'écriture est dans ce cas considérée comme étant une image binaire (ou avec des niveaux de gris) à deux dimensions dans laquelle la séquence temporelle du tracé des points a été perdue. Cette méthode se prête bien pour la reconnaissance des textes ayant été déjà imprimés ou écrits à la main depuis un moment.

En clair, l'approche utilisée dans cette reconnaissance peut être classée parmi les méthodes les plus générales utilisées pour reconnaître une écriture. Elle utilise les algorithmes de traitement

d'images : l'interprétation de l'information est indépendante de la source de génération. Ceci permet d'entrevoir de larges champs d'applications qui ne nécessitent pas un traitement spécifique en temps réel. En conséquence, cela exclut le recours à cette méthode dans les activités bancaires ou postales de reconnaissance des signatures et dans lesquelles l'information temporelle, qui est un élément déterminant dans ce cas de figure, n'est pas prise en compte dans le processus de traitement.

En résumé, dans le mode hors-ligne, l'ordre temporel du tracé de l'information n'est pas présent. Pour la reconnaissance, on doit plutôt chercher dans l'image-écriture les tracés décrivant le message écrit. Par contre dans le mode en-ligne, l'ordre du tracé, qui se présente comme une suite ordonnée de points-images, aide à enlever une grande ambiguïté liée à la reconnaissance de certains caractères.

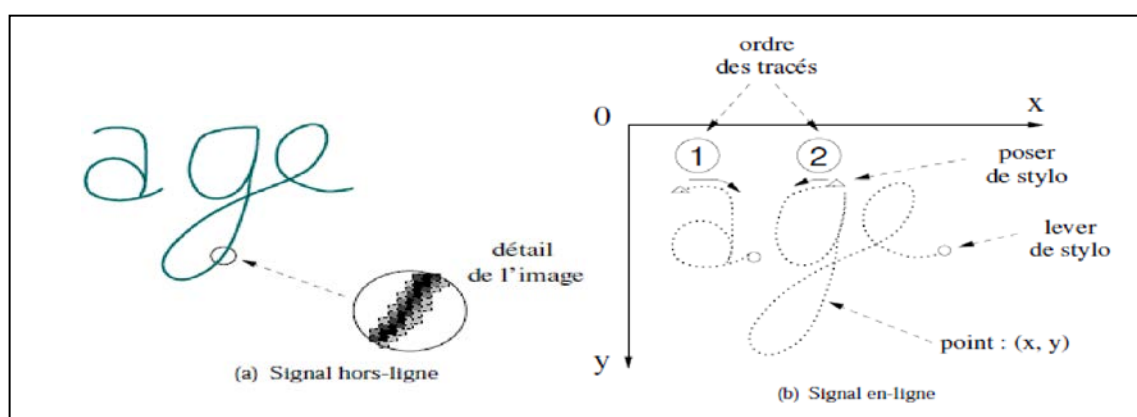


Figure I.1 (a) Ecriture hors ligne (ensemble de pixels)
(b) Ecriture en ligne (ensemble de coordonnées).

I.2.2 Style de l'écriture

Le style de l'écriture, qui est propre à la personne qui rédige, constitue un défi majeur dans la reconnaissance de l'écriture. En effet, le style de l'écriture, qui varie d'une personne à une autre, rend difficile le traitement de l'information en termes de clarté et de visibilité. D'autres difficultés liées au style de l'écriture concernent la langue traitée, l'existence de contraintes imposées au scripteur (à l'image des cadres, des cases ou les suites de points que l'on trouve dans les documents-formulaires), ou encore l'inexistence totale de contraintes (comme c'est le cas dans les documents fibres), constituent autant de défis qui s'ajoutent à la problématique de reconnaissance de l'écriture du point de vue du style de l'écriture.

A l'inverse, l'écriture imprimée, qui présente une régularité exemplaire contrairement à l'écriture manuscrite, a fait que les solutions proposées pour la reconnaissance ont donné des résultats très promoteurs.

En considérant les différents types de difficultés rencontrées dans le traitement de l'écriture hors-ligne, nous pouvons classer les méthodes de reconnaissance en plusieurs types :

I.2.2.1 Reconnaissance de caractères ou l'analyse de documents:

Dans le premier cas, le texte étant court, le traitement consiste dans un premier temps à repérer à travers ces quelques lignes les mots qui les composent. Dans un second temps, il s'agit de repérer dans chaque mot trouvé les caractères qui le composent. Dans le deuxième cas, s'agissant d'un document bien structuré, il importe de connaître au préalable la typographie ainsi que la mise en page de ce document avant d'entamer la procédure de traitement.

A. Reconnaissance de l'imprimé ou du manuscrit :

Suivant que l'on est en présence d'un texte imprimé ou d'un texte manuscrit, les techniques utilisées pour la reconnaissance diffèrent selon qu'on est dans le premier ou le second cas de figure. Dans un document imprimé, (Figure I.2.a) les caractères sont, dans le cas le plus général, alignés du point de vue horizontal et séparés par un espace régulier du point de vue vertical. Ces caractéristiques simplifient grandement le processus de reconnaissance combien même certaines particularités viennent distinguer certains documents selon que le texte est écrit avec une ou plusieurs fontes ou bien que ces fontes soient séparées ou accolées. En effet, la forme d'un caractère est déterminée par un style calligraphique qu'on appelle fonte et qui s'impose comme un prototype indispensable de distinction entre les caractères.

Pour résumer, les textes imprimés peuvent être *monofonte*, *multifonte* ou *omnifonte*.

Un système est dit monofonte lorsqu'il ne traite qu'un seul type de fonte à la fois. L'apprentissage n'est pas alors difficile à réaliser et il est permis d'atteindre des taux de reconnaissance très grands avec ce type de systèmes.

Nonobstant, un texte multifonte fait appel à plusieurs fontes dans le même document. Dans ces conditions, un système multifonte doit être en mesure de reconnaître un assortiment de quelques fontes parmi un ensemble de fontes préalablement apprises. Ce faisant, il doit donc être capable de gérer les ambiguïtés qui découlent des ressemblances intrinsèques qui existent entre les différentes formes de fontes.

Finalement, un système est appelé omnifonte s'il est capable de reconnaître tout type de fonte combien même ces fontes lui sont inconnues [1]. Autrement dit, l'étape d'apprentissage n'est pas absolument nécessaire.

Dans le cas où le document est un manuscrit, (Figure I.2.b et c) les caractères sont liés entre

eux (on dit qu'ils sont ligaturés) et leur dessin n'est pas très régulier. Il est inégalement proportionné. Cela résulte de la variabilité inter et intra-scripteurs. Afin de baliser la reconnaissance de l'écriture, différentes techniques de délimitations spécifiques sont généralement employées souvent combinées avec des connaissances contextuelles du document.

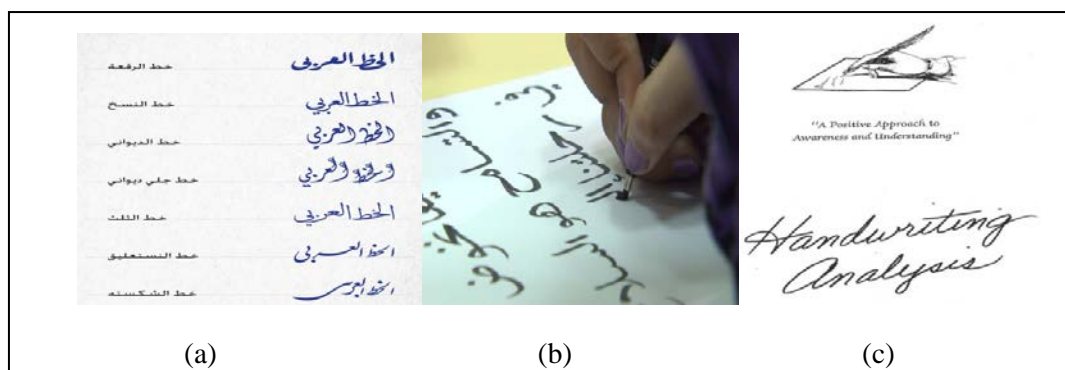


Figure. I.2 Styles d'écritures (a)Ecriture Arabe imprimé (b) Ecriture Arabe manuscrite (c)Ecriture latine manuscrite

I.3 Les difficultés de reconnaissance de l'écriture manuscrite

I.3.1 Nombre de scripteurs

Dans un système de reconnaissance de l'écriture, lorsque le nombre de scripteurs augmente, la reconnaissance devient de plus en plus difficile à gérer. En d'autres mots, autant le nombre de variétés de scripteurs est important (Droitier/Gaucher, Masculin/Féminin, Tranche d'âge, Types de profession), autant les styles d'écriture deviennent de plus en plus différents. Généralement, on classe les différents systèmes selon le critère suivant:

Système mono-scripteur: est un système destiné à apprendre et à reconnaître l'écriture d'un seul utilisateur.

Système multi-scripteur: Ce système est d'un niveau plus général que le précédent. Il permet l'apprentissage et de la reconnaissance de l'écriture d'un certain nombre de scripteurs bien précis. Dans ce cas, chaque classe de scripteurs est symbolisée par un modèle qui rassemble l'ensemble des variabilités liées aux habitudes de rédaction de chaque scripteur.

Système omni-scripteur: Ce système est bien plus général et que les deux précédents et naturellement bien plus proche de la lecture humaine. Toutefois, il est caractérisé par un degré de complexité maximale. L'objectif assigné à ce système étant d'apprendre à reconnaître l'écriture de tout scripteur. Ce type de systèmes exige un corpus d'apprentissage beaucoup plus important et nettement diversifié que les deux précédents systèmes afin d'aboutir en toute un modèle en mesure d'absorber cette variabilité abondante des styles d'écriture.

I.3.2 Taille du vocabulaire traité

Dans ce genre d'applications, la base de données des mots à reconnaître regroupe un lexique de taille limitée (<100 mots). La complexité du système à concevoir est ainsi moindre car la réduction du nombre de mots tend à réduire l'encombrement de la mémoire et permet l'utilisation de méthodes de reconnaissance directes et donc rapides. Cette taille réduite autorise un balayage systématique de l'ensemble des mots du lexique pour une meilleure reconnaissance. De tels systèmes de reconnaissance peuvent être rencontrés, à titre d'exemple, dans les services bancaires qui utilisent un vocabulaire réduit : les montants manuscrits dans les chèques.

Les applications à vocabulaire très étendu

Dans ce genre d'applications, le nombre de mots à reconnaître est composé d'un lexique de taille plus importante (>100 mots). Cette contrainte prélude à des problèmes d'encombrement de la mémoire et un allongement du temps d'accès à la mémoire pour chaque mot. Dans ce cas, il est judicieux de faire appel à des approches de type arborescentes avec des raffinements au fur et à mesure de l'analyse afin de permettre de reconnaître efficacement les caractères de l'écriture.

I.3.3 Contraintes morphologiques

Les difficultés de mise en œuvre d'un système de reconnaissance d'écriture augmentent en fonction de sa capacité, plus ou moins grande, à prendre en considération plusieurs types de modèles distincts et à tolérer des variantes par rapport à ces prototypes afin d'imposer de moins en moins de contraintes à l'utilisateur. On peut distinguer trois niveaux successifs :

- Un premier niveau correspondant à la reconnaissance d'un nombre limité d'entités isolées.
- Un second niveau correspondant à la reconnaissance d'entités plus globales.
- Le troisième niveau est celui de la reconnaissance globale des textes manuscrits qui peuvent comporter plusieurs types d'écritures différentes comme les retouches.

I.3.4 La disposition spatiale du manuscrit

La disposition spatiale du manuscrit est susceptible d'être influencée par des contraintes externes et des contraintes internes. Les contraintes externes peuvent être imposées au scripteur tandis que les contraintes internes peuvent découler des habitudes intrinsèques propres à chaque scripteur. Citons à titre d'illustration les différents types de contraintes internes et externes.

Les contraintes externes conduisant à une écriture:

- Pré-casée : L'utilisateur doit s'efforcer d'écrire à l'intérieur de cases prédéfinies dans le formulaire.
- Zonée : Le scripteur est obligé d'écrire dans des zones bien délimitées.
- Guidée : L'utilisateur est guidé dans son écriture par l'existence d'une ligne-support qu'il est tenu de suivre pour une meilleure lisibilité du texte.
- Générale : Le scripteur dispose d'un espace totalement libre dans lequel il peut écrire librement ses observations sans aucune contrainte spatiale.

Parmi les contraintes internes, on trouve:

- Lettres séparées: les mots écrits par le scripteur sont formés de caractères non attachés.
- Groupes de lettres liées à l'intérieur d'un mot : certaines lettres d'un mot sont liés et d'autres ne le sont pas.
- Mots entièrement liés: toutes les lettres d'un même mot sont liés.
- Les trois cas possibles à l'intérieur d'un même texte manuscrit : le texte comprend des mots avec des lettres séparées, des mots où les lettres sont complètement attachées et d'autres mots ou les lettres sont partiellement attachées.

I.3.5 Stratégies de reconnaissance de l'écriture

Lors de la reconnaissance de l'écriture, le système de reconnaissance dispose *a priori* de deux stratégies distinctes (stratégie globale et stratégie locale) qui peuvent être utilisées selon le type du texte en présence et pour lequel il peut être fait appel à l'une ou l'autre des stratégies selon la nature et la taille de l'unité considérée dans la modélisation mise en œuvre.

Primo : Le système reconnaît le mot comme une entité entière et indivisible. Dans ce cas, il s'agit d'une *approche globale* ou holistique.

Secundo : Le système reconnaît le mot à partir de ses caractères qui ont été segmentés au préalable. Il s'agit, dans ce cas, d'une *approche analytique*.

Une troisième alternative consiste à n'utiliser que certaines propriétés de l'une et l'autre des deux stratégies et raffine sa description du mot par rebouclage. Nous parlerons alors de systèmes basés sur la lecture humaine.

Le choix de l'une ou l'autre des deux approches (globale ou analytique) est, généralement, imposé par la taille du vocabulaire. Dans le cas où le système de reconnaissance est prévu pour de petits lexiques, il est alors envisageable de créer autant de modèles que de mots possibles. L'approche globale peut convenir dans ces conditions. Par contre, lorsqu'il s'agit de traiter des documents présentant un vocabulaire de plusieurs milliers de mots, il n'est pas

concevable de créer un modèle pour chaque classe ; il est plus pertinent de modéliser au niveau d'entités plus petites comme les caractères. Dans ce cas, le système doit segmenter l'écriture en entités de base.

I.3.5.1 L'approche globale:

Appelée encore méthode holistique. Cette approche considère les mots comme des entités à part entière, isolées les unes des autres. Chaque mot (ou entité) est décrit en faisant abstraction des caractères qui le composent. L'avantage notable de cette technique est de conserver spécialement le caractère dans son contexte avoisinant. Ceci permet une modélisation plus efficace qui se prémunit contre les variations de l'écriture et des dégradations éventuelles qui peuvent l'entacher. Néanmoins, cette méthode recèle des inconvénients qui limitent son utilisation. En effet, de par sa conception cette méthode est gourmande en espace mémoire, en temps de calcul. D'autre part, la complexité du traitement croît linéairement avec le nombre de mots du lexique considéré, d'où une limitation explicite du vocabulaire dans cette technique. [2-8].

I.3.5.2 L'approche analytique:

Dans l'approche analytique, le mot est partitionné en caractères ou en fragments morphologiques distinctifs, appelés graphèmes, dont la taille est inférieure à celle des caractères. La reconnaissance consiste d'abord à reconnaître, en premier lieu, les entités segmentées puis dans un second temps à étendre ce résultat vers une reconnaissance complète du mot. La difficulté de cette approche est directement reliée à la complexité de la segmentation [9].

D'une manière générale, il existe deux approches de segmentation : l'approche implicite et l'approche explicite.

I.4 Principales phases d'un système de reconnaissance d'écriture hors ligne

La reconnaissance de l'écriture consiste en la transformation d'un signal-écriture, se présentant sous forme imprimée ou manuscrite, en une représentation symbolique, en prenant en considération ses diverses natures spatiales. Pour ce faire, un certain nombre d'étapes à mettre en œuvre sont indispensables. L'ensemble de ces étapes constituent en général la structure fondamentale du système de reconnaissance d'écriture que l'on peut schématiser par le diagramme de la **Figure I.3** ci-dessous.

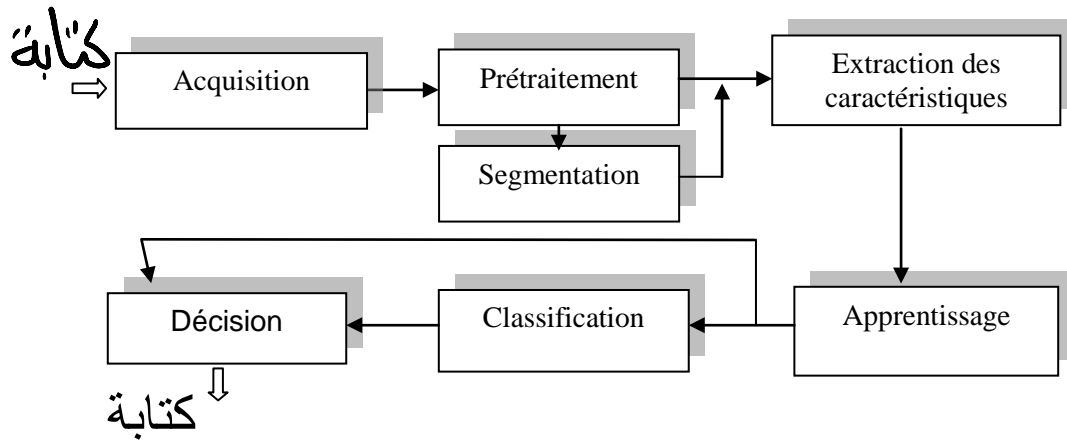


Figure I.3 Structure générale d'un système de reconnaissance d'écriture

L'*acquisition* est la phase dont le but est la numérisation du signal de l'entrée au moyen d'un scanner classique.

La phase des *prétraitements et normalisation* a pour objectif la réduction des bruits et la diminution des variabilités de l'écriture afin de n'en conserver que les informations pertinentes.

L'étape de *segmentation* permet, quant à elle, la décomposition du texte-image en plusieurs entités (mots, caractères ou graphèmes) en vue de réduire ultérieurement la complexité des systèmes de traitement.

La phase d'*extraction des primitives* permet de faire, à partir de la représentation de l'image, une description synthétique de la forme à reconnaître dans un espace à plusieurs dimensions.

La phase *apprentissage* consiste à découvrir les modèles les plus aptes à convenir aux signaux-entrées du système. À l'issue de cet apprentissage, le résultat obtenu est une base de données qui constituerait la base de référence du système.

Pendant la phase de *classification*, la description du caractère à reconnaître, récupéré de la base de test, est comparé aux descriptions des caractères de la base de référence.

Outre la présence de ces phases, il est possible de rencontrer dans certains systèmes de reconnaissance une autre phase : la phase de *post-traitement*. Cette étape de traitement, qui est optionnelle, a pour but de vérifier et à corriger, éventuellement, l'hypothèse de reconnaissance générée par le module de classification en utilisant des informations contextuelles telles que : les informations lexicales, syntaxiques, sémantiques, pragmatiques, etc.

À présent, nous allons passer en revue les techniques de mise en œuvre de chaque phase constituant le système de reconnaissance.

I.4.1 Acquisition

Pour les systèmes de reconnaissance de l'écriture Hors-ligne, les scanners et les caméras sont les systèmes d'acquisition les plus répandus. Ils sont sollicités afin de permettre la conversion des informations contenues dans le document papier en une forme d'image numérique (bitmap). Cette étape est cruciale, dans la mesure où c'est à ce niveau que la perte d'informations capitales peut s'opérer en raison d'un mauvais paramétrage de ce matériel. L'on doit donc se préoccuper sérieusement, lors de la préparation des documents à saisir, du choix et du paramétrage du matériel de saisie (scanner) ainsi que du format de stockage des images. Pour ces raisons, est-il utile de rappeler les définitions d'un certain nombre des caractéristiques essentielles d'une image numérique.

Résolution: une résolution normale doit être de 300 dpi (dpi : dots per inch = nombre de points par pouce est une unité de précision communément utilisée pour définir la résolution). Toutefois, lorsque la taille de l'écriture est petite, il importe d'augmenter la résolution.

Niveau d'éclairage: L'ajustement du scanner de manière à rendre l'image plus claire fait réduire, il est vrai, le bruit. Toutefois, cela engendre également la disparition de minces traits.

I.4.2 Prétraitement et normalisation

Comme cela a été mentionné dans la section précédente, de nombreux facteurs influent sur la variation et la qualité de l'écriture tels que la personnalité de l'écrivain, le contexte d'écriture, les dispositifs de capture (différente résolution et/ou méthode de capture), ... Par voie de conséquence, une étape de prétraitement est souvent nécessaire lors de l'opération de reconnaissance de l'écriture afin de réduire les bruits et la distorsion qui apparaissent dans l'écriture, de diminuer la variation des styles afin de faciliter les traitements ultérieurs tels que la segmentation et l'extraction de primitives. De manière générale, les méthodes de prétraitement peuvent être divisées en deux catégories : élimination du bruit (filtrage, lissage, ...) et normalisation de la forme (correction de la pente, normalisation de la taille, ...).

I.4.2.1 Filtrage

Le filtrage consiste à atténuer le bruit qui entache une image originale afin d'améliorer sa qualité visuelle. Le filtrage est obtenu en changeant la valeur du niveau de gris d'un pixel donné en fonction de la valeur des pixels voisins. Deux types de filtrages coexistent: filtrage linéaire et filtrage non linéaire. Citons à titre d'exemple: Le filtre linéaire moyen qui corrige une valeur par la valeur moyenne locale et le filtre médian, les filtres morphologiques, etc. le choix du filtre dépend du type de l'image traitée

I.4.2.2 Binarisation

Pour les images acquises en niveaux de gris, la binarisation de l'image est une étape préalable nécessaire qui s'impose avant le début de tout traitement. Elle permet de mieux distinguer les caractères du fond. La binarisation consiste à attribuer à chaque pixel une valeur binaire soit 0 soit 1. Le '0' représente le noir (ou plus spécifiquement les caractères du texte) et le '1' représente le blanc (ou tout simplement le fond de la page). Avant d'aborder l'étape de binarisation, l'on se doit d'effectuer en premier lieu l'opération de seuillage [10].

Seuillage : L'opération de seuillage consiste à déterminer la valeur seuil à partir de laquelle tous les pixels ayant un niveau de gris inférieur à cette valeur sont forcés à zéros '0'. Autrement dit, ils sont remplacés par du noir et formeront donc le tracé de l'écriture. Parallèlement, tous les pixels dont le niveau de gris est supérieur à ce seuil prendront la valeur un '1' (qui symbolise le blanc de la page).

On distingue deux types de seuillage : le seuillage global et le seuillage adaptatif.

Le seuillage global prévoit de prendre un seuil identique pour toute l'image. Chaque pixel de l'image est comparé à ce seuil. Pour le seuillage automatique d'une image à niveaux de gris [1], une méthode consiste à tracer un histogramme de ces niveaux de gris et choisir ensuite le seuil au fond de la vallée qui sépare le pic correspondant au niveau de gris du fond et le premier pic suivant (voir Figure I.4).

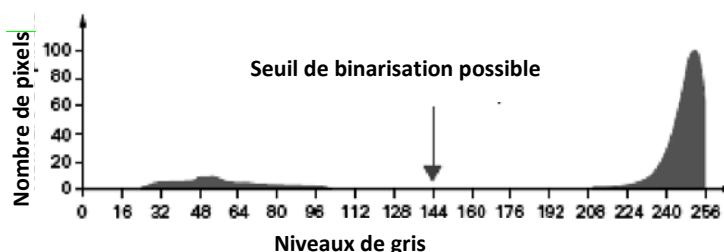


Figure I.4 Choix de seuil sur l'histogramme de l'image à niveaux de gris

Les limites du seuillage peuvent être résumées dans les points suivants :

- Mauvaise qualité de l'impression : Dans le cas où la qualité d'impression du texte n'est pas uniforme tout au long de la page, une mauvaise appréciation de la valeur du seuil peuvent engendrer la perte partielle d'un certain nombre de caractères ;
- Présence d'un bruit de fond : Des taches parasites risquent d'apparaître si jamais le fond est non homogène ou bruité ;
- Présence d'un fond coloré : Lorsque la page contient des encarts de couleur (texte imprimé sur fond coloré), tout le texte présent dans ces blocs risque de disparaître à jamais.

Les photographies qui subiraient un tel traitement de seuillage ne seraient plus

reconnaissables. Nous verrons alors apparaître une mosaïque de zones noires et blanches. Afin de pallier à ce genre de problèmes, la solution consisterait à recourir à un seuillage adaptatif d'une part et la séparation préalable des photographies du texte d'autre part.

Seuillage adaptatif: Le seuillage adaptatif consiste à faire évoluer le seuil localement et régulièrement en fonction du niveau de gris des pixels du document. Plusieurs solutions de seuillage adaptatif ont été élaborées afin de restituer de manière correcte le texte original dans tous les cas de figure. Pour ce faire, on distingue des méthodes qui ont recours au découpage de l'image en régions et des méthodes locales proches du traitement du signal.

1.4.2.3 Lissage

Le lissage sert à rattraper les imperfections et anomalies qui résultent suite à l'étape de binarisation en procédant, au niveau local sur les pixels, à des opérations de nettoyage et de bouchage. Ces opérations ont été introduites, dans le cas de la reconnaissance de l'écriture, par [1]. Ces démarches sont fondamentales dans la mesure où elles tendent à minimiser au plus haut point les discontinuités insérées dans l'image, à rendre le pourtour de l'écriture plus régulier et éliminer en outre l'effet d'escalier.

Le nettoyage est obtenu de plusieurs manières et dans tous les cas il est nécessaire d'apposer à chaque fois un masque 3 x 3 (Figure I.5) qui provoque l'élimination des tâches de faibles dimensions et autres excroissances présentes dans la forme.

A ce propos, nous ferons appel aux notions suivantes pour décrire les masques de transformation de voisinage : '1' pour désigner un point qui appartient à la forme ; '0' pour désigner un point qui appartient au fond et finalement X pour désigner un point quelconque.

En fonction du type de bruit à supprimer, différentes techniques peuvent être préconçues [1]. Nous proposons, à présent, une technique qui permet de nettoyer les petites tâches et d'éliminer les excroissances de la forme.

- L'application sur l'image d'un masque de type **-a-**, suivant les huit directions, conduit à l'élimination des points qui sont isolés de la forme ou bien qui sont situés à l'extrémité des contours.
- L'application sur l'image d'un masque de type **-b-**, suivant les huit directions, permet de supprimer les points formant des angles droits ou encore les excroissances du contour.
- Pour l'opération de bouchage, il suffit d'égaliser dans un premier temps les contours et de boucher ensuite les trous qui apparaissent à l'intérieur de la forme des caractères en les remplissant de points noirs. A cet égard, nous distinguons 2 types de bouchage :

- Le bouchage de trous isolés qui nécessite l'utilisation d'un masque de type **-c-**.
- La correction des irrégularités des contours à la décomposition de la forme qui nécessite l'utilisation d'un masque de type **-d-**.

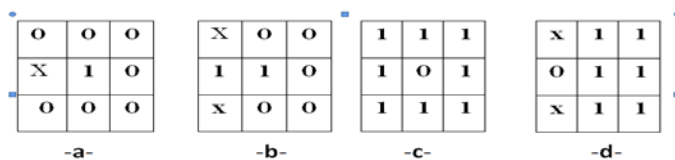


Figure I.5. Les différents masques pour les opérations de nettoyage d'un mot

Il est utile de rappeler qu'il est fait habituellement appel aux opérateurs de la morphologie mathématique pour les opérations de lissage [1]. Ainsi, l'opérateur de fermeture sert à supprimer les trous de petites dimensions et à remplir les régions vides du contour, tandis que l'opérateur d'ouverture est prévu pour découper les isthmes étroits et supprimer les petits îlots, les pics anguleux ainsi que les capes.

I.4.2.4 Normalisation

Une étape primordiale dans la reconnaissance de l'écriture consiste à rendre ses différentes opérations le moins possible sensibles aux variations de styles chez les scribes. Pour ce faire, c'est à l'étape de normalisation qu'est dévolue cette mission où elle y est assignée à faire réduire au maximum les effets de variabilités liées aux styles, tailles et orientations d'écriture du scribe. L'importance de cette phase réside dans le fait que l'étape ultérieure d'extraction des primitives soit la moins sujette aux influences des variations de styles dans les écritures.

Les différents types de traitements de normalisation sont résumés ci-après :

- Normalisation de la taille ;
- Correction de l'inclinaison des lignes (*Skew correction*) ;
- Correction de l'inclinaison des caractères (*Slant correction*) ;
- Estimation de la ligne de base.

Un exemple de normalisation des différents types de variations de style que l'on peut rencontrer dans la pratique, et qui sont traitées avec différentes techniques, est présenté dans [11-13].

I.4.2.5 Squelettisation

La squelettisation (ou amincissement) est une opération fondamentale que l'on rencontre dans plusieurs systèmes de reconnaissance d'écriture [14-19]. Elle s'occupe d'amincir les formes à leurs plus simples expressions (squelettes), d'exprimer les relations structurelles dans la forme et de réduire en fin de compte l'espace et le temps de traitement en simplifiant les structures des données.

De manière générale, les systèmes off-line de reconnaissance d'écriture manuscrite utilisent de manière intensive la squelettisation afin de réduire, au maximum, les effets de variabilité intrinsèques aux différents styles d'écriture [17]. Ces systèmes s'appuient également sur le squelette d'un mot de sorte à en déduire une représentation on-line [19].

I.4.2.6 Détection des lignes de référence

A partir de l'histogramme horizontal de l'image, on commence par diviser chacun des mots en trois zones horizontales. On obtient ainsi une zone des extensions hautes, une zone médiane et une zone des extensions basses (Figure I.6).

La zone médiane est, sans conteste, la plus importante. En effet, c'est à ce niveau que l'on rencontre, non seulement, la concentration la plus importante de pixels mais également le pic le plus haut de l'histogramme horizontal de l'image. Ce dernier est généralement considéré comme étant la ligne de base médiane [13], [22]. En outre, cette zone médiane est délimitée par deux lignes : la ligne de base supérieure et la ligne de base inférieure. La zone des extensions hautes est située au dessus de la ligne de base supérieure. Parallèlement, la zone des extensions basses est située en dessous de la ligne de base inférieure.

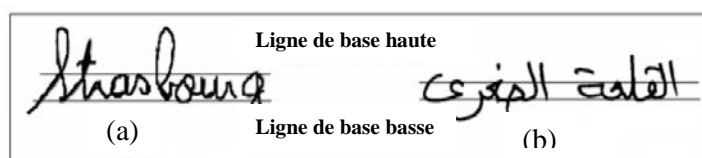


Figure I.6 Exemples de lignes de base d'écriture (a) cas de l'écriture latine, (b) cas de l'écriture arabe

I.4.3 Segmentation

La segmentation est l'étape qui consiste à découper l'image d'une écriture en entités (mots, caractères ou graphèmes). Ces entités font partie d'un alphabet prédéfini selon l'objectif recherché. D'un autre côté, la segmentation permet également de simplifier le degré de complexité des modules de traitement auxquels on fera appel ultérieurement pour la suite des opérations [20], [21].

Pour cette raison, il est admis que la segmentation est considérée comme une étape très

cruciale dans tous les systèmes de reconnaissance [22]. En effet, la séparation des lignes, des mots, des pseudo-mots, des caractères et des graphèmes sont des opérations très laborieuses et excessivement coûteuses en temps de calcul, tant les styles d'écriture sont nombreux, les lignes sont parfois embrouillées et les caractères généralement attachés les uns aux autres (comme c'est le cas pour l'écriture arabe qui est semi-cursive, par excellence).

Pour effectuer la segmentation, les chercheurs se retrouvent le plus souvent devant un choix décisif à prendre qui, s'il n'est pas correctement pris, risque de déboucher sur deux résultats tout deux problématiques : une sur-segmentation ou une sous-segmentation. La sur-segmentation peut aboutir à une fausse détection. En d'autres termes, le caractère est coupé en plusieurs petits segments qui peuvent rendre compliquée la reconnaissance de l'écriture et fournir de fausses interprétations. La sous-segmentation, quant à elle, risque d'aboutir à une non détection du tout. En effet, l'on pourrait se retrouver avec des entités de segmentation qui seraient à cheval sur deux caractères consécutifs. Ce qui poserait un problème réel car on ne sera pas en mesure de reconnaître correctement certains caractères qui composent ce mot.

La littérature spécialisée contient généralement deux genres de techniques de segmentation : la segmentation explicite et la segmentation implicite. Dans la première technique, le système de reconnaissance essaie de trouver explicitement les Points de Segmentation Potentiels (PSP) avant de découper les données d'entrée en une séquence de caractères ou de graphèmes. Le système de reconnaissance reposant sur la seconde technique utilise, par contre, une fenêtre glissante pour découper les données d'entrée en une séquence d'observations.

Nous allons maintenant présenter brièvement ces deux méthodes de segmentation dans les sections suivantes.

I.4.3.1 Segmentation explicite

En fait, il existe deux méthodes de segmentation explicite : la première méthode est purement explicite, la seconde a recours à un duo segmentation/reconnaissance.

Dans les méthodes de segmentation explicites pures, le mot manuscrit entré est instantanément décomposé en caractères individuels. L'on alors obtient en sortie une séquence de graphèmes dont chacun d'entre eux est susceptible d'être considéré comme un caractère éventuel. Par la suite, un système de reconnaissance est utilisé pour reconnaître chaque fragment ainsi segmenté [23-26].

Les systèmes de reconnaissance basés sur une segmentation explicite pure requièrent une procédure de segmentation très au point. Cette étape est primordiale car les erreurs de

segmentation peuvent aboutir à des reconnaissances erronées des mots. Toutefois, étant donné les complications inhérentes à la technique elle-même, il n'existe guère d'algorithme de segmentation irréprochable.

Dans le but de trouver une solution à ce problème, des techniques explicites de segmentation/reconnaissance ont été mises au point [27], [28]. Ces méthodes ont pour objectif d'intégrer un système de reconnaissance au cours du processus de segmentation lui-même, dans le but de diminuer les erreurs provoquées par les limites des techniques de segmentation. Plusieurs méthodes de segmentation peuvent être trouvées dans la littérature, à l'exemple des méthodes qui reposent sur l'analyse du contour [20] (Cette technique a été en fait imaginée pour une reconnaissance des données hors ligne) ainsi que les techniques fondées sur la trajectoire [30], [31] (Cette technique a été en fait inventée pour le signal en ligne). Les méthodes de reconnaissance produisent un ensemble de graphèmes utilisés en plus afin de créer un treillis à des niveaux L qui correspondent au mot manuscrit d'entrée.

I.4.3.2 Segmentation implicite

En réalité, les systèmes basés sur la segmentation implicite n'ont recours à aucun module de segmentation à proprement parler. Ils font, en fait, appel à une fenêtre glissante qui analyse les données de l'écriture manuscrite. Dans ces conditions, un ensemble de fenêtres segmentées sont produites à la fin de cette analyse. A chaque position d'une fenêtre, une liste de caractéristiques en est alors déduite et qui serviront à alimenter un modèle de classification dynamique formé au préalable pour chacun des mots du lexique.

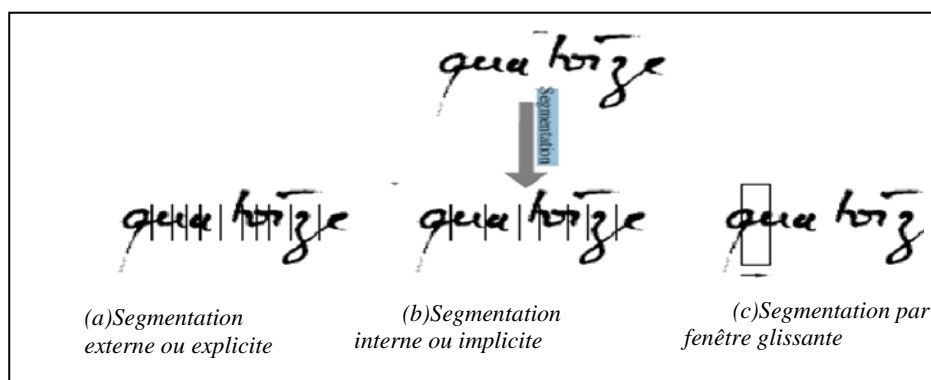


Figure I.7 Différentes segmentations pour le mot quatorze [32]

I.4.4 Extraction de primitives

L'extraction des primitives, habituellement dénommées caractéristiques, revient à transformer les données d'entrée (mots, caractères, graphèmes) en un vecteur de primitives de dimensions fixes. C'est une étape très importante et très délicate dans les systèmes de reconnaissance.

L'objectif attendu de cette étape est la recherche de l'information pertinente, discriminante et de dimension limitée pour la suite des opérations de classification. Le choix de cette information ne doit pas être fait au détriment d'un risque éventuel de perdre des informations essentielles et significatives. En effet, un choix erroné des primitives influencerait de manière négative et notable sur les résultats et ce même si on utilise un classifieur très performant. Selon [29] l'extraction des caractéristiques est un « *problème d'extraction à partir de l'image, de l'information la plus pertinente, pour un problème de classification donné, c'est-à-dire celle qui minimise la variabilité intra-classe et qui maximise la variabilité interclasse* ». Cette étape est cruciale pour les raisons que l'on expose ci-dessus :

- Une mauvaise représentation des données par des primitives (déficit en caractéristiques) mène à des résultats non satisfaisants dans la reconnaissance.
- Une répétition des redondances de la représentation mène à des confusions (la liste des échantillons valables pour l'apprentissage et classification étant limitée).

Pour toutes ces raisons, très nombreux sont les travaux récents [33-35] qui sont orientés vers la recherche de techniques d'extraction de primitives efficaces, et ce eu égard à l'importance de la sélection d'un groupe de primitives discriminantes pour la conception des systèmes de reconnaissance solides et performants (qui possèdent de forts taux de reconnaissance).

Dans la littérature, les primitives sont en général regroupées en trois catégories fondamentales: les caractéristiques statistiques, les primitives structurelles (topologiques) et les transformations globales.

I.4.4.1 Caractéristiques statistiques

Les primitives statistiques servent à distribuer l'information sur l'image en entier. Elles sont en général engendrées par l'estimation de la densité de pixels qui appartiennent à l'image en entier (caractères ou mots) ou encore à seulement quelques parties de l'image. Pour ce faire, on se sert des outils statistiques tels que l'entropie, la moyenne, la variance, etc. associées à des caractéristiques locales [36],[37] qui peuvent être, par exemple, l'histogramme directionnel qui fournit le nombre de pixels sur une ligne dans une direction quelconque de l'image.

Le principe du "zoning" permet de fournir, au niveau local, les moyennes de pixels d'une fenêtre dans l'image.

- Les projections verticales et horizontales des densités de pixels noir/blanc représentant le nombre de pixels pour chaque ligne ou chaque colonne de l'image.
- Les directions des contours dans une fenêtre locale.

- L'histogramme des transitions noir/blanc et vice-versa qui, comme son nom l'indique, permet de fournir le nombre de transitions 0-1 et 1-0 entre pixels.

La description des primitives statistiques est axée sur les densités de pixels, du fait que ces caractéristiques sont considérées comme étant moins discriminantes comparées aux autres types. En effet, celles-ci sont porteuses de peu d'informations sur le signal d'écriture. Récemment, des recherches publiées dans la littérature tendent vers des primitives hybrides qui font intervenir une combinaison de différents types (statistiques, géométriques, structurelles, ...) et ce dans le but de compenser mutuellement les faiblesses des techniques des unes par les autres.

I.4.4.2 Caractéristiques structurelles

Les propriétés topologiques et géométriques de la forme sont représentées par les primitives structurelles. Ces caractéristiques sont engendrées à partir de la représentation de la forme par le squelette ou le contour et non à partir de l'image brute de l'écriture. Ces caractéristiques concernent particulièrement les *concavités, convexités, occlusions, ascendants, descendants, composantes connexes, segments de droites et leurs attributs (position, orientation, ...), mesures de pentes, arcs, boucles, croisements, jonctions des traits, paramètres de courbures, angularités, points extrêmes et points terminaux, longueur et épaisseur des traits, surfaces et périmètres*. Contrairement au type précédent, les primitives structurelles ont une capacité discriminative très importante. Cette propriété a une influence tangible sur la rapidité de la décision de l'appartenance d'une classe lors de la classification. Par voie de conséquence, une utilisation maladroite de ces primitives (mauvaise détection de leur présence) engendre fatalement des résultats non satisfaisants lors du processus de reconnaissance.

I.4.4.3 Les transformations globales et les moments

Les transformations globales étaient jadis usuellement très utilisées dans le domaine du traitement de signal. Elles étaient utilisées pour le changement d'espace de représentation de l'image (caractères ou mots) afin de faciliter l'extraction des caractéristiques principales. A ce titre, les méthodes les plus communément employées dans la reconnaissance de l'écriture sont:

- **La transformation de Hough** : Mise au point au départ par *Paul Hough* en 1962 [38] afin de déceler la présence d'éventuelles formes communes (droites, cercles, ellipses, ...). Dernièrement, cette technique est intensivement employée dans le domaine de la reconnaissance d'écriture en raison de sa robustesse et de son efficacité dans

l'extraction des primitives du niveau haut.

- **La transformée de Fourier** : est certainement la transformée la plus utilisée dans le domaine de reconnaissance de l'écriture [39]. Elle est invariante aux rotations et aux changements d'échelle. Dans cette transformation, les descripteurs de Fourier, qui sont représentés par les coefficients complexes des séries de Fourier, sont ordinairement les caractéristiques recherchées.
- **La transformée en ondelettes** : est très similaire à celle de Fourier [39]. Néanmoins, elle n'est pas invariante à la translation. C'est pour cette raison qu'elle soit moins employée dans le domaine de la reconnaissance. En effet, un léger décalage modifie complètement les coefficients d'ondelettes. Ces derniers permettent l'obtention d'une information fréquentielle localisée pour un signal en entrée (image de caractères ou de mots).

Nous rappelons que dans cette catégorie figure également la grande famille des *moments invariants*. Ces outils ont des caractéristiques très intéressantes car ils sont invariants à la rotation, à la translation et au changement d'échelle [40]. Les primitives engendrées à partir de cette famille résultent d'une transformation réalisée globalement au niveau des pixels de l'image. Elles sont principalement calculées à l'aide des mesures statistiques de la distribution des pixels autour du centre de gravité du caractère. Les *moments invariants* les plus utilisées dans la reconnaissance de l'écriture sont :

➤ **Les moments de Hu**

Les sept moments invariants proposés par Hu représentent une famille de primitives statistiques largement utilisée dans la reconnaissance de l'écriture. Ces primitives sont invariantes par translation, rotation et changement d'échelle. Les moments invariants fournissent une information générale sur la distribution des pixels d'un caractère autour de son centre de gravité [41]. Les sept moments introduits par Hu sont présentés dans ce qui suit :

L'équation générale des moments est la suivante :

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (I.1)$$

Avec p et $q \in \mathbb{N}$ et $f(x, y)$ est le niveau de gris de l'image.

Le *centre de gravité* d'une image « \bar{x}, \bar{y} » est défini par les relations suivantes :

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (I.2)$$

Où m_{00} est la surface de la forme dans le cas d'une image binaire.

Afin de rendre les moments m_{pq} invariants par rapport à la translation, on les définit en choisissant « \bar{x}, \bar{y} » comme origine. On parle alors de moments centrés μ_{pq} dont la définition formelle est la suivante :

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q \cdot f(x, y) dx dy \quad (I.3)$$

Pour une image discrète, les relations (II.5) et (II.6) deviennent :

$$m_{pq} = \sum_x \sum_y x^p y^q \cdot f(x, y) \quad (1.4)$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q \cdot f(x, y) \quad (1.5)$$

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

$$\mu_{00} = m_{00}, \mu_{10} = \mu_{01} = 0, \mu_{11} = m_{11} - (m_{00} \cdot \bar{x} \cdot \bar{y})$$

Les moments centrés normalisés sont définis par :

$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \text{ Avec } \gamma = \frac{p+q}{2} + 1 \quad (I.6)$$

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

$$\begin{aligned} \phi_1 &= n_{20} - n_{02} \\ \phi_2 &= (n_{20} - n_{02})^2 - 4n_{11}^2 \\ \phi_3 &= (n_{30} - 3n_{12})^2 + (3n_{11}^2 - n_{03})^2 \\ \phi_4 &= (n_{30} - 3 \cdot n_{12})^2 + (3 \cdot n_{21}^2 - n_{03})^2 \\ \phi_5 &= (n_{30} - n_{12}) \left[(n_{30} - 3n_{12})^2 - (3n_{21} - n_{03})^2 \right] (n_{21} - n_{03})(3n_{11} - n_{03}) [3(n_{30} + n_{12})^2 - \\ & (n_{21} + n_{03})^2] \\ \phi_6 &= (n_{20} - n_{02})^2 [(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03}) \\ \phi_7 &= (3n_{21} - n_{03})(n_{30} + n_{12}) [(n_{30} - 3n_{12})^2 - (3n_{21} - n_{03})^2] + 3(n_{12} - n_{30})(n_{21} + \\ & n_{03}) [3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \end{aligned}$$

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

➤ *Les moments de Zernike*

Les moments de Zernike ont été introduits pour la première fois en 1934 par Zernike [42], tandis que leur utilisation dans l'analyse des images a été introduite par Teague [43]. Les moments de Zernike se sont révélés meilleurs et plus robustes en présence de bruit. Du moment que leurs fonctions de moment soient déterminées en faisant appel à la représentation en coordonnées polaires dans l'espace d'image, les moments de Zernike sont ainsi souvent employés dans les opérations de reconnaissance qui nécessitent une invariance dans la rotation. Les moments de Zernike se révèlent être une bonne représentation des caractéristiques et produisent nettement plus d'informations sur l'image faciale tout en

réduisant la dimension du vecteur caractéristique ce qui conduit à des résultats tout simplement meilleurs.

Les polynômes de Zernike se présentent comme un ensemble de polynômes orthogonaux définis à l'intérieur du cercle unité. Les moments de Zernike bidimensionnels d'ordre p avec une répétition q de la fonction d'intensité d'image $f(r, \theta)$ sont définis comme suit :

$$z_{pq} = \frac{p+1}{\pi} \int_0^{2\pi} \int_0^1 R_{pq}^*(r) e^{-jq\theta} f(r, \theta) r dr d\theta, |r| \leq 1 \quad (\text{I.7})$$

$$\text{Ou } R_{pq}^*(r) = \sum_{k=0}^{(p-|q|)/2} \frac{(-1)^k (p-k)! r^{p-2k}}{k! ((p+|q|)/2 - k)! ((p-|q|)/2 - k)!}$$

I.4.5 Classification

La classification consiste en l'élaboration d'une règle de décision transformant les attributs qui caractérisent les formes en appartenance à une classe (passage d'un espace de codage vers un espace de décision). Avant qu'un modèle de décision n'intègre un système de reconnaissance de l'écriture, il faut veiller auparavant à ce que deux étapes soient d'abord exécutées : l'étape d'*apprentissage* et l'étape de *test*.

L'étape d'apprentissage est une opération qui permet de caractériser les classes de formes de sorte à bien distinguer les familles de formes homogènes. Ce traitement est une étape cruciale dans le système de reconnaissance. On distingue deux types d'apprentissage : apprentissage *supervisé* et apprentissage *non supervisé* [44].

Dans la méthode de l'apprentissage supervisé, un spécimen typique représentant l'ensemble des formes à reconnaître est présenté au module d'apprentissage. Chaque forme est alors marquée d'une vignette par un opérateur que l'on appelle généralement professeur. Cette vignette permet de préciser au module d'apprentissage la classe dans laquelle le professeur souhaiterait que la forme soit classée. Cette phase d'apprentissage est prévue pour rechercher les similitudes entre les éléments d'une même classe d'une part et les dissimilitudes entre les éléments de classes différentes afin de produire la meilleure subdivision possible de l'espace des représentations. Les paramètres qui décrivent cette partition sont alors sauvegardés dans une table d'apprentissage à laquelle le module de décision reviendra par la suite pour trier les formes qui lui seront présentées.

Dans la méthode de l'apprentissage non supervisé, il est fourni au système de reconnaissance un ensemble de formes non étiquetées. Pendant l'étape de la classification, il sera procédé à l'identification automatique des formes qui appartiendraient à la même classe.

Dans le domaine de la reconnaissance de l'écriture, ce sont les techniques qui ont recours à l'apprentissage supervisé qui sont le plus généralement employées et plus précisément, dans

le cas où les caractères manuscrits sont isolés les uns des autres parce que les classes naturelles sont connues d'avance et sont en nombre restreint.

Pendant l'étape de test on procède à l'évaluation de la performance du classifieur afin de réaliser un apprentissage spécifique. Cette étape est d'une importance vitale parce qu'elle peut complètement fausser le choix des primitives ou encore celui de la méthode d'apprentissage. En pratique, il n'est pas toujours aisé de deviner de prime abord les primitives les plus adéquates et la technique d'apprentissage qui convient le mieux au problème posé. De cela résulte l'intérêt de faire des approches par itérations successives. Celles-ci consistent à trouver les primitives qui soient les plus pertinentes à la reconnaissance et à trouver ensuite la solution convenable au système avant de procéder à son test pour évaluer sa performance vis-à-vis du système en considérant l'ensemble de ces primitives. Il est évident que les performances désirées du système ne soient pas atteintes du premier coup, il importe de recommencer une fois encore et essayer à nouveau une autre liste de primitives et/ou combiner les primitives extraites avec d'autres qui sont nouvelles.

Les étapes de l'apprentissage et de test sont exécutées en tout premier lieu et ce avant toute intégration du module de décision dans le système de reconnaissance. En tout état de cause, il est permis et même fortement conseillé de réitérer les phases d'apprentissage et de test autant de fois que c'est nécessaire jusqu'à ce qu'à l'atteinte des performances recherchées. Il faut noter que la performance du classifieur dépend étroitement de la performance atteinte lors de ce calcul. Aussi est-il important de bien construire le classifieur. A cet effet, trois méthodes sont offertes pour la mise au point du classifieur, les approches structurelle, statistique et hybride.

I.4.5.1 Les approches statistiques

Il s'agit dans cette méthode de trouver les caractéristiques relatives à une forme afin de les analyser d'une manière statistique. Par voie de conséquence, cette technique exige un grand nombre d'exemples dans le but d'effectuer un apprentissage le plus correct possible des lois de probabilité de toutes les classes. En d'autres termes, cette méthode profite des techniques d'apprentissage automatiques qui reposent sur des assises théoriques bien rodées telles que la théorie de la décision bayésienne, les méthodes de classification non supervisées et l'analyse en composantes principales. Il existe deux familles de méthodes essentielles qui emploient les approches statistiques en l'occurrence les méthodes *paramétriques* et les méthodes *non paramétriques*.

- Les approches paramétriques considèrent, par hypothèse préalable, que les classes analysées obéissent à une distribution de probabilité suivant une certaine forme connue au départ. Dans ces conditions, la prise de décision revient à trouver la classe qui répond au mieux à la forme inconnue et qui présente la probabilité d'appartenance maximale. Ces approches nécessitent des bases d'apprentissage de taille assez considérable afin d'aboutir à une juste estimation des paramètres de la distribution supposée. A cet effet, l'approche statistique comprend la règle de Bayes [12], les méthodes neuronales et finalement les chaînes de Markov [19].

- Pour les approches non paramétriques, les lois de probabilité ne sont pas connues pour une des classes. Aussi, il importe d'ajuster des frontières de décision entre les différentes classes. Pour ce faire, les méthodes les plus connues sont : la méthode du plus proche voisin [12] d'une part et la méthode d'appariement de graphes [12] d'autre part. Pour de plus amples informations concernant ces deux techniques, il serait utile de consulter les travaux de Kunt [20] qui a décrit un ensemble d'approches statistiques dans la reconnaissance des formes.

Quoique les deux méthodes statistiques soient de nature distincte, il est possible de les combiner pour le même domaine d'applications. Le choix d'une méthode plutôt qu'une autre peut être assujéti à des limites propres aux deux méthodes telles que la dimension de la base d'apprentissage de travail, le temps de calcul demandé ainsi que la taille mémoire requise. L'utilisation simultanée des deux approches peut s'avérer être la solution idéale pour l'opération de reconnaissance de l'écriture.

I.4.5.2 Les approches structurelles

Les approches structurelles tentent d'établir une relation entre la structure des formes sous analyse avec la syntaxe d'un langage formel. La description des formes est effectuée au moyen de phrases. Au même moment, le problème de classification se trouve réduit à un exercice d'analyse de grammaire (*parsing*). De toutes les manières, les approches quelles soient syntaxiques ou structurelles tendent à décrire des formes compliquées au moyen de formes simples. Ces dernières, que l'on appelle aussi caractéristiques, sont instantanément déduites des données qui se présentent à l'entrée du système. Ce qui distingue principalement les approches structurelles des approches statistiques est que les caractéristiques extraites à partir de la méthode structurelle sont des formes de base très simplifiées alors que dans les techniques statistiques, ces caractéristiques sont des données de type mesure. Autre élément de différence entre les deux approches est que la présente technique fait intervenir la notion d'ordre dans sa description d'une forme. En définitive, les méthodes structurelles les plus

communément utilisées s'appuient sur le calcul de distance d'édition entre deux chaînes et la programmation dynamique.

A. *Les réseaux de neurones*

Les réseaux de neurones (RN) artificiels sont des réseaux de processeurs élémentaires extrêmement reliés entre eux et fonctionnant en parallèle. Chacun des processeurs effectue son propre calcul et fournit en sortie un résultat à partir des informations qu'il reçoit sur ses entrées. Les réseaux de neurones ont connu un succès fantastique dans les années 90 en particulier en raison de la mise au point d'un algorithme d'apprentissage robuste et pratique dans sa mise en œuvre : la rétro-propagation du gradient [45].

L'objectif recherche dans un réseau de neurones est de lui enseigner à ce qu'il délivre en sortie les informations souhaitées pour un ensemble de données en entrée. Pour ce faire, l'on s'appuie sur un nombre très élevé d'exemples qui lui permettent d'ajuster au fur et à mesure les paramètres - les poids des connexions - de sorte à récupérer en sortie les valeurs voulues en fonction des valeurs des entrées. De nombreuses topologies de réseaux de neurones sont disponibles :

- Les réseaux multicouches : Organisés en couches, chaque neurone reçoit habituellement sur son entrée les sorties de tous les neurones de la couche inférieure. Dans ce type de réseau, il n'y a ni cycles ni connexions intra-couche. Aussi, détermine-t-on une "couche d'entrée", "une couche de sortie" et n "couches cachées". Les réseaux multicouches sont souvent utilisés parce que l'apprentissage est facile à réaliser grâce à l'algorithme de rétropropagation du gradient
- Les réseaux à connexions locales : Dans ce type de réseau, la même structure en couche utilisée précédemment est conservée. Néanmoins, le nombre de connexions entre les couches est limité : un neurone n'est pas nécessairement branché à tous les neurones de la couche antérieure.

Au niveau d'un neurone d'une couche cachée, on dispose d'une fonction qui effectue la somme des valeurs des entrées pondérées par leur poids respectif. Cette somme se transforme ensuite en une sortie en une fonction d'activation.

B. *Machines à Vecteurs de Support*

Les Machines à Vecteurs de Supports (SVM) [45], sont une famille de méthodes à noyaux permettant de trouver efficacement des solutions à des problèmes de classification ou de régression. En fait, ces méthodes, mises au point par *Vladimir Vapnik*, s'inspirent de la théorie

de l'apprentissage statistique. Cette méthode a rencontré dernièrement énormément de réussite. Elle utilise des outils efficaces, comme ceux de la librairie LIBSVN [46], qui s'adaptent aisément à des tests sur un système à base de SVM et qui ont de cette manière contribué encore plus à l'essor de cette méthode.

En résumé les SVM, qui sont des classifieurs à 2 classes, ont pour but de rendre maximale la marge entre les deux classes. Son principe consiste à recourir à une transformation non linéaire sur des données afin d'assurer leur "immersion" dans un espace de plus grandes dimensions. En effet, des données qui seraient non linéairement séparables dans l'espace d'origine seront plus faciles à classer dans un espace de plus grandes dimensions (dans l'idéal, ces données deviennent linéairement séparables, ce qui permet ainsi de définir un hyperplan de séparation). L'idée de base consiste à estimer alors la frontière de séparation linéaire qui maximiserait la marge dans l'espace de plus grandes dimensions (voir figure I.8). En d'autres mots, l'hyperplan de séparation linéaire dans l'espace de plus grandes dimensions, sert d'une certaine façon à établir une frontière de séparation non linéaire dans l'espace de départ.

Pour le calcul de la frontière de séparation, on se base sur les vecteurs de support. Parmi ces vecteurs, seul, un sous-ensemble de données d'entraînement remplirait le rôle de vecteurs de supports. C'est ce sous-ensemble qu'on utilise pour effectuer la classification.

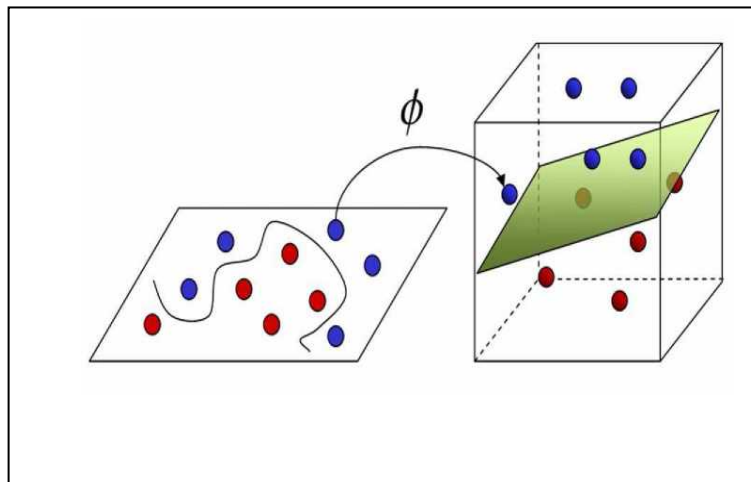


Figure I.8 Principe du SVM : Trouver une frontière de séparation linéaire dans un espace de grandes dimensions revient à trouver une frontière non-linéaire dans l'espace de départ.

Soit l'ensemble d'entraînement à l échantillons :

$$\{(x_i, y_i)\}, \text{ Avec } i = 1, \dots, l, \text{ et } y_i \in \{+1; -1\}$$

La décision s'obtient en évaluant le signe de :

$$f(x) = \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b \quad (\text{I.8})$$

Les vecteurs de support sont les vecteurs x_i pour lesquels les α_i sont non-nuls, et :

$$K(x, x_i) = \langle \varphi(x), \varphi(x_i) \rangle \quad (\text{I.9})$$

$K(x, x_i)$ est une fonction noyau [64]. L'utilisation d'une fonction noyau appliquée aux données dans l'espace initial en lieu et place du calcul du produit scalaire dans l'espace de grandes dimensions est appelée "kernel trick". Cette astuce permet de substituer un produit scalaire dans un espace de grandes dimensions par une fonction noyau, plus rapide à calculer. De cette manière, un classifieur linéaire peut aisément être transformé en un classifieur non linéaire. Autre avantage des fonctions noyau est qu'il n'est pas indispensable d'expliquer la transformation. Seul, le résultat du produit scalaire compte. Les noyaux le plus souvent utilisés sont :

- Le noyau sigmoïde:

$$K(x_i, x_j) = \tanh(K(x_i, x_j) + \theta), \quad (\text{I.10})$$

- Le noyau linéaire:

$$K(x_i, x_j) = x_i x_j, \quad (\text{I.11})$$

- Le noyau Polynomial:

$$K(x_i, x_j) = (x_i x_j + 1)^p, \quad (\text{I.12})$$

- Le noyau Radial Basis Function:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad (\text{I.13})$$

K, θ, p, σ^2 sont des paramètres des noyaux, dont on détermine l'optimum sur un problème donné à l'aide d'une grille.

Les SVM sont intrinsèquement un système de classification à deux classes. *J. Milgram* [125] a étudié les différentes stratégies de combinaison afin de traiter le cas des multi-classes. Il a aussi proposé une méthode afin d'estimer les probabilités de classification à l'aide des SVM en vue de les intégrer dans un environnement probabiliste.

I.4.5.3 Les approches stochastiques

Les approches stochastiques font appel à un modèle, pour la reconnaissance, qui prend en compte l'étendue de la variabilité de la forme. La forme est ici considérée comme un signal continu observable dans le temps à différents endroits qui constituent les états "d'observation". Le modèle décrit ces états en ayant recours aux probabilités des transitions d'états et des probabilités d'observation par état. S'ensuit ensuite une étape de comparaison dont la mission se résume à rechercher, dans ce graphe d'états, le chemin de probabilité le plus fort qui correspond à une suite des éléments observés dans la séquence d'entrée. Ces techniques sont performantes et sûres en raison de l'existence d'algorithmes d'apprentissage

efficaces. Il est vrai que leur apprentissage est quelque peu lent, toutefois la reconnaissance est par contre extrêmement rapide parce que les modèles contiennent en général très peu d'états et le temps de calcul se trouve relativement accéléré.

A. *Les modèles de Markov cachés*

En raison de leur capacité d'intégration du contexte et de leur absorption du bruit, les modèles de Markov cachés (Hidden Markov Models) sont considérés comme étant une technique qui a donné un élan fabuleux à la reconnaissance de l'écriture manuscrite arabe [35].

Le modèle de Markov caché est un processus doublement stochastique. Il contient un processus sous-jacent non observable pouvant être retrouvé à partir d'un deuxième processus stochastique qui fournit les séquences d'observation. Dans les approches de reconnaissance de l'écriture par le modèle HMM, on s'ingénie à modéliser la séquence de caractères à l'aide de modèles de Markov cachés. La couche cachée du modèle se distingue par la séquence d'étiquettes constituant la séquence. Par ailleurs, la couche observable est donnée par une séquence d'observations que l'extraction de caractéristiques fournira via à un balayage séquentiel de l'image ou du signal temporel.

I.4.6 La décision

Les méthodes de reconnaissance rencontrées ci-dessus sont employées afin d'arrêter les règles de prise de décision. Plusieurs types de décisions peuvent être pris par le système :

- **Caractère ou mot reconnu** : Si le système de reconnaissance réussit à associer un prototype, et un seul, au caractère (ou au mot) à reconnaître, il prendra alors une décision unique.
- **Caractère ou mot ambigu** : Si le système adjoint au caractère (ou bien au mot) à reconnaître plusieurs prototypes, il présentera dans ce cas un ensemble de choix ayant des confiances du même ordre.
- **Caractère ou mot rejeté** : Si le système ne réussit guère à relier un prototype de sa base de données au caractère (ou au mot) à reconnaître, aucune décision de classification ne sera prise.

On définit alors un ensemble de facteurs de performance du système tels que :

- *le taux de reconnaissance* : Il traduit le pourcentage de caractères ou de mots ayant été reconnus parmi ceux qui ont été présentés à l'entrée du système.
- *le taux de substitution* : Il traduit le pourcentage de caractères ou de mots ayant été acceptés par le système mais rangés de façon incorrecte

- *le taux de rejet* : Il traduit le pourcentage de caractères ou de mots ayant été rejetés (parce que non reconnus) parmi ceux qui ont été présentés à l'entrée du système.

I.4.6.1 Evaluation des performances d'un classifieur

Ce paragraphe aborde les critères le plus souvent utilisés afin d'estimer la robustesse des systèmes de reconnaissance dans leur phase de généralisation. Le but étant d'acquérir une évaluation fidèle autant que possible du comportement du système dans des conditions réelles d'utilisation. Pour ce faire, des tests habituels tels que le taux de reconnaissance et le taux d'erreurs sont presque régulièrement utilisés [23]. Toutefois, d'autres critères à l'image de la fiabilité ou de la complexité véhiculent également des informations significatives.

A. Taux de reconnaissance et taux d'erreurs

Les taux de reconnaissance et d'erreurs servent à apprécier la qualité du classifieur f_M par rapport au problème pour lequel il a été préconçu. Ces taux sont évalués en se référant à une base de tests qui renferme des formes d'écriture dans le même espace de représentation E que celles utilisées pour l'apprentissage. Elles sont également marquées suivant leur classe réelle d'appartenance dans le but de juger les réponses du classifieur. Afin que l'évaluation du taux de reconnaissance soit la plus vraisemblable possible, il est important que le reconnaiseur n'est jamais eu recours aux échantillons de cette base pour effectuer son apprentissage (la base de tests ne doit avoir aucun spécimen commun avec ceux de la base d'apprentissage ou éventuellement ceux des autres bases de validation). En outre, cette base de tests doit être raisonnablement représentative du problème de classification.

De manière générale, lorsque les échantillons étiquetés et mis à disposition sont relativement abondants, ils sont alors scindés en deux groupes disjoints en prenant soin de respecter les proportions par classes de la base initiale. Le premier groupe est utilisé pour établir la base d'apprentissage tandis que le second aide à constituer la base de test. Le découpage le plus couramment utilisé est les deux-tiers pour la base d'apprentissage et le tiers restant pour la base de test. Les performances en termes de taux de reconnaissance sont alors calculées en présentant au classifieur chacun des exemples e_j de la base de test et en comparant la classe donnée en résultat $f_M(e_j) = s$ à la vraie classe de e_j .

En supposant que la base de tests contienne N individus, et que parmi ceux-ci $N_{corrects}$ sont biens classés par le système, le taux de reconnaissance τ_{reco} s'exprime alors naturellement par

$$\tau_{reco} = \frac{N_{corrects} \times 100}{N} \quad (I.14)$$

Le taux d'erreur τ_{err} est alors de la même manière défini à partir du nombre d'individus N_{err} mal classés.

$$\tau_{err} = \frac{N_{err} \times 100}{N} \quad (I.15)$$

I.5 Domaines d'applications de la reconnaissance de l'écriture

La reconnaissance automatique hors ligne de l'écriture suscite un intérêt incontestable dans la mise au point de multiples applications telles que celles que l'on rencontre dans plusieurs segments d'activités. Loin d'être exhaustif, citons certaines de ces applications :

- Authentification et lecture de chèques bancaires au moyen de la reconnaissance des montants rédigés en toutes lettres ou bien encore ceux qui sont écrits en chiffres.
- Lecture automatique des formulaires et des documents administratifs.
- Lecture des adresses postales et tri automatique du courrier.
- Echange de fichiers informatisés à distance dans le domaine des télécommunications.
- Indexation et archivage automatique de documents.
- Reconnaissance automatique des documents techniques (schémas électroniques, dessins techniques, plans architecturaux, plans cartographiques, etc.).
- Transcription assistée par ordinateur appliqué à la reconnaissance du manuscrit mono-scripteur ce qui permettra la rétro-conversion en documents électroniques de manuscrits anciens.
- Reconnaissance des numéros des plaques minéralogiques pour le contrôle routier.
- Authentification et identification de manuscrits et identification du scripteur.
- Recherche d'informations dans une base de documents manuscrits en vue d'une identification d'un scripteur.

I.6 Conclusion

Ce chapitre a servi à brosser, de manière rapide, le tableau de notre domaine d'intérêt : la reconnaissance de l'écriture manuscrite. Nous avons d'abord présenté dans la première partie un certain nombre de travaux de recherche les plus remarquables dans ce domaine de la reconnaissance automatique de l'écriture. Ensuite, nous avons énuméré les différentes méthodes utilisées, ainsi que les problèmes les plus saillants auxquels fait-on face lors de la construction de tels systèmes. Les différentes phases qui constituent un système de reconnaissance (acquisition, prétraitement, extraction des primitives et classification) ont été abordées plus loin et explicitées. Finalement, nous avons rappelé des applications réelles de ce domaine constatées dans notre vie quotidienne.

Chapitre **2**

*La Reconnaissance
de L'écriture
Arabe Manuscrite*

II.1 Introduction

La reconnaissance de l'écriture arabe fait naturellement partie d'un cadre beaucoup plus général de reconnaissance de l'écriture cursive avec tout ce que ce type d'écriture implique comme difficultés et spécificités intrinsèques. Seulement, à la différence de la plupart des autres scripts, la reconnaissance de l'écriture arabe manuscrite reste encore jusqu'à présent un défi scientifique majeur qui sévit dans toute sa complexité et qui tarde encore à offrir une solution satisfaisante. Les raisons de ces avancées lentes dans cette catégorie de scripts sont évidentes et sont singulièrement dues à la complexité morphologique de l'écriture arabe et aux grandes variétés de styles d'écriture qui posent de multiples problèmes aux systèmes de reconnaissance automatique. Ce chapitre tient comme objet de faire une analyse des différentes méthodes et outils employées dans la reconnaissance de l'écriture arabe manuscrite et de considérer ensuite les différentes complications auxquelles les chercheurs ont dû faire face avant de concevoir et proposer en fin de parcours des systèmes de reconnaissance robustes.

II.2 Présentation et problématique de l'écriture arabe

L'écriture arabe dérive de l'écriture nabatéenne. Elle est l'ascendante directe de l'écriture araméenne ancienne qui, elle-même, est une variété de l'alphabet phénicien. A partir de là, et en subissant autant de dérivations que d'altérations dans la forme, allait se développer l'alphabet arabe qu'on connaît, qui subsistera longtemps, et presque invariablement jusqu'à aujourd'hui, proche de son ancêtre. En résumé, l'écriture arabe a donc vu le jour aux environs du septième siècle émanant de l'écriture cursive nabatéenne. Elle possède plusieurs particularités que l'on peut présenter ci-dessous :

1. L'alphabet arabe comporte 28 lettres dont les formes changent selon que l'on est au début d'un mot, à son milieu, à sa fin, ou bien que cette lettre soit isolée (voir tableau II.I). Parmi ces 28 lettres, on en distingue 22 qui présentent chacune quatre formes d'écriture différentes.
2. Les six lettres restantes : "ذ", "د", "ر", "ز", "أ", "و" et "و" ne peuvent être jamais reliées aux lettres qui leur succèdent. Aussi, n'ont-elles que deux formes possibles : elles peuvent être ou bien isolées au début d'un mot ou bien isolées à la fin. A noter que de par leur nature, ces lettres conduisent à une coupure dans le mot. De sorte que dans l'écriture arabe, un mot est composée d'une séquence d'entités connexes entièrement séparées appelées pseudo-mots (ou PAWs : *Pieces of Arabic Words*) qui peuvent contenir 2 ou 3 lettres, voire plus, qui n'ont pas obligatoirement un sens intelligible. Pour

résumer, un mot se compose d'un ou de plusieurs pseudo-mots, chaque pseudo-mot est une suite de lettres accolées les unes aux autres, ce qui confère à cette écriture son aspect cursif. Précisons enfin qu'un caractère isolé peut, lui-même, constituer un pseudo-mot à lui tout seul.

La **figure II.1** illustre parfaitement un exemple dans lequel le premier mot (en partant de la droite vers la gauche), le deuxième mot, le troisième mot ainsi que le quatrième sont respectivement constitués de 2, 4 et 3 pseudo-mots. Rappelons qu'un pseudo-mot est une unité connexe qui regroupe une ou plusieurs lettres sous forme d'une séquence. Dans l'écriture manuscrite spécifiquement, l'espacement entre les différents pseudo-mots d'un même mot n'est pas forcément systématiquement supérieur à l'espacement entre deux mots différents. Ceci ajoute une difficulté supplémentaire pendant l'étape de segmentation.

3. L'écriture arabe s'écrit de la droite vers la gauche et ce, contrairement aux écritures occidentales.
4. Aucune différence entre les lettres manuscrites et les lettres imprimées. Les concepts de lettre capitale et lettre minuscule n'existent pas dans l'écriture arabe. Elle est donc monocamérale.
5. La plupart des lettres s'attachent entre elles, aussi bien même en écriture imprimée qu'en écriture manuscrite. Ce qui confère à l'écriture arabe son caractère cursif [46]. La **figure II.2** illustre un exemple emprunté à une sourate du Coran, le Livre sacré.
6. Les caractères arabes ne possèdent pas de taille invariable (aussi bien en hauteur qu'en largeur). Un caractère change dans la taille d'une lettre à une autre. Sa forme change également d'une lettre à une autre à l'intérieur du même caractère.



Figure II.1 Exemple de pseudo-mots constituant les mots arabes

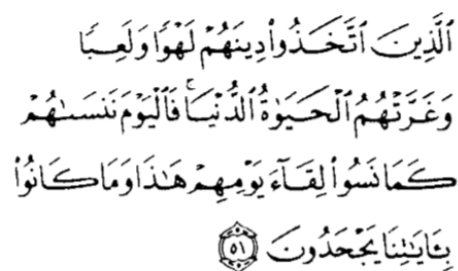


Figure II.2 Cursivité de l'écriture arabe

Tab II.1: les caractères arabes et leurs différentes formes, (b) les caractères additionnels « ة » et « لا », (c) et (d) Hamza et Madda et les positions qu'elles occupent en association avec Alif, Waw et Ya.

caractère	Initiale	milieu	Finale	Isolé
Alif			ا	ا
Ba	ب	ب	ب	ب
Ta	ت	ت	ت	ت
Tha	ث	ث	ث	ث
Nun	ن	ن	ن	ن
Ya	ي	ي	ي	ي
Jim	ج	ج	ج	ج
Ha	ح	ح	ح	ح
Kha	خ	خ	خ	خ
Dal			د	د
Thé			ذ	ذ
Ra			ر	ر
Za			ز	ز
Waw			و	و
Sin	س	س	س	س
Chin	ش	ش	ش	ش
Sad	ص	ص	ص	ص
Dhad	ض	ض	ض	ض
Tad	ط	ط	ط	ط
Dhad	ظ	ظ	ظ	ظ
Ayn	ع	ع	ع	ع
Ghyn	غ	غ	غ	غ
Fa	ف	ف	ف	ف
Qaf	ق	ق	ق	ق
Kaf	ك	ك	ك	ك
Lam	ل	ل	ل	ل
Mim	م	م	م	م
Ha	ه	ه	ه	ه

(a)

Caractère	Initiale	Milieu	Finale	Isolé
Alif + ~			آ	آ
Alif + ء			إ	إ
			أ	أ
Waw + ء			ؤ	ؤ
Ya+ ء		ئ	ئ	ئ

(b)

Caractère	Initiale	Milieu	Finale	Isolé
Ta marbouta				ة
Lamalif				لا

(c)

Caractère	Initiale	Milieu	Finale	Isolé
Lamalif			لا	لا
Lamalif +~			لا	لا
Lamalif +ء			لا	لا
			لا	لا

(d)

La reconnaissance de manuscrit arabe est beaucoup plus ardue que la reconnaissance des manuscrits d'autres écrits tels que le latin [47] parce que cette écriture possède des caractéristiques spécifiques qui la rendent plus compliquée. Parmi ces caractéristiques, citons encore :

- Les signes diacritiques ont un rôle décisif dans la distinction entre certains groupes de lettres. En outre, leur manipulation est loin d'être aisée. En effet, les étapes de prétraitements, telles que la squelettisation et la suppression du bruit en particulier ont un effet négatif sur la forme des signes diacritiques qui se retrouvent altérés voire même totalement supprimés. D'un autre côté, la variation dans les styles d'écriture des diacritiques accroît la complexité de reconnaissance de ces signes. Par ailleurs, la superposition verticale de ces diacritiques par rapport à la lettre originale, qui est pratiquement garantie dans les écrits imprimés, ne l'est point dans les manuscrits (voir [figure II.3.a](#)).

A l'instar de l'écriture latine, l'écriture arabe possède des descendants. Ces derniers, en se prolongeant horizontalement sous la bande de base, risquent d'introduire une superposition verticale entre le descendant et la lettre suivante ce qui rend, non seulement la segmentation encore plus complexe, mais également l'analyse de l'histogramme de projection horizontale.

- L'existence du pseudo-mot conduit lui-même à une segmentation spontanée de l'écriture. Pourtant dans le manuscrit, l'espace qui apparaît alors entre les mots (espace intra-mot) n'est pas obligatoirement plus petit que l'espace naturel qui sépare deux mots (espace entre-mots). Cet état de fait rajoute encore une complication de plus dans la segmentation en pseudo-mots et en mots. En plus, cette décomposition en pseudo-mots peut elle-même laisser apparaître plusieurs angles d'inclinaisons à l'intérieur d'un même mot. Ceci posera également problème durant l'extraction de la ligne de base.

- Ligature verticale: les connexions entre les lettres arabes d'un pseudo-mot se situent habituellement au niveau de la ligne de base. Pourtant, un certain nombre de lettres peuvent être reliées verticalement ce qui conduit ainsi à des ligatures verticales, en général, très difficiles à segmenter

- Levée de plume : Deux familles de lettres peuvent engendrer l'obligation de la levée du stylo avant de continuer l'écriture. La première famille comprend les lettres qui s'écrivent de gauche à droite telles que : « ط » « ح » et « ض ». La seconde concerne les lettres qui s'écrivent de haut en bas telles que la lettre « ك ». L'apparition de l'une de ces lettres au milieu ou à la fin des pseudo-mots oblige le scripteur à lever son stylo provoquant parfois des coupures embarrassantes à l'intérieur des pseudo-mots qui doivent être traités pendant la reconnaissance. En effet, cet état de fait crée donc des composantes connexes additionnelles et conduit fatalement au problème de sur-segmentation des pseudo-mots (voir [figure II.4.c](#)).

- Par contre, les connexions non souhaitées entre les pseudo-mots, introduites par le type d'écriture du scripteur, sont indubitablement très fréquemment rencontrées dans le manuscrit arabe. Ce genre de problème apparaît habituellement entre les successions des descendants

qui peuvent se joindre ainsi qu'au niveau des ligatures verticales. Cet aspect de l'écriture rend malaisée l'extraction des pseudo-mots. En d'autres mots, cela conduit inévitablement à des sous-segmentations en pseudo-mots et en mots (voir [figure II.4.b](#)).

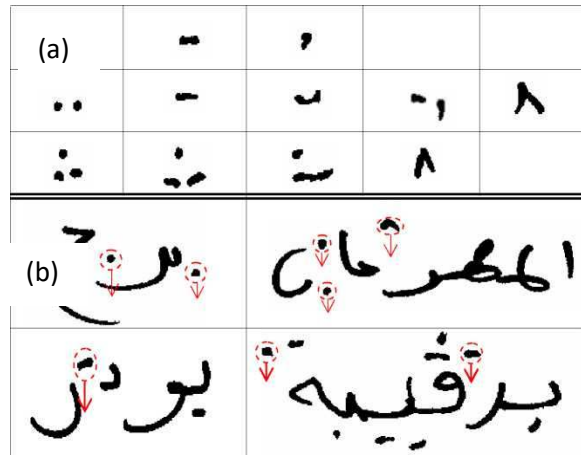


Figure. II.3 - Difficultés liées aux points diacritiques. (a) différents styles d'écriture des points diacritiques ; (b) problème d'association des diacritiques à leur lettre originale.

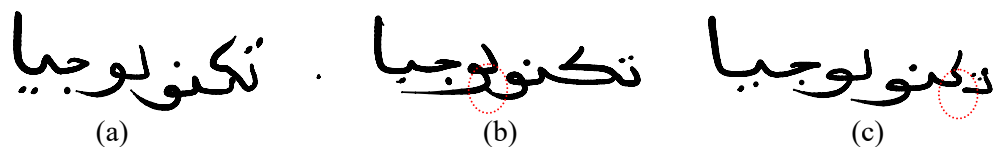


Figure. II.4 Les problèmes liés à l'extraction des pseudo-mots. (a) un mot arabe correctement écrit (3 PAW) ; (b) liaison indésirable entre deux caractères et sous segmentation en PAWs (2 PAWs) ; (c) coupure indésirable et sur-segmentation en PAWs (4 PAWs).

L'introduction des élongations horizontales, autre caractéristique couramment utilisée dans le manuscrit arabe, correspond à l'insertion d'une ou plusieurs élongations entre les caractères d'une même chaîne. Ces élongations se rencontrent le plus souvent à gauche du caractère courant ou encore à gauche des ligatures verticales. La [figure II.5](#) montre un mot arabe contenant deux ligatures horizontales des deux caractères "ء" et "ج".

Tab II.2 Différents hambages et jambages dans les lettres arabes

	Hambages					
Lettres	ا	ل	ك	و	و	ط
	Jambages					
Lettres	و	و	م	و	و	و
و	و	و	ح	ع	ع	ع

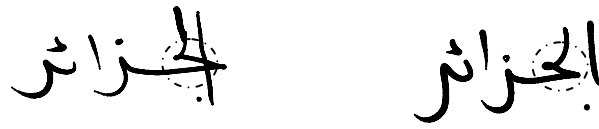


Figure II.5 : Ligatures verticales et inversion de l'ordre du tracé.

Tout comme dans l'écriture latine, l'écriture arabe est pourvue d'ascendants et de descendants. En arabe, les descendants ou les jambages peuvent continuer horizontalement sous la bande de base, ce qui conduit à une superposition verticale entre la lettre qui comprend le descendant et la lettre suivante. Le **tableau II.2** fournit toutes les lettres arabes qui possèdent des hampes et des jambages.

- L'alphabet arabe dispose de 15 lettres parmi les 28 possédant un ou plusieurs points. Ces points diacritiques sont placés soit au dessus soit en dessous de la lettre à laquelle ils sont associés. A noter qu'il n'y a pas de lettres en arabe disposant de points des deux côtés de la forme. Le nombre de points que l'on peut placer au dessus d'une lettre est soit 1, 2 ou 3 points. De même, le nombre de points que l'on peut placer en dessous du caractère est toujours 1 ou 2 points. Ces points permettent de distinguer la prononciation des lettres arabes. Le **tableau II.3** présente toutes les lettres qui possèdent des points diacritiques ainsi que leur nombre et leur position.

Tab II.3 Lettres arabes ayant des points diacritiques.

Position	Au dessus	Au dessous
Nombre de points		
Un point diacritique	خ ز ض ظ غ ف ن	ب ج
Deux points diacritiques	ت ق	ي
Trois points diacritiques	ث ش	

II.3 Principales bases de données existantes

L'existence de bases de données universelles dans le domaine de la reconnaissance de l'écriture manuscrite a joué un rôle primordial dans l'évaluation et la comparaison des résultats auxquels sont arrivés les chercheurs de par le monde. Malheureusement, la reconnaissance automatique de l'écriture arabe manuscrite a été le parent pauvre parmi les différents systèmes de reconnaissance et a donc été caractérisée, durant de longues années, par l'absence d'une base de données de référence qui aurait permis aux différents chercheurs de comparer objectivement les résultats de leurs activités. Nous allons présenter dans ce qui suit les bases de données les plus réputées.

II.3.1 CENPARMI 1 Database

Les chercheurs du Centre for Pattern Recognition and Machine Intelligence (CENPARM) [48] ont parachevé, durant l'année 2000, la création d'une base de données pour la reconnaissance des chèques arabes manuscrits, scannés en niveaux de gris avec une résolution de 300 dpi. Une collection de 3000 images de chèques a permis à ces chercheurs d'élaborer quatre bases de données : Une base de 2 499 montants littéraux, une base de 2 499 montants numériques écrits avec les chiffres indiens, une base de 29 498 pseudo-mots et une base de 15 175 chiffres indiens isolés.

II.3.2 IFN/ENIT Database

La base de données IFN/ENIT [49] comprend des noms manuscrits arabes des villes et villages tunisiens collectés auprès de 411 scripteurs. Les images sont binaires et sont réalisées avec une résolution de 300 ppp. Cette base d'images a été mise au point pour l'apprentissage et le test des systèmes de reconnaissance de l'écriture arabe manuscrite. Les données de la base IFN/ENIT ont été aussi utilisées pour la reconnaissance des différents scripteurs. Cette base de données, de par sa taille et sa disponibilité (elle est en effet à accès libre pour les travaux académiques) s'est imposée comme la base de référence des travaux en reconnaissance de scripteurs arabes et en reconnaissance d'écriture arabe manuscrite.

II.3.3 AHDB database

La base de données AHDB (*Almaadeed et al.*, 2002)[36] dispose de mots arabes et de textes rédigés par 100 scripteurs. En premier lieu, les auteurs ont sélectionné, à partir de documents récupérés de la toile d'internet, 20 mots qui y sont les plus récurrents. Lors de l'étape suivante, ces chercheurs ont établi un formulaire en six pages. Les trois premières pages de ce formulaire contiennent 96 mots, 67 d'entre eux sont des mots manuscrits qui sont les plus employés dans les écritures manuscrites. Les 29 autres mots restants ont été sélectionnés à partir des mots les plus courants en écriture arabe. La quatrième page comprend trois phrases composées de mots manuscrits, de nombres et de quantités ayant le plus de chance d'être présents sur les chèques. La cinquième page est alignée, elle est prévue pour être renseignée au gré du scripteur sur tout sujet de son choix. 105 formulaires ont été scannés à l'aide d'un scanner à 600 dpi en format TIF.

II.3.4 CEDARABIC Database

La base de données CEDARABIC [50] a vu la participation de 10 scripteurs. Chacun d'entre eux a rédigé 10 pages entières et différentes de texte manuscrit. Chaque page contient entre 150 et 200 mots, soit un total de 100 pages contenant en tout quelque 20 000 mots. Les documents ont été ensuite scannés en 300 dpi. Pour compléter cela, une annotation faite à la

main a complété ces documents : segmentation en mots, mais également séquence de lettres arabes, ainsi que la prononciation et la traduction en anglais. Cette base de données est donc spécialement adaptée pour des travaux sur la localisation de mots (word spotting) [50] ainsi que la reconnaissance de l'écriture [51].

II.3.5 IFN/Farsi database

Cette base de données a été publiée par Mozaffari et al. en 2008 [52]. Elle contient des images de 1080 différents noms de villes et de provinces iraniennes. Elle a été réalisée par 600 scripteurs environ. Chacun d'entre eux a été prié de renseigner deux formulaires distincts avec les noms de 24 villes présélectionnés ainsi que les codes postaux correspondants. La base IFN/Farsi contient un total de 7271 mots, 23545 pseudo-mots et 43501 caractères. Elle a été automatiquement annotée au niveau caractère. L'annotation comporte : la séquence/nombre de caractères, le nombre de pseudo-mots ainsi que le nombre/position des diacritiques. Cette annotation est différente de celle contenue dans la base IFN/ENIT, en ce sens qu'elle ne contient pas la position de la ligne de base. Nonobstant, elle y recèle une information très utile pour les systèmes de reconnaissance : la position (U : up et D : down) et le nombre de signes diacritiques. Cette information peut conduire à une réduction lexicale [54].

II.3.6 KHATT Database

KHATT [55] est une nouvelle base de données qui contient des images de textes arabes manuscrits. Elle est employée pour la reconnaissance des scripteurs, la segmentation des textes en lignes ainsi que la reconnaissance de textes manuscrits. Elle est devenue publique en septembre 2012. La base de données KHATT dispose de 4000 images de paragraphes en niveaux de gris. Ces images représentent des textes scannés à diverses résolutions (200, 300 et 600 ppp). 1000 scripteurs, de différents âges et origines et appartenant à 18 pays différents, ont pris part dans la construction de cette base de données. Parmi les 1000 scripteurs, 677 étaient de sexe masculin et 323 étaient de sexe féminin. 928 scripteurs étaient droitiers alors que 72 autres étaient gauchers. 2000 images, parmi les 4000 de la base, se réfèrent au même texte qui couvre la totalité des caractères et chiffres arabes. Les 2000 images restantes se rapportent à des textes libres écrits par les scripteurs sur des sujets qui les intéressent individuellement.

II.3.7 AHTID/MW Database 2012

La base de données AHTID/MW (Arabic Handwritten Text Images Database written by Multiple Writers) [56] est une collection de documents manuscrits rassemblés par les chercheurs du Laboratoire MIRACL, ISIMS, Université de Sfax, Tunisie en collaboration

avec les chercheurs de l'Institut de la Technologie de Communications (IFN), Braunschweig, Allemagne. Cette base de données est accessible pour les besoins de la recherche dans le domaine de la reconnaissance de l'écriture arabe manuscrite, la segmentation des mots et l'identification des scripteurs. La base AHTID/MW contient en tout 22896 mots repartis sur 3710 lignes de texte rédigés par 53 scripteurs différents avec une précision de taille : la langue maternelle de ces scripteurs est l'arabe. Ces images ont été scannées en niveau de gris avec une résolution de 300 dpi en format bmp. Cette base est subdivisée en quatre parties distinctes. Chaque image de mot/ligne dans la base de données AHTID/MW est totalement décrite au moyen d'un fichier XML qui contient des informations sur la séquence des PAW (Party of Arabie Word), ainsi que la séquence des caractères pour chaque PAW.

II.3.8 QUWI Database

QUWI [57] est une base de documents manuscrits hors ligne récente. Elle a été réalisée par une équipe de chercheurs de l'Université du Qatar. Cette base de données renferme des documents rédigés en arabe et en anglais. Il est possible de l'utiliser pour évaluer les performances des systèmes de reconnaissance. Elle est composée de manuscrits écrits par 1017 scripteurs de d'âges, nationalités, sexes et niveaux d'éducation tous différents. Les scripteurs ont été priés à copier un texte particulier et à générer un texte aléatoire. Ceci a permis à la base de données d'être sollicitée aussi bien en mode dépendant du texte qu'en mode indépendant du texte. Il est à noter qu'une partie de cette base de données (475 scripteurs) a fait l'objet d'un concours sur la prédiction du genre à partir de textes manuscrits qui s'est déroulée dans le cadre de la conférence ICDAR 2013.

II.4 Surveys existants

Plusieurs *surveys* sont proposés dans le domaine de la reconnaissance de l'écriture arabe manuscrite dans le but de présenter l'état de l'art ainsi que les différentes techniques développées.

II.4.1 B. Al-Badr et S. A. Mahmoud, 1995

Cette étude a résumé les travaux de recherche sur la reconnaissance du texte arabe depuis la toute première publication [58]. Les auteurs ont ainsi analysé et classifié les résultats de cette recherche en cinq catégories en fonction des cinq phases qui constituent le processus de reconnaissance : prétraitement, segmentation, extraction de primitives, classification et post-traitement.

Conformément aux conclusions des auteurs de cette étude, la différence fondamentale entre la reconnaissance de l'écriture latine et l'écriture arabe réside dans la phase de segmentation.

Ces auteurs recommandent que des recherches poussées soient entreprises et doivent être orientées dans le développement d'algorithmes de segmentation plus efficaces. Ils conseillent par ailleurs de profiter des avantages de la combinaison des différents types de primitive avec les différents classifieurs dans le but d'accroître le taux de reconnaissance. En revanche, ils regrettent l'absence de la phase de post-traitement dans les procédures de reconnaissance actuelles. Ils préconisent également au recours à une base de test commune pour l'évaluation des résultats comme un préalable important. Aussi, invitent-ils les chercheurs dans ce domaine à exploiter, dans la phase d'apprentissage, des documents altérés dans un premier temps avant de s'attaquer plus tard à la reconnaissance de textes plus complexes tel que le *Coran*.

II.4.2 A. Amin, 1998

Dans [48], *Adnan Amine* passe en revue les différentes techniques de reconnaissance de l'écriture arabe : les méthodes en ligne et hors ligne ainsi que les approches globale et analytique. Il met en lumière les difficultés rencontrées dans la segmentation, qu'elle soit implicite ou explicite, des mots en caractères. Il examine ensuite les différentes approches utilisées pour la reconnaissance ainsi que les problèmes rencontrés. Cette étude comprend aussi une classification des systèmes de reconnaissance conformément au type des primitives utilisées (globales ou locales) et au type du classifieur utilisé (RNA ou HMM). Notons finalement que tous les systèmes étudiés dans ce *survey* ont abordé de façon exclusive la reconnaissance de l'écriture arabe imprimée (mots et caractères).

II.4.3 M. S. Khorsheed, 2002

M. S. Khorsheed fait passer en revue, dans [59], un *survey* des méthodes de prétraitements, de segmentation, d'extraction des caractéristiques et de classification. L'auteur a fait implémenter de multiples techniques de prétraitements et de segmentation et a précisé les résultats obtenus sur des mots imprimés et des textes arabes manuscrits. Il a aussi établi une brève description de quatre OCR de textes imprimés arabes : TextPert, ICRA, OmniPage, et Al-Qari' al-Ali.

II.4.4 N. Essoukri Ben Amara et F. Bouslama, 2003

Les auteurs de cette étude ont démontré les avantages que l'on peut tirer de la combinaison de plusieurs sources d'informations dans un système de reconnaissance d'écriture arabe [60]. Ils proposent, à cet effet, trois niveaux de combinaison :

Le premier niveau concerne la combinaison de plusieurs familles de caractéristiques. Cela est motivé d'un côté, par l'insuffisance d'une seule méthode à englober la variation

intra-classes, et d'un autre côté par la sensibilité de certaines techniques aux bruits et distorsions.

Le deuxième niveau concerne la combinaison de plusieurs classifieurs. En effet, il n'existe pas un classifieur universel aussi bon en termes d'apprentissage/génération qu'en temps de calcul.

Le troisième niveau concerne la combinaison de plusieurs familles de caractéristiques et des classifieurs.

Les auteurs révèlent ensuite un grand nombre de systèmes de reconnaissance qui ont exploité ces différentes configurations de combinaison. Ils arrivent finalement à la conclusion que cette technique de combinaison résout les problèmes des variations de styles dans l'écriture et se montre performante vis-à-vis des bruits et des distorsions. En définitive, ils proposent d'exploiter encore la logique floue et le recours à des post-traitements dans les techniques OCR arabes.

II.4.5 L. M. Lorigo et V. Govindaraju, 2006

Liana M. Lorigo et Venu Govindaraju ont brossé un tableau complet des différentes étapes de la chaîne de reconnaissance de l'écriture arabe [69]. Le travail accompli par ces auteurs [38] est regardé comme étant le *survey* le plus complet à ce jour dans le domaine de la reconnaissance de l'écriture arabe manuscrite. Ils ont révélé la plupart des problèmes typiques en relation avec la reconnaissance de l'écriture arabe. Ils ont aussi fait le tour de quelques bases de données : AHDB, CENPARMI et IFN/ENIT. Par ailleurs, ils ont exposé la chaîne de reconnaissance de l'écriture arabe manuscrite au regard des étapes suivantes :

- Représentation (Squelette, contour, ou pixels) ;
- Extraction des primitives (statistiques et structurelles) ;

Classifieurs (les systèmes à base de règles, Réseaux de Neurones, Modèle de Markov Cachés et Approches hybrides).

II.4.6 M. T. PARVEZ et S.A. Mahmoud, 2013

M.T. Parvez et S.A. Mahmoud ont réalisé un *Survey* remarquable qui fait état des récentes avancées dans le domaine de la reconnaissance de l'écriture arabe [63]. Leur travail débute par un exposé des caractéristiques particulières de l'écriture arabe, suivi par la proposition d'un modèle général de reconnaissance en hors ligne du manuscrit arabe. Ils abordent aussi les bases de données les plus couramment employées dans la reconnaissance de l'écriture arabe manuscrite. Ils présentent aussi les différentes étapes de prétraitements tels que la représentation des textes (squelette et contour), la détection de la ligne de base, la correction

de l'inclinaison et finalement la segmentation des mots en caractères. Plus loin, ces chercheurs ont classifié les différents travaux de recherche sur l'extraction des primitives en deux catégories: statistiques et structurelles. Ils rappellent également les études d'un très grand nombre de chercheurs ayant employé différentes techniques de classification telles que les HMMs, les réseaux de neurones, SVM, KNN et la combinaison de classifieurs. Enfin, ils dévoilent également quelques approches de post-traitements employées souvent dans le cadre de la reconnaissance de l'écriture arabe.

II.4.7 A. LAWGALI , 2015

Cette étude a résumé les travaux de recherche sur la reconnaissance du texte arabe Selon aux conclusions de auteur [61] de cette étude les problèmes de reconnaissance des mots arabes manuscrits montre que la plupart des techniques développées jusqu'à présent sont basées sur la reconnaissance de mots entiers sans segmentation. Cela est dû à la nature cursive et aux ligatures particulières de l'écriture arabe. De plus, la plupart des algorithmes de segmentation actuellement proposés ne résolvent pas le problème du chevauchement des caractères dans l'écriture arabe, ni ne tentent de reconnaître les mots après leur segmentation. Certains chercheurs ont tenté de reconnaître les caractères après le processus de segmentation, mais l'identification des caractères ne fournit toujours pas une solution complète car cela peut entraîner des erreurs de classification de certains caractères dans un mot.

De plus, l'objectif de ces systèmes est d'identifier le mot, pas un groupe de caractères. D'autres chercheurs ont supposé que les caractères étaient déjà segmentés afin d'éviter le processus de segmentation et ils ont proposé des techniques pour les caractères arabes isolés. Cependant, reconnaître les caractères après le processus de segmentation implique plus de défis.

II.5 Processus de reconnaissance d'écritures arabes manuscrits

Les méthodes classiques de prétraitements jadis mis au point pour l'écriture latine et chinoise n'ont pas montré leur efficacité pour l'écriture manuscrite arabe. En effet, les difficultés liées à l'écrit manuscrit arabe sont très nombreux. On peut citer à titre d'exemple l'estimation de la ligne de base qui, à la différence de ce que l'on peut trouver dans l'écriture latine, n'est pas évidente et l'on se trouve très vite confronté au problème des descendants étendus horizontalement sous la bande de base. Rappelons également la complexité de distribution, loin d'être régulière, des pseudo-mots d'un même mot qui peut générer différents angles d'inclinaison au sein de ce même mot. Soulignons aussi la phase critique de squelettisation et d'élimination des bruits pouvant déboucher sur une altération de la forme des points diacritiques.

Par voie de conséquence, une bonne prise en compte des singularités de l'écriture arabe exige le développement d'étapes de prétraitements entièrement voués à ce script. Le cas échéant, une adaptation *ad hoc* des méthodes existantes est plus que nécessaire.

A cet égard, nous abordons dans cette section les méthodes de prétraitement développées spécifiquement pour l'écriture arabe. Nous rappelons également différentes adaptations à certaines méthodes classiques.

II.5.1 Le prétraitement

Le prétraitement est une phase capitale dans tous les systèmes de reconnaissance d'écriture. Son importance est encore plus décisive pour le traitement du manuscrit arabe qui possède ses propres caractéristiques nécessitant la mise au point de techniques de prétraitements spécifiquement orientés vers ce script ou tout du moins initiées par des adaptations adéquates des méthodes existantes. Dans cette section, nous ferons un tour d'horizon des techniques de prétraitement spécialement conçues pour l'écriture arabe et nous parlerons également des adaptations réalisées sur certaines techniques classiques de prétraitement en direction de l'écrit arabe.

II.5.1.1 Squelettisation

La squelettisation, très courante dans le domaine de la reconnaissance de l'écriture, est une méthode qui introduit des modifications dans la représentation dans le but d'extraire les caractéristiques morphologiques et à réduire la masse d'informations à analyser. L'un des travaux pionniers de squelettisation adaptée particulièrement au script arabe est abordé dans [40]. Les auteurs y exposent un algorithme très distinctif : le Clustering Fuzzy Isodata pour la squelettisation des caractères arabes, appelé CBSA (Clustering Based Skeletonization Algorithm).

Dans [41], les auteurs introduisent un algorithme de squelettisation de l'écriture arabe manuscrite qui a recours à la technique du codage à l'aide de différentes couleurs afin de poinçonner, enlever, ou récupérer des pixels dans une image d'un document écrit en arabe. Cette technique permet d'obtenir de manière simple et efficace un squelette de cette image de bonne qualité et ce, en comparaison aux autres algorithmes qui font appel à une morphologie complexe et à des calculs mathématiques très coûteux en temps de calcul. Selon ces auteurs, cet algorithme assure de manière excellente l'intégrité de la forme de l'image d'origine. Le squelette ainsi obtenu peut être incorporé de manière efficace dans un système de reconnaissance du manuscrit arabe. Certains chercheurs sont arrivés à extraire un squelette au moyen d'une version modifiée de l'algorithme de *Hilditch* [17] utilisée pour la squelettisation

d'une écriture latine.

II.5.1.2 Correction de l'inclinaison des lignes

L'inclinaison des lignes dans un texte arabe est globalement homogène si le document est incliné dans son ensemble. Cette hypothèse est entièrement observée dans les textes imprimés. Dans ces conditions, une correction de l'inclinaison globale du document aboutit à des résultats convenables. Pour ce faire, les approches les plus souvent employées font appel à la transformée de Hough [64] ainsi qu'aux histogrammes de projection. En revanche, dans le cas des textes manuscrits libres, plusieurs types angles d'inclinaison risquent de coexister sur la même ligne. Afin de pallier à cette difficulté, dans [65], les auteurs ont utilisé la chaîne de Freeman pour détecter les traces quasi-verticales. Pour calculer l'angle d'inclinaison moyenne du mot, tous les contours sont parcourus pour calculer l'angle d'inclinaison pour chaque trace quasi-vertical. Ensuite, la moyenne pondérée (des angles d'inclinaisons multipliées par la hauteur euclidienne de la ligne verticale) des angles d'inclinaison individuels, est prise et la moyenne globale d'angle d'inclinaison du mot est déterminée.

Les documents analysés sont générés automatiquement à partir d'une sélection aléatoire de mots de la base IFN/ENIT (Figure II.6.a). L'inclinaison du mot est corrigée en faisant appliquer à l'image une rotation jusqu'à ce que la ligne de base devienne horizontale (figure II.6.b). Au final, un regroupement des mots du texte, dont les lignes sont plus ou moins horizontales, est opéré.

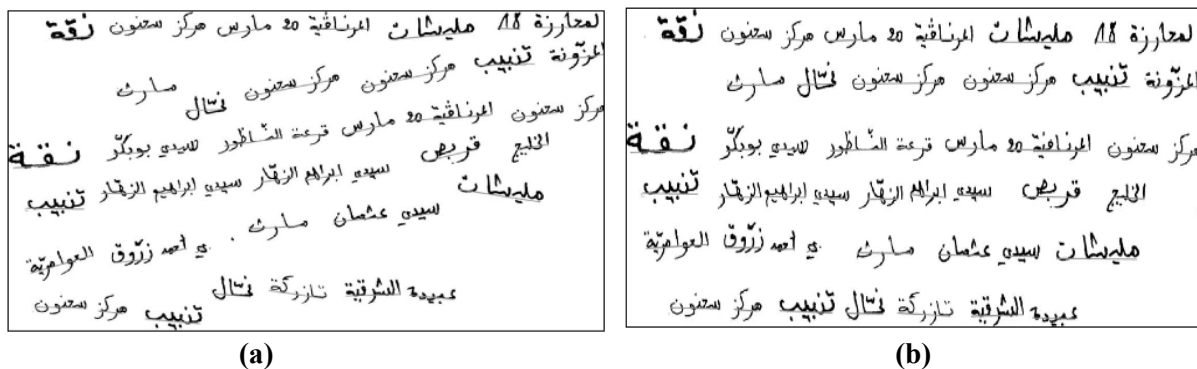


Figure. II.6 Correction de l'inclinaison des lignes d'un document construit automatiquement par collection des images de mots de la base IFN/ENIT (figure extraite de [65]).

II.5.1.3 Estimation de la ligne de base

La détection de la ligne de base est une approche souvent employée dans la reconnaissance de manuscrit arabe. En effet, cette ligne y recèle des données fondamentales sur l'orientation de l'écriture et le positionnement des points de connexion entre les lettres. Par voie de conséquence, son exploration est utilisée intensivement pour la normalisation de la taille et de

et la correction de l'inclinaison de l'écriture), pour la segmentation des mots en caractères et pour l'extraction des primitives [61-64].

En utilisant la ligne de base, les caractères et la forme sont classés en trois groupes ascendants, descendants et marques spéciales appelées diacritiques tels que points, shadda (Zigzag) et maddah, et ces groupes peuvent être construits à partir d'un trait ou d'un petit élément ou d'un caractère complet. La langue arabe formes de caractères basés sur la ligne de base montrée dans la voir [figure II.7](#). Les ascendants (*Jambages*) se trouvent au-dessus de la ligne de base, les descendants (*Hampes*) se trouvent sous la ligne de base, et diacritiques se trouvent soit au-dessus soit sous la ligne de base selon le caractère.

Dans [64] et [65], les auteurs démontrent l'impact et l'importance de la qualité de la ligne de base sur les performances des techniques de reconnaissance qu'ils ont utilisée. Il y a une multitude de techniques qui sont employées pour l'extraction de la ligne de base.

Al-Shatnawi [67] ont divisé les algorithmes de détection des lignes de base en texte arabe en trois catégories: méthodes basées sur des projections horizontales, méthodes basées sur des représentations de contours de mots et méthodes basées sur l'analyse en composantes principales (ACP). Une modification de l'approche simple basée sur la projection horizontale a été utilisée dans *Al-Khateeb* et al. [94] où la recherche de pic n'a été effectuée que dans la moitié inférieure de l'image du mot. Cependant, cette modification peut bien fonctionner lorsque la ligne de base est relativement droite. La transformation de l'image du mot en un espace de paramètres de Hough avec une détection maximale ultérieure pour la détection de base a été utilisée dans [54], [65], [68] et [83] a présenté une méthode de détection de la ligne de base, où ils ont localisé les minimums locaux dans le contour et utilisé une approche de maximisation pour localiser la ligne de base correcte.

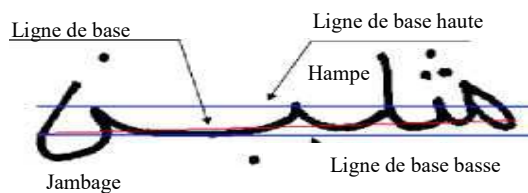


Figure II.7 lignes de base

A. Méthodes basées sur la projection horizontale :

C'est la méthode la plus utilisée pour l'extraction de la ligne de base d'un écrit arabe [68]. Cette méthode, simple à implémenter, fournit de bons résultats aussi bien pour le texte imprimé [48] que pour le texte manuscrit si celui-ci est de bonne qualité, qu'il est composé de mots assez longs et que l'écriture est bien droite [50],[65] [69]. (voir [figure II.8](#)).

Par ailleurs, pour les mots inclinés, *Pechwitz* et ses collaborateurs [71] combinent cette

technique avec la transformée de Hough (voir [figure II.9](#)).

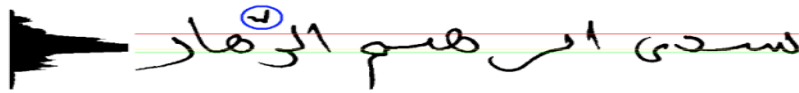


Figure II.8 Estimation de la ligne de base par la projection horizontale

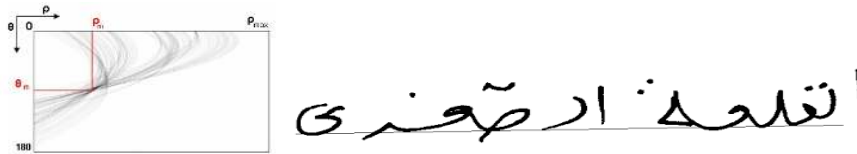


Figure II.9 Extraction de la ligne de base par la la transformée de Hough

Dans l'écriture arabe, un des défis majeurs de la technique de la projection horizontale réside dans l'existence d'un grand nombre de signes diacritiques et de descendants avec des parties horizontales quelque peu longues qui sont capables de produire potentiellement des pics parasites et altérer ainsi l'histogramme (voir [figure II.8](#)). *Menasri* et son équipe contournent cette difficulté en éliminant, en premier abord, les signes diacritiques. Dans une seconde étape, ils considèrent l'emplacement des boucles comme référence pour retrouver le pic maximal [69][127]. L'emploi de la position des boucles est dicté par le constat suivant : dans l'écriture arabe les boucles se trouvent toujours à proximité de la ligne de base. Par conséquent, la zone de recherche du maximum de l'histogramme est limitée à une bande horizontale large de trois fois la hauteur maximale des boucles et centrée sur les boucles



Figure. II.8. Problèmes liés à la méthode de projection horizontale : pics parasites dûs à la succession de descendants avec des portions horizontales suffisamment longues. [69]

Certains auteurs [69] ont tiré profit de ce constat en réduisant la région de recherche du pic maximum par la reconnaissance des boucles. Le point faible de cette approche réside dans le fait que cette méthode, pour qu'elle soit efficace, exige la présence des boucles dans la forme du mot. En effet, ces auteurs s'abstiennent d'aborder le cas des mots sans boucle ou bien le cas des mots ayant des boucles colmatées ou encore celui des mots avec une boucle qui se trouve en haut du mot. Tous ces mots, qui font partie du vocabulaire de la base IFN/ENIT), ne vérifient pas l'hypothèse du départ. Cette approche de reconnaissance n'est pas donc vérifiable dans ces cas-ci.

B. Méthode basée sur le squelette

Pechwitz et al ont suggéré une méthode qui fait appel au squelette pour l'extraction de la ligne de base [49]. Cette technique utilise les étapes suivantes :

- L'extraction du squelette du mot au moyen d'une approximation polygonale ;
- Créent le squelette du mot au moyen des approximations polygonales ;
- Ensemble des caractéristiques calculées, basées sur le squelette polygonal, ces caractéristiques représentent les caractéristiques pertinentes de la ligne de base ;
- Toutes ces caractéristiques sont utilisées dans une première approximation de la ligne de base. Les lignes du squelette qui sont quelques peu horizontales sont marquées. La position y de leur point central est ensuite pondérée par certaines mesures. La somme de ces points pondérés détermine ensuite, en première approximation, la position y de la ligne de base. Celle-ci sert à décrire enfin une bande de base horizontale dont la hauteur correspondrait au tiers de la hauteur totale du mot ;
- L'extraction des points qui supportent la ligne de base finale : les points centraux des lignes horizontales situés dans la bande de base ou dans ses proximités, les minimums des boucles qui sont dans cette bande et les points supérieurs des longs tracés courbes relatifs aux descendants ;
- La ligne de base détectée en utilisant une régression linéaire en fonction des caractéristiques pertinentes sélectionnées qui se convertiront au point d'estimation de la ligne de base à l'étape précédente.

C. Méthode basée sur le contour

Des chercheurs ont révélé une autre technique d'extraction de la ligne de base qui s'appuie sur le contour du mot [64]. Pour ce faire, une première appréciation de la ligne de base est obtenue en appliquant la régression linéaire sur tous les minimums locaux du contour. Cette ligne les autorise de déterminer les minimums locaux appropriés qui sont situés à proximité de la ligne de base. Cette dernière est obtenue finalement par l'application d'une seconde régression linéaire des minimums locaux les plus saillants. *Farooq* [65] signale l'obligation d'effectuer cette technique sur un grand nombre de minimums locaux.

D. Méthode hybride

Dans [69], *Farès Menasri* et ses collaborateurs effectuent une combinaison entre l'histogramme de projection horizontale, le squelette et le contour afin d'extraire la ligne de base. La ligne obtenue est donnée par l'interpolation linéaire des points supports fournis à partir du squelette et du contour du mot. Des contraintes propres à la morphologie du

manuscrit arabe sont ensuite effectuées afin de vérifier la sélection des points supports. Cette sélection s'opérant dans une bande horizontale ayant été définie au préalable.

Pour obtenir cette bande, ces chercheurs effectuent un calcul du pic maximal de l'histogramme de projection horizontale. Comme cela a été dit plus haut, la position des boucles est prise en compte afin d'éviter les pics indésirables engendrés par l'alternance des descendants avec des parties horizontales plutôt longues. Ensuite et à partir du maximum de l'histogramme, ils définissent un seuil haut et un seuil bas qui leur facilitent l'obtention d'une bande de base approximative. De manière empirique, ces seuils haut et bas ont été respectivement fixés à 20% et à 45% de la valeur du pic principal. A l'inverse de ce que *Pechwitz* a recommandé [49], [65] envisagent des lignes de base droites, qui sont généralement, peu courantes dans une écriture manuscrite libre dans laquelle un mot est susceptible d'avoir plusieurs inclinaisons. Jusqu'à preuve. Comme résultat, la ligne de base ne se réduit pas alors à une seule ligne droite mais se présente plutôt comme une succession de segments de droite. Chaque segment est relié par deux points supports consécutifs. Les auteurs ont alors rapporté qu'un taux de 88.4% a été atteint en appliquant une procédure d'évaluation identique à celle adoptée par [49].

E. Méthode basée sur l'analyse des composants principaux (ACP)

L'analyse en composantes principales peut être utilisée pour compresser et reconnaître les images de deux et trois dimension, ainsi que pour la détection de la ligne de base arabe [67]. Deux travaux de *Al-Shatnawi* [67] et *K.A. Berthe* [11] pour détecter la ligne de base de l'écriture arabe basée sur le principe de l'analyse des composants, cette méthode basée sur l'angle détecté par le principe de l'analyse des composants. Il est utilisé pour déterminer la direction de la ligne de base du mot arabe en fonction de la distribution des pixels de premier plan ou d'arrière-plan. Ce travail se résume en quatre étapes:

- 1- Effectuez l'analyse des composants principaux pour la distribution des pixels de premier plan ou d'arrière-plan.
- 2- Déterminer la direction d'estimation de la ligne de base en utilisant le vecteur propre PCA.
- 3- Faites pivoter les images selon la direction de l'estimation et calculez les angles de rotation.
- 4- Appliquer une projection horizontale pour déterminer le pic le plus long et détecter la position verticale de la ligne de base.

Cette méthode fonctionne à la fois avec les pixels de premier plan ou d'arrière-plan et donne de bons résultats avec des pixels d'arrière-plan plus qu'avec le premier plan. Il fonctionne

également avec les diacritiques ou sans diacritiques, mais il fonctionne mieux sans diacritiques. Il peut également être appliqué dans le texte imprimé.

II.5.2 Estimation de la bande de base

Selon *M. Pechwitz* et ses collaborateurs [71], l'extraction de la partie supérieure de la bande de base est une opération bien plus ardue que celle de l'évaluation de la ligne de base elle-même. Dans leur étude, ils abordent pour cela deux techniques : la première consiste à supposer que la ligne de base supérieure est située sur une ligne parallèle à la ligne de base et qu'elle se trouve à environ 40% à mi-chemin entre la hauteur de la ligne de base et le sommet du mot. Cette méthode a fourni des résultats à de 74.3% pour un décalage de 7 pixels. La seconde technique utilise l'histogramme de projection et la transformée de Hough. Cette dernière a enregistré un résultat de 59.3% pour un décalage de 7 pixels.

Les résultats de ces deux expériences ne sont pas tout à fait satisfaisants. Aussi, les auteurs conseillent ainsi de ne plus prendre en considération l'extraction de la bande de base supérieure dans les systèmes de reconnaissance de l'écriture arabe [69]. En revanche, ils préconisent l'utilisation de la partie supérieure de la bande de base uniquement pour la normalisation verticale des mots.

II.5.3 Segmentation des mots en caractères

La segmentation des mots cursifs en lettres ou morceaux de lettres (graphèmes) est une étape indispensable pour les systèmes de reconnaissance analytique. Elle désigne le découpage du mot en ses éléments constitutifs afin de les identifier isolément. Dans le cas idéal, ces éléments présentent des lettres ou des graphèmes, dans le cas contraire, deux problèmes sont envisagés : la sur-segmentation lorsque l'élément constitutif est lui-même fragmenté, et la sous-segmentation lorsque plusieurs éléments constitutifs n'ont pu être isolés.

Au fil des années, les chercheurs à travers le monde ont mis au point énormément de techniques en vue de segmenter les mots. Parmi celles-ci, les méthodes dont le principe s'appuie sur la projection verticale ou encore celles qui utilisent l'histogramme de l'image sont toutes les deux reconnues comme étant celles qui sont les plus employées dans le domaine du manuscrit arabe. Il existe également d'autres techniques qui font appel soit au contour, soit à la squelettisation, soit aux réseaux de neurones, ou encore à base de graphes et autres techniques morphologiques [72-79]. (Nous aborderons cette phase plus en détail dans le prochain chapitre).

II.5.4 Extraction de primitives

Une étape d'extraction de caractéristiques est réalisée de manière à obtenir l'information la

plus discriminante pour la tâche de reconnaissance et également pour réduire le volume d'informations qui sera fourni au système. C'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. Les techniques d'extraction des primitives dans le domaine de la reconnaissance du manuscrit arabe se distinguent en fonction de la nature des caractéristiques utilisées. De manière générale, ces dernières sont classées en trois grandes catégories : caractéristiques statistiques, caractéristiques structurelles et transformations globales. Nous montrons dans le tableau qui suit les principales caractéristiques qui sont obtenues dans les systèmes de reconnaissance hors ligne des mots arabes manuscrits.

Tab II.4 Primitives utilisées dans les systèmes de reconnaissance de mots arabes manuscrits.

SYSTÈME/ANNÉE	CARACTERISTIQUES	CARACTÉRISTIQUES
Miled [80] /1997	Structurelles	Les points et les signes diacritiques Les points de segmentation décisifs et leurs positions. Le nombre de graphèmes dans chaque tracé. La longueur du mot en graphèmes. Le nombre de tracés sans compter les Alifs. Le nombre de hampes dans chaque tracé.
Miled [81]/ 1998		Une séquence chronologique de droite à gauche d'indices visuels codés.
Khorsheed [82]/ 1999		Les boucles et les transformations de segment.
Khorsheed [3]/2003		Points diacritiques, points extrêmes, points de jonctions et boucles
Farah [7] /2006		La chaîne de code de Freeman, nombre d'ascendants, nombre de descendants, boucles et
Al-Abed [83]/2007		Primitives de directions calculées à partir du squelette du mot.
Zavorin[102] /2008		Boucles, points diacritiques, points extrêmes et de jonctions
Al-Nuzaili[146] /2016		Points diacritiques, Boucles, formes de caractères
F. Zaiz [165] /2016		La direction, Taille relative des segments, Nombre de boucles, Forme des boucles, Angles des segments
Burrow [84] /2004	Statistiques	Les moments de Zernike
A. Mubarak [85]/ 2012		La technique de zoning

Chergui [86]/ 2012		Moments de Tchebichef Moments de Zemike Moments de Hu
Azizi [87]/2002	Statistiques et Structurelles	Moments de Hu Zoning, ascendants, descendants, boucles et points diacritiques.
Al-Maadeed [4]/2004		Les boucles et leurs positions d'intersection ,les points diacritiques et leur position, les points extrêmes, les points de jonction Les moments invariants, et la densité de l'histogramme horizontal
Farah [88] /2005		Les ascendants, les descendants et les boucles. La technique de Zoning.
Gasmi [89]/ 2005		Le nombre des pseudo-mots, le nombre d'ascendants, le nombre de descendants, les boucles et les points diacritiques. La densité des pixels.
Benouareth [90]/2008		La technique de Zoning Points caractéristiques Concavités
R.Alhadj [91]/2009		Distribution de pixels et concavités
Kessentini [92]/2010		Directions, contours, et densités de pixels
Abdel Azeem [93]/ 2013		Concavités Gradient
Al Khateeb [94]/2009		Globales et statistiques
Broumandnia[95]/2008	Globales	La transformée en ondelettes
Khorsheed [3]/ 2000		La transformée de Fourier.
Chen [96] /2010		Gabor
Nemmour [97]/ 2011		La transformée de Ridglet
Elzobi [98]/ 2014		Gabor La transformée Wavelet
El Qacimy [99] /2015		DCT
A. Saidani[100] /2015		HOG

II.6 Différentes approches et systèmes existants

Des *Survey* sur la reconnaissance de l'arabe sont disponibles dans [38],[48],[58],[63][67],[101],[158]. Trois stratégies principales ont été appliquées à la reconnaissance des caractères arabes imprimés et manuscrits. Celles-ci peuvent être classées en deux catégories: les approches globales (holistiques) et les approches analytiques.

Pour l'approche holistique, la reconnaissance s'effectue globalement sur toute la représentation du mot sans tentative d'identification individuelle des caractères.

Avec l'approche analytique, la reconnaissance n'est pas effectuée directement au niveau des mots mais à un niveau intermédiaire traitant d'unités ou de segments. Dans cette stratégie, les mots ne sont pas considérés comme un tout, mais comme des séquences d'unités ou de segments de petite taille (graphèmes).

II.6.1 Approche holistique Par HMM

Au regard des triomphes formidables enregistrés par les HMM dans le domaine de la reconnaissance de la parole et la reconnaissance de l'écriture latine, de nombreux chercheurs ont eu recours à ces techniques pour les exploiter dans la reconnaissance des arabes manuscrits. *El-Hajj* [102],[91]. *Khorsheed*[03], *Kessentini*[92], *Al-Ma'adeed* [4][37][151], *Saon*[12], *Ahmad* [30], *Augustin*[32], *Slimane*[46], *Abdel*[93], *X. Dupre* [134].

II.6.2 Approche holistique par RNN

Nombreux sont les chercheurs qui ont su exploiter les réseaux de neurones afin d'améliorer les performances des systèmes de reconnaissance hors ligne appliqués aux mots arabes manuscrits. Parmi ces auteurs, *Alm'addeed* [83]. [97-99]. *Alex Graves* [100]

II.6.3 Approche holistique par (SVMs)

La SVM est une technique de classification spécialement bien adaptée afin de prendre en compte des données de très grandes dimensions, telles que les textes et les images. Ces caractéristiques ont amené les chercheurs à les utiliser dans le domaine de la reconnaissance du manuscrit arabe. *H. Nemour* [107],[108]. *A. Mubaak* [85], *EI qacimy*[99].

II.7 Conclusion

Ce chapitre nous avons présenté une vue globale des notions de bases des systèmes de reconnaissance de l'écriture arabe manuscrite. Nous avons, en premier lieu, exposé les différentes propriétés morphologiques de l'écriture arabe manuscrite ainsi que les difficultés rencontrées lors de la conception d'un système de reconnaissance performant et une vue globale sur les systèmes de reconnaissance d'écriture arabe et quelques bases de mots existantes. Enfin, les aspects et les différentes méthodes employées dans la reconnaissance de l'écriture arabe manuscrite.

Chapitre **3**

*La Segmentation de
L'écriture Arabe
Manuscrite*

III.1 Introduction

Un des défis majeurs que l'on doit à tout prix surmonter lors de la réalisation d'un système automatique de reconnaissance de l'écriture manuscrite concerne la segmentation des caractères. En effet, son impact sur la performance de ce système n'est plus à démontrer. Un choix judicieux de la segmentation augmente considérablement cette performance. L'on peut partager les techniques de segmentation [103-107] en quatre grandes familles distinctes. *Primo* : la *segmentation par dissection*, qui consiste à isoler diverses unités d'écriture avant l'opération de reconnaissance. *Secundo* : la *segmentation par reconnaissance* qui fait appel à des techniques de reconnaissance [108], [109]. Cette approche est spécifiquement utilisée dans l'écriture cursive *Tertio* : la *segmentation hybride* qui conjugue la technique de dissection avec des méthodes de recherche et enfin la *segmentation holistique* qui ignore tout simplement la segmentation des caractères et qui procède plutôt à la reconnaissance du mot en entier [110]. Sur le plan pratique, la segmentation et la reconnaissance sont intimement interdépendants. Ainsi, la segmentation est d'autant plus efficace que lorsqu'elle répond de manière satisfaisante aux contraintes contextuelles aussi bien au niveau du mot (telles que le nombre de caractères qui composent ce mot) qu'au niveau de la phrase (telles que le nombre de mots constituant cette phrase). En conséquence, l'interdépendance entre la segmentation du caractère et la reconnaissance du caractère est un aspect primordial dont il faut en tenir compte lorsqu'il s'agit d'entreprendre une segmentation des caractères efficiente [111]. A cet égard, nous présentons dans ce chapitre l'état de l'art en ce qui concerne les techniques de segmentation de caractères les plus importantes qui ont été appliquées sur les écrits manuscrits. Nous ferons également un tour d'horizon sur les techniques utilisées pour séparer les caractères se chevauchant et qui sont extraites de la littérature spécialisée.

III.2 Segmentation de l'écriture cursive

La segmentation est, par définition, une opération qui consiste à découper des régions connexes qui appartiennent à un même environnement spatial. Ainsi, la segmentation de caractères permet de fragmenter un mot manuscrit en un ensemble d'unités d'information de base qui serviront à la reconnaissance de ce mot.

La segmentation du caractère est une étape primordiale dans la reconnaissance automatique des caractères. Malheureusement la segmentation est loin d'être une opération aisée. Elle est encore plus compliquée dans le cas de l'écriture cursive. Il est unanimement admis, dans la communauté des chercheurs spécialisés dans la reconnaissance de l'écrit manuscrit, qu'il est pratiquement impossible de segmenter de manière directe un mot cursif en lettres distinctes.

En effet, afin de pouvoir segmenter efficacement un mot, il est indispensable de l'identifier en premier lieu. Par ailleurs, pour identifier un mot, il faudrait au préalable le segmenter correctement. C'est ce qu'on appelle le paradoxe de *Sayre* [162]. Afin de contourner ce dilemme, il est impératif de sectionner le mot en groupes de lettres.

X. Dupré [55] a réalisé une étude exhaustive sur la segmentation de l'écriture cursive à laquelle on peut se référer. On y distingue deux types de segmentation :

- la segmentation en graphèmes [92], connue aussi sous l'appellation segmentation explicite, le mot est segmenté en sous-parties pouvant être pratiquement des lettres (figure III.3)
- l'analyse par fenêtres glissantes, également connue sous l'appellation segmentation implicite, le mot est découpé en bandes verticales.

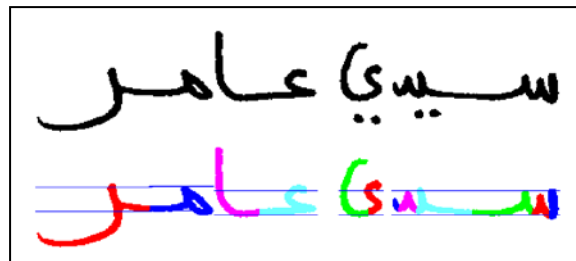


Fig. III.1 Exemple de segmentation en graphèmes

III.3 Difficultés dans la segmentation des caractères

La segmentation de caractères est une étape très ardue dans le processus de reconnaissance automatique de l'écriture manuscrite et ce pour plusieurs raisons : la dégradation éventuelle de l'image acquise, le type d'écriture cursif, le grand nombre de variétés dans les styles d'écriture, le chevauchement de caractères, la fragmentation de caractères, ou encore l'appartenance des caractères à un fond texturé [112], [113]. Sur un autre plan, ce qui rend l'opération de segmentation encore plus laborieuse est sa grande dépendance vis-à-vis de la reconnaissance due à la similarité visuelle de certains caractères. Par voie de conséquence, il est fatalement très difficile de prendre la bonne décision quant au choix de l'hypothèse qui peut être potentiellement candidate aux dépens d'autres hypothèses pour un caractère donné. Il faut également savoir qu'un grand nombre de caractères sont visuellement ressemblants. De même qu'il existe des caractères qui replient exactement les parties d'autres caractères. Tout cela ne fait, en fait, qu'accroître la difficulté de cette tâche [114]. Notons enfin que si, par exemple, la segmentation envisagée admette plusieurs coupes, certaines lettres peuvent alors être divisées en sous-lettres.

III.4 Les méthodes de segmentation des caractères

La localisation des mots dans les textes manuscrits est un problème qui relève du paradoxe de *Sayre* (1973) « Il faut segmenter le tracé pour reconnaître une lettre, mais il faut reconnaître une lettre pour segmenter le tracé » [162]. En effet, si théoriquement les règles de disposition de l'écriture imposent de marquer des espaces entre les mots plus importants (lue les espaces entre caractères ce qui est le cas pour l'imprimé), en pratique ces règles ne sont pas toujours vérifiées sur les écritures manuscrites non contraintes. Et on doit admettre que pour résoudre le problème, il est nécessaire de demander l'aide d'un système de reconnaissance. De ce fait la localisation des mots dans une ligne de texte est un problème qui s'apparente à celui de la reconnaissance des caractères dans les mots.

Jadis, les approches utilisées dans la segmentation de caractères étaient fondamentalement fondées sur l'analyse du profil de projection et sur l'estimation du pitch (qui constitue le nombre de caractères par unité de distance horizontale). Ces méthodes ont fourni des résultats satisfaisants quant à la reconnaissance des caractères imprimés qui présentent implicitement des espacements réguliers. L'examen des parties connexes est ensuite utilisé afin de traiter les ligatures et la police italique. Aujourd'hui, des améliorations ont été en outre apportées à ces techniques, lesquelles combinent la segmentation avec la reconnaissance suivi de l'utilisation des résultats de la reconnaissance aussi bien que du contexte linguistique comme feedback afin d'améliorer encore plus les décisions de segmentation [114].

Fondamentalement, les techniques de segmentation peuvent être globalement catégorisées en quatre grandes familles à savoir (figure III.2) : la segmentation par *dissection*, la segmentation par *reconnaissance*, la segmentation *hybride* et la segmentation *holistique*.

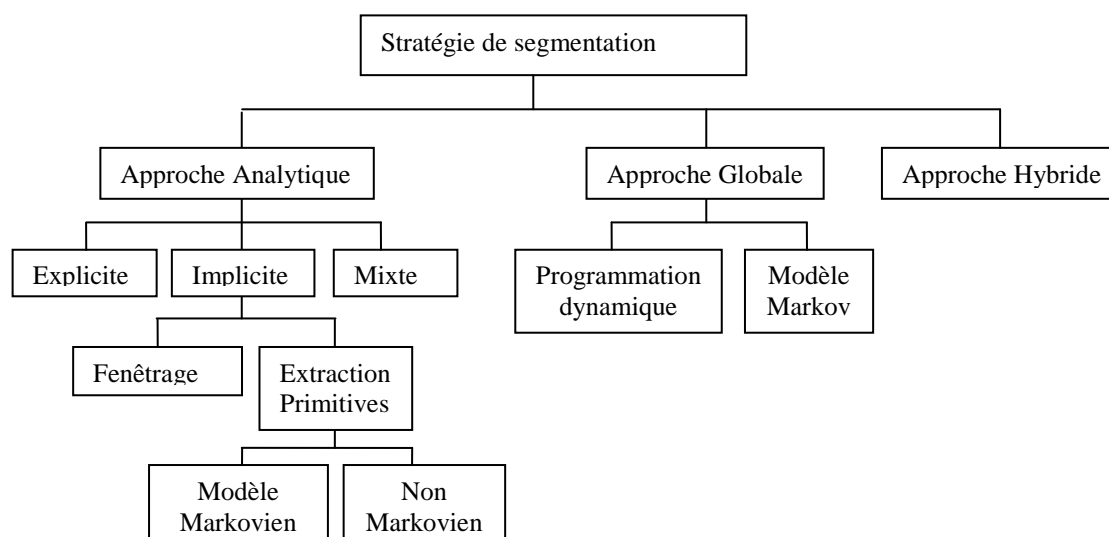


Figure III.2: Classification des méthodes de segmentation de caractères [127]

III.4.1 La segmentation par Techniques de dissection

La segmentation par dissection est une opération préalable à celle de la reconnaissance. Elles sont donc indépendantes, du point de vue temporel, l'une de l'autre. Par dissection, on entend la démarche qui consiste à repérer les lieux exacts où doivent s'effectuer la scission des caractères. Il est ensuite procédé à une identification des segments basée sur les propriétés des caractères. Ce processus de découpe de l'image en éléments significatifs est appelé "dissection". Finalement, La dissection est un processus intelligent dans lequel une analyse de l'image est effectuée; Cependant, la classification en symboles n'est pas impliquée à ce stade.

Les techniques de la segmentation par dissection font donc appel à :

- L'analyse des composants connexes,
- L'espace blanc et l'estimation de pitch,
- Landmark (points de repère),
- L'analyse des profils.

Nous décrivons à présent un certain nombre de méthodes de cette catégorie de segmentation.

III .4.1.1 L'analyse des composants connexes

La composante connexe est la région maximale de pixels connexes (non séparés par un contour), la définition de l'ensemble de composantes connexes est la division de l'image en segments [116].

Différents algorithmes, basés sur différentes structures, sont disponibles afin de déterminer les composants connexes d'images binaires tels que :

- *La diffusion récursive* des composants connexes qui consiste à balayer de manière récursive la matrice d'image.
- *Les chaînes de codage* qui consistent à tracer le contour autour d'un composant dans la matrice d'image.
- *Analyse des composants connexes basée RLC* (Run-Length-Contour).

Pour déterminer les composants connexes, il convient de faire appel à l'algorithme ci-dessous dont le principe est fort simple :

- Repérer le prochain pixel non étiqueté dans la matrice d'image.
- Diffuser récursivement aux voisins.
- Marquer les pixels visités par l'étiquette courante.
- Une fois le composant est épuisé, sélectionner la prochaine étiquette et répéter

III .4.1.2 L'espace blanc et l'estimation du pitch

La notion de l'espace blanc vertical entre les caractères successifs est un concept important employé afin de fractionner les caractères, en particulier, pour les documents présentant une

police imprimée dans lesquels chaque caractère occupe un espace de largeur bien définie. Le pitch, ou encore le nombre de caractères par unité de distance horizontale, fournit une base à même de permettre de trouver les points de segmentation. La combinaison de l'espace blanc et de l'estimation du pitch sont susceptibles de donner des résultats de segmentation fiables sur des caractères bien espacés ayant une largeur fixe. A noter que l'estimation du pitch a permis aussi une segmentation exacte dans le cas de la présence dans le texte de caractères fragmentés ou se chevauchant [114].

Dans de nombreuses applications d'impression sur machine impliquant des jeux de polices limités, chaque caractère occupe un bloc de caractères de largeur fixe. La hauteur, ou le nombre de caractères par unité de distance horizontale, fournit une base pour l'estimation des points de segmentation. La séquence de points de segmentation obtenue pour une ligne d'impression donnée devrait être approximativement équidistante à la distance correspondant à la hauteur. Ceci fournit une globale base de la segmentation, car les points de séparation ne sont pas indépendants.

L'application de cette règle permet une segmentation correcte dans les cas où plusieurs caractères de la ligne sont fusionnés ou cassé. Si la plupart des segmentations peuvent être obtenues en trouvant des colonnes de blanc, la grille régulière des limites entre caractères peuvent être estimées. Les points de segmentation ne se trouvant pas près de ces limites peuvent être rejetés comme probablement dû à des caractères brisés. Les points de segmentation manqués en raison de la fusion de caractères peuvent être estimé également, et une analyse locale menée afin de décider où scinder le mieux le composite

III .4.1.3 Segmentation à partir Les profils de projection verticales

Cette méthode. Elle consiste à calculer des histogrammes de projection dans plusieurs directions proches de la verticale. Les droites choisies sont celles qui interceptent le moins de pixels noirs, avec une contrainte d'espacement régulier dans l'image.

Afin de segmenter les lignes d'un document, on a le plus souvent recours au profil de projection horizontal. Malheureusement, ces profils ne fournissent pas des résultats probants lorsqu'il s'agit de les appliquer à une écriture cursive, ou encore lorsqu'on est en présence de caractères italiques ou qui se chevauchent [119].

Le Profil horizontal est défini comme étant la somme des pixels noirs qui sont perpendiculaires à l'axe des x. Ceci est représenté par le vecteur p_h , de taille M, défini comme suit:

$$p_h[i] = \sum_{j=1}^N s[i, j] \quad (\text{III. 1})$$

Où : $S(N, M)$ une image binaire de N lignes et M colonnes.

Les espaces existants entre deux mots consécutifs voire même les tracés verticaux sont aisément détectables par le profil de projection horizontal [120].

Le Profil vertical défini comme étant la somme des pixels noirs qui sont perpendiculaires à l'axe des y . Ceci est représenté par le vecteur p_y , de taille N , défini comme suit:

$$p_v[i] = \sum_{j=1}^M s[i, j]. \quad (\text{III. 2})$$

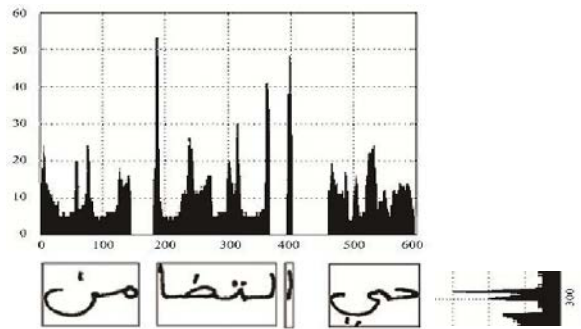


Figure III.3 Profils de projection horizontal et vertical [119].

Dans la **figure III.3**, le profil de projection vertical fait apparaître deux crêtes. L'une correspondant à la ligne des x et l'autre à la ligne de base. Parallèlement, le profil de projection horizontale montre des régions sans aucun pic qui coïncident avec les blancs qui séparent deux mots consécutifs ou encore les blancs qui séparent les caractères d'un même mot. L'utilisation de la méthode de la projection ou de l'histogramme simplifie le problème de segmentation des caractères dans un système 1D à la place d'un système 2D. Dans [40], Najoua et Noureddine ont présenté une méthode basée sur l'histogramme modulé, ainsi que le nombre de segments noirs dans une ligne de pixels. L'algorithme consiste des étapes suivantes: identification des lignes de texte, segmentation lignes de texte en morceaux de mots arabes (PAW), séparant chaque PAW en composants connectés, localisant approximativement son des différents personnages dans le PAW, calculant la nombre maximum de segments noirs dans une ligne de pixels, extraire les primitives et enfin utiliser un correcteur d'erreur détecter les erreurs de segmentation. Cette technique a été appliquée sur des polices arabes imprimée.

III .4.1.4 Détection des bornes limites

Les attributs qui caractérisent une image, telles que les ascendantes et les descendantes d'un caractère, sont susceptibles d'être exploitées dans un mot manuscrit comme bornes limites afin de concourir à la segmentation des caractères de ce mot. Ce faisant, la segmentation de caractères basée sur la détection des ascendantes et des descendantes a été appliquée aussi bien au texte imprimé qu'à l'écriture cursive. Une illustration de la détection des bornes limites pour la segmentation de caractères est montrée dans [118].

III.4.2 La segmentation par reconnaissance

La segmentation basée sur la reconnaissance est une autre approche dans laquelle le système recherche des sous-composants dans le texte cursif qui correspondent aux lettres de son alphabet. A ce titre, une lettre ne peut être segmentée sans être identifiée auparavant. La segmentation et la reconnaissance des caractères ne sont pas des étapes totalement séparées avec un degré de dépendance variable. La sélection initiale des points peut être faite de différentes manières [111] [124]. Par exemple, en utilisant une fenêtre mobile avec une largeur prédéfinie qui divise le mot en plusieurs morceaux qui se chevauchent sans tenir compte de son contenu. Ensuite, une méthode de reconnaissance itérative ou parallèle est utilisée pour rechercher une notation de classification "satisfaisante" pour les sous-composants communs, généralement en générant un réseau de toutes ou de plusieurs combinaisons possibles de l'ensemble de candidats initial. La décision finale est déterminée par le meilleur chemin à travers le réseau. Tout en évitant d'utiliser des méthodes de dissection complexes, ces techniques reposent largement sur la précision du classifieur, qui affecte considérablement la précision de la segmentation.

III .4.2.1 Les approches de fenêtrage

Le principe de base dans cette technique est l'emploi d'une fenêtre glissante de largeur variable sur les caractères à identifier dans le but d'extraire des séquences de segmentation expérimentales. Celles-ci doivent être ensuite validées par la reconnaissance des caractères [111] [112]. Deux approches ont été retenues dans cette technique : l'approche séquentielle et l'approche parallèle. La première approche examine les mots itérativement en glissant la fenêtre de gauche à droite. La seconde approche procède, quant à elle, de façon plus globale en faisant générer une grille d'attributs aux différentes combinaisons possibles des lettres. La décision finale est prise en considérant le chemin optimal à travers le graphe des attributs possibles [114].

III .4.2.2 Les approches basées sur les attributs caractéristiques

Dans cette technique, les attributs caractéristiques de l'image sont dans un premier temps extraits. Les correspondances éventuelles entre les attributs et les lettres sont ensuite générées pendant la deuxième étape. Finalement, il est procédé à la recherche de la combinaison qui présente les meilleurs résultats [112]. Ce faisant, l'on peut dire que ce type d'approche réalise une segmentation implicite par le biais d'une classification des sous-ensembles d'attributs spatiaux de l'image. Deux catégories d'approches coexistent dans cette méthode : *Les approches basées sur le modèle de Markov caché* (HMM) et les approches *non-markoviennes*. Dans la première approche, l'objectif des modèles de Markov cachés est de

modéliser les variations de l'écriture cursive comme étant une structure probabiliste non directement observable. sont représentées par des variables aléatoires, dont la distribution dépend de l'état. Elles constituent une représentation séquentielle des caractéristiques de l'image d'entrée. Ces caractéristiques peuvent représenter :

- la variation du langage lettre par lettre.
- La transition à l'intérieur du caractère état par état.
- La variation dans le mot, état par état dans l'ensemble de mots admissibles d'un lexique

Par ailleurs, les modèles de Markov cachés peuvent être utilisés pour la segmentation des caractères en raison des contraintes contextuelles liées aux documents analysés telles que la fréquence des mots et des lettres. Aussi, est-il possible de formuler des règles syntaxiques à l'aide des probabilités de transition entre les différents états d'un caractère [126], [127].

Les approches de type *non markoviennes* à l'image de la relaxation probabiliste et du test d'hypothèse ont été exploitées dans la segmentation par reconnaissance. Ainsi, Hayes [130] a adopté à cet effet une méthode de relaxation probabiliste afin de déchiffrer les mots manuscrits. Pour ce faire, le modèle utilise une description hiérarchique des mots dérivés à partir de la relaxation d'une représentation squelettique. La relaxation a été accomplie sur le graphe des primitives ainsi que des lettres dans lesquelles toutes les segmentations potentiellement intéressantes ont été représentées [114].

III.4.3 La segmentation hybride

Les méthodes hybrides opèrent une combinaison entre les méthodes de dissection et les méthodes de recherche des points de coupure par reconnaissance de manière hybride. A cet égard, un algorithme de dissection est appliqué à l'image. Le nombre d'endroits dans lesquels les découpes sont pratiquées doit être suffisamment élevé de manière à ce que les frontières réelles qui séparent les caractères soient inévitablement comprises dans les coupures réalisées. Une fois que cela est réalisé, la segmentation optimale est alors obtenue par l'évaluation individuelle des sous-ensembles des découpes ainsi faites. Chaque sous-ensemble suppose l'existence d'une hypothèse de segmentation. La classification est ainsi appliquée pour estimer les différentes hypothèses afin de choisir la segmentation la plus encourageante [111] [131].

Un exemple pertinent qui relate le processus adopté dans ces méthodes est l'algorithme de segmentation exposé dans [140] qui établit une séquence de chiffres et de symboles, recouvre ensuite un feed-back au moyen du module de reconnaissance afin de réévaluer les partitions et recommencer une nouvelle fois le processus si nécessaire. Ainsi, la segmentation est

considérée comme conforme si tous les chiffres sont reconnus avec une grande marge de confiance.

Le système débute la segmentation par la séparation des caractères évidents. De la sorte, les chiffres qui ne sont pas suffisamment reconnus sont considérés comme étant composés de caractères liés. Ainsi, les blocs, qui ont été repoussés, sont eux-mêmes fractionnés et réexaminés dans une boucle de feed-back jusqu'à ce que la solution finale soit obtenue pour tous les segments. A noter que le système est en mesure de réunir également les fragments de chiffres avec les chiffres voisins afin de produire de nouveaux segments.

Dans le cas où ces segments ne sont pas encore reconnus, différentes techniques de fusion pourraient alors être envisagées. Il faut savoir que sur le plan pratique, plusieurs chiffres peuvent toucher leurs voisins directs et génèrent, en conséquence, un composant connexe. Dans ce cas de *figure*, le système recourt à un algorithme de division du contour afin de repérer les différents chemins possibles permettant la séparation des caractères recouverts, tels que l'analyse du contour, l'algorithme hybride Drop Falling ou bien encore l'algorithme étendu de Drop Falling que nous allons de suite exposer.

Pour augmenter en efficacité, cet algorithme procède de telle sorte que les vérifications de reconnaissance superflues soient, dans la mesure du possible, ignorées. Aussi, emploie-t-il à cet effet une approche de type heuristique afin de prendre un chemin donné susceptible de fournir les meilleures informations en se servant des différents attributs structurels. A ce titre, de multiples caractéristiques peuvent être adoptées dans le but de sélectionner les meilleurs chemins de coupure, en l'occurrence :

1. *Nombre de coupures faites pour diviser le segment*
2. *Longueur de la coupure*
3. *Bordure de la coupure*

La segmentation, opérée sur une partition composée de deux caractères (deux chiffres dans ce cas) à analyser, est considérée comme exacte si les deux chiffres sont distinctement bien identifiés. Dans le cas contraire, si aucun des deux chiffres n'a été convenablement identifié (niveau de confiance faible), alors d'autres chemins alternatifs seront envisagés.

Si seulement l'un des deux chiffres a été identifié, dans ce cas la partie restante sera elle-même considérée comme étant un groupe constitué de deux caractères (ou chiffres) accolés. Le procédé de segmentation, effectué précédent, sera réitéré de manière récursive sur cette partie jusqu'à ce qu'une solution satisfaisante soit obtenue.

Dans le but d'être encore plus précis, durant la reconnaissance de l'ensemble des chiffres constituant le montant d'un cheque par exemple, l'algorithme refait plusieurs essais de

fusion/séparation des caractères. Au cas où, à la suite d'une segmentation, il subsisterait encore des morceaux de caractères dont les dimensions seraient trop faibles, alors ces morceaux seront considérés comme étant des "fragments de caractère". A ce moment, ces fragments seront associés avec le segment voisin tout en respectant les critères de proximité et de chevauchement entre les caractères.

III.4.4 Segmentation holistique (implicite)

L'approche holistique ou la segmentation implicite (tente d'effectuer simultanément la segmentation et la reconnaissance), ou la segmentation libre, segmentée par les découvertes dans les études psychologiques du processus de lecture humaine. Dans les techniques sans segmentation, les mots sont reconnus dans leur ensemble sans être divisés en caractères, comme c'est le cas dans les applications de traitement de la parole. L'algorithme, présenté dans [133], est une illustration parfaite de la segmentation holistique cette technique est basée sur le principe des chaînes de Markov cachées ou de la programmation dynamique [131]. L'algorithme, utilisé dans cette technique, tente de reconnaître le mot dans sa globalité. De sorte que l'obligation d'opérer une segmentation du mot en caractères n'est plus de mise dans cette technique. La segmentation est complètement contournée. Toutefois, cette méthode fait appel par contre aux caractéristiques des mots. Il faut savoir, en effet, que ces derniers sont représentés par une liste d'attributs telles que les ascendantes, les descendantes, les boucles, ... etc. Pour reconnaître un mot, le principe de la reconnaissance holistique effectue une comparaison d'un ensemble d'attributs simples extraits à partir du mot entier avec une liste d'un lexique de chaînes représentant la forme théorique des mots possibles.

III.5 La technique détection des points de coupures

Les techniques de segmentation, dans leur grande majorité, procèdent à la reconnaissance de blocs isolés de Pixels et les considèrent en tant que composants connexes. Malheureusement ce n'est pas toujours le cas dans la réalité quotidienne des documents analysés. En effet, dans le cas d'une écriture cursive, plusieurs caractères sont accolés les uns aux autres, voire même, certains caractères se chevaucheraient complètement ou partiellement avec leurs voisins. Dans ce cas de figure, les techniques de segmentation considèreront ainsi ces blocs comme des composants connexes qui seront difficilement identifiables par tout système de reconnaissance. Pour pallier à ces difficultés, des algorithmes de segmentation effectuent afin de repérer les chemins possibles séparant les caractères en chevauchement. Parmi ces algorithmes, citons la méthode basées sur *l'analyse du contour*, *Segmentation à partir du squelette*, *Segmentation à base de fenêtres glissantes*...etc.

III.5.1 Analyse du contour

Ces méthodes basées sur des algorithmes de segmentation effectuent une division du contour des composants connexes afin de repérer les chemins possibles séparant les caractères en chevauchement. Parmi ces algorithmes, citons la méthode du *Hit-and-deflect*, l'analyse du contour, l'algorithme hybride *drop-and-falling*, l'algorithme étendu *drop-and-falling*, ...etc.

III.5.2 Segmentation à partir du squelette

A partir du squelette, on essaie de localiser certains motifs afin d'en déduire les candidats de points de coupures potentiels (voir [figure III.4](#)). La localisation de ces motifs conduit à des calculs de courbures et d'angles qui sont ensuite confrontés à des seuils ajustés de sorte à parvenir au résultat souhaité. A ce propos, *X. Dupré* a souligné dans ses travaux de recherche [\[134\]](#) que ce procédé reste inadéquat dans environ 10% des cas.

Elzobi et al. [\[148\]](#) proposent une approche de segmentation topologique optimisée. D'abord, le système applique une opération d'amincissement sur une image filtrée en utilisant un filtre médian pour obtenir une représentation squelette du contenu traité. Ensuite, applique une segmentation à deux étapes. Premièrement, il analyse l'axe des x des composants connectés pour extraire les sous-mots des images sans chevauchement. Puis, il extrait des segments de base en fonction des caractéristiques topologiques pour les représentants des caractères.



Figure III.4. Segmentation à partir du squelette

III.5.3 Segmentation à partir du contour

De nombreuses techniques utilisant le traçage de contour ont été développées pour résoudre le problème de la segmentation des caractères en chevauchement. Parmi ces techniques, citons la méthode du *Hit-and-deflect*, Minima des profils horizontaux et verticaux, l'algorithme hybride *drop-and-falling*, l'algorithme étendu *drop-and-falling*, ...etc.

III .5.3.1 Minima des profils horizontaux et verticaux

Il existe un cas de figure où le profil de projection vertical ne fournit pas des points de segmentation corrects. Cela arrive quand des caractères voisins se chevauchés horizontalement, [\[137\]](#) [\[138\]](#), [\[139\]](#).

III .5.3.2 L'algorithme Hit-and-deflect

Le principe des algorithmes *Hit-and-deflect* repose sur une idée assez simple. Cette idée consiste à tenter de trouver le chemin optimal de coupure d'un composant connexe en guidant un genre de "pointeau" à travers le composant connexe, tout en évaluant si les caractères analysés sont séparés ou en chevauchement [134] connexe. L'algorithme clôture le processus lorsqu'il accède au niveau supérieur de l'image [135].

III .5.3.3 L'algorithme Drop-and-Falling

L'algorithme *Drop and Falling* [135] peut être assimilé à une goutte d'eau située à un emplacement supérieur de l'image qui glisse vers le bas tout au long d'une coupure. Dans le cas où cette goutte d'eau se trouve prisonnière dans un coin donné, cela implique que l'on peut considérer qu'une coupure traverse cette partie de la figure de haut en bas.

III .5.3.4 L'algorithme des points de coupure critique

Cet algorithme est le produit d'une observation minutieuse de l'écriture cursive et qui peut s'énoncer ainsi : A chaque fois que les caractères se chevauchent, la jonction entre ces caractères renferme une ou plusieurs zones concaves tout au long de la coupure. Par voie de conséquence, si les points d'un coin concave se retrouvent identifiés, ceux-ci peuvent être, à leur tour, exploités afin de découvrir les points exacts où la coupure pourrait être entamée. En pratique, cet algorithme enclenche le processus de recherche à partir du point du centre horizontal du caractère connecté et sonde de haut en bas tous les points critiques de coupure. Il prospecte ainsi, à partir du centre horizontal, toutes les paires de points de coupure jusqu'à ce qu'une paire adéquate soit retrouvée. L'algorithme considère finalement que la meilleure paire a été trouvée qu'après la validation d'un certain nombre de critères tels que leur distance, leur distance par rapport au centre et l'angle de la coupure. A noter enfin que les coupures qui traversent un espace blanc sont interdites.

Dans [141] on examine d'abord la partie inférieure du contour pour voir s'il y a des caractères qui se touchent. La segmentation des caractères en contact est obtenue en traçant la partie inférieure du contour de droite à gauche. Les valeurs les plus basses dans la direction verticale sont enregistrées. On constate que le point de contact se produit entre deux de ces points et est la valeur la plus élevée dans la direction verticale qui remplit la condition de seuil mentionnée ci-dessus. Une fois qu'un point de contact est trouvé, les caractères sont séparés. Cela conduit à diviser le contour en deux parties ou plus en fonction du nombre de caractères en contact. En conséquence, le contour de la première partie est extrait à nouveau et la partie inférieure de ce contour est examinée. Cependant, les procédures de segmentation et de

reconnaissance sont interconnectées de sorte que chaque caractère segmenté est d'abord éclairci et que ses caractéristiques sont ensuite extraites. S'il n'est pas reconnu, sa segmentation est annulée et le point de segmentation suivant est pris et une tentative de reconnaissance est répétée. Cette procédure d'autocorrection est nécessaire pour accroître la précision des procédures de segmentation et de reconnaissance.

S. Madhvanath et al. ont prôné dans leur étude un procédé de segmentation que l'on peut appliquer aux contours [142]. Ainsi, ils avancent que les points de coupure entre les graphèmes sont potentiellement les meilleurs candidats. Ils se basent, pour cela, sur les extrema locaux du contour qui, eux, sont associés selon un critère de proximité.

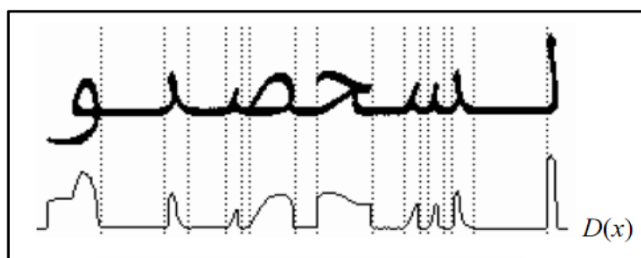


Figure III.4. Segmentation à partir du contour

III.5.4 Segmentation à base de fenêtres glissantes

Le principe de la fenêtre glissante [127] consiste à couper simplement l'image en bandes verticales. Cette découpe n'est pas nécessairement régulière, elle peut être ainsi irrégulière. Les bandes consécutives peuvent, éventuellement, se recouvrir partiellement. Outre au fait qu'il soit indépendant de la connexité, ce procédé a l'avantage d'être également simple et robuste au bruit. Néanmoins, l'inconvénient de cette technique réside dans le fait que la séquence générée contienne beaucoup de bruit dans le cas du recouvrement de deux lettres consécutives. Cela est tout aussi vrai lorsqu'on est en présence de lettres se superposant verticalement, mais qui ne se touchent pas forcément. C'est le cas, par exemple, de la barre de la lettre 'س' avec une lettre suivante, ou encore les descendants comme 'ج' ou 'ي' en arabe.

III.6 Exemples de méthodes de la segmentation de caractères

III.6.1 La segmentation par HMMs

Les HMMs ont été utilisés avec succès dans les applications de reconnaissance vocale [35],[37],[90-92]. Dans [136], HMM a été utilisé pour la segmentation de caractères arabes imprimée hors ligne. Une fois les traits secondaires supprimés, une fenêtre glissante a été utilisée pour balayer chaque ligne de droite à gauche. Les caractéristiques ont été extraites des fenêtres puis appliquées aux modèles HMM. M différents nombres de HMM ont été construits pour chaque caractère ou ligature. En utilisant seulement deux polices, les polices arabes

simplifiées et traditionnelles de Microsoft Word application, un taux de précision de segmentation de 99% a été enregistré.

Les performances du système diminuaient à mesure que le nombre de modèles augmentait lorsque davantage de styles de police et de traits secondaires étaient utilisés. Dans le système de reconnaissance vocale basé sur HMM a bien exécuté les tâches de OCR avec un minimum de modifications et sans pré-segmentation des données. [86],

El-Hajj et al. dans [102], a proposé un système de reconnaissance de l'écriture hors ligne 1D HMM utilisant une approche analytique. Le système est supporté par un ensemble de fonctionnalités robustes, indépendantes du langage, extraites sur des images binaires. Des paramètres tels que les lignes de base inférieure et supérieure sont utilisés pour dériver un sous-ensemble d'entités dépendantes de la ligne de base. Ainsi, la variabilité lexicale due aux parties inférieure et supérieure des mots est mieux prise en compte. De plus, le système proposé apprend les modèles sans pré-segmentation en caractère.

Le travail de *S. Alma'adeed, et al* [7], qui a débuté en 2002 et s'est poursuivi plus avant, ainsi que d'autres chercheurs de [4][151] ont proposé un système de reconnaissance optique des caractères arabe analogue au système de reconnaissance vocale automatique basé sur HMM. Les auteurs ont tenté de surmonter certains des problèmes suivants rencontrés dans les systèmes de reconnaissance optique des caractères en arabe: connexion, pointillage, graphèmes multiples, ligatures et diacritiques. Le système proposé a été testé sur trois polices différentes avec des ligatures. Les résultats montrent que le taux d'erreur moyen des mots était de 3,3% et le taux d'erreur des caractères était de 0,69%.

III.6.2 La Segmentation basée sur la correspondance de modèles

La correspondance de modèle est une technique de traitement d'image numérique permettant de rechercher de petites parties d'une image qui correspondent à une image de modèle. La correspondance de modèles a été utilisée dans les applications de reconnaissance de caractères. Il est utilisé comme une technique pour reconnaître des caractères ou des mots avec une référence à une base de données stockée contenant un ensemble d'images pour les caractères ou les mots. Cette technique peut être considérée comme une technique de segmentation, mais les caractères ou les segments doivent être choisis manuellement et stockés pour être utilisés à des fins de comparaison. Dans [108] une technique a été proposée qui recherche l'apparition d'un angle formé par la jonction de deux caractères à la ligne de base. Utiliser une fenêtre 7 x 7 le voisinage des mots, est examiné pour décider de la segmentation. Bien que cette méthode ait réussi à obtenir un bon résultat, son succès à trouver

le bon angle dépend très fortement du bruit de l'image.

La correspondance de modèles n'est pas une technique appropriée pour la reconnaissance de caractères arabes écrite ou imprimée à la main en raison des styles différents d'écriture, en outre, il faut également du temps pour vérifier tous les modèles, en particulier avec trop de styles et de tailles de police. Par conséquent, un tel système est très lent et peu pratique.

III.7 Conclusion

Dans ce chapitre nous avons essayé d'exposer les différentes méthodes utilisées dans la segmentation de mot. Ces méthodes ont connu beaucoup de progrès ces derniers temps. Des techniques variées influencées par l'évolution dans les domaines tels que la reconnaissance de la parole et la reconnaissance en ligne des caractères ont émergés.

La segmentation de caractères est une tâche pénible due à la similarité visuelle de caractères aussi bien que sa forte dépendance avec la reconnaissance ; on ne peut pas reconnaître les caractères chevauchés sans faire leur segmentation aussi leur segmentation est difficile sans faire leur reconnaissance.

La difficulté dans la réalisation d'une segmentation performante dépend généralement de la nature du document à lire et de sa qualité. Le taux de mauvaise segmentation croit progressivement à partir de l'écriture imprimée à l'écriture manuscrite jusqu'à l'écriture manuscrite cursive où la difficulté devient plus importante. La performance d'un système de reconnaissance de l'écriture ne dépend pas seulement des résultats de la phase de segmentation, mais aussi du type de classifieur utilisé pour la reconnaissance de ces segments.

Chapitre **4**

*Contribution à la
Reconnaissance des
Montants Littéraires
Arabes*

IV.1 Introduction

Au cours des deux dernières décennies, les systèmes de reconnaissance des textes manuscrits arabes cursifs ont reçu une attention considérable dans de nombreuses applications réelles qui agrémentent notre quotidien telles que la conversion automatique d'images numérisées d'un texte imprimé en un texte pouvant être lu par un ordinateur, la lecture et l'authentification de chèques bancaires, le tri postal [143], la reconnaissance des plaques d'immatriculation et autres applications commerciales [144], [145].

Le système de reconnaissance des textes manuscrits arabes cursifs est souvent une tâche très difficile nécessitant des compétences spécialisées [145], [146]. De nombreux obstacles intimement liés à l'écriture ont eu un impact négatif sur les avancées dans ce domaine pour de multiples raisons telles que la grande variabilité du texte arabe manuscrit en termes de formes d'écriture, d'échelles et d'orientations, la nature cursive et la taille variable des différents styles d'écriture, l'ambiguïté existant entre certains scripts arabes [147], [148].

Pour toutes ces raisons, la maximisation du taux de reconnaissance nécessite des règles très spécifiques. De manière générale, le système de reconnaissance des textes manuscrits arabes cursifs est confronté à ses propres défis, la reconnaissance des montants littéraux en est un exemple. Sinon, il partage les mêmes obstacles mentionnés ci-dessus mais avec moins de difficultés. Ceci est dû au vocabulaire réduit utilisé dans les montants littéraux arabes [146] qui se compose d'environ 61 mots.

Récemment, la croissance des systèmes automatisés dans les pays du Moyen-Orient et principalement dans leurs secteurs financiers tels que les banques et les bureaux de poste, d'une part, l'énorme exigence des systèmes de reconnaissance automatique des chèques arabes, d'autre part, ont attiré l'attention de nombreux chercheurs de par le monde. En effet, une masse de près de 100 milliards de chèques circulent chaque année dans le monde et la plus grande partie de ces chèques est encore traitée manuellement par des opérateurs humains [149]. Différentes approches de reconnaissance ont donc été proposées pour améliorer les performances en termes de précision et de robustesse face aux changements d'écriture [150].

Il est à noter que l'amélioration du taux de reconnaissance dépend fortement des performances qui peuvent être enregistrées dans trois étapes principales, à savoir : le prétraitement, l'extraction de caractéristiques et la classification. Notre recherche se concentrera particulièrement sur l'étape d'extraction des caractéristiques et l'étape de classification. La différence fondamentale entre le système de reconnaissance proposé et ceux existant dans la littérature réside dans l'utilisation d'un modèle d'extraction de caractéristiques basé sur les PAWs sans segmentation en caractères.

En général la langue arabe se compose de caractères écrits horizontalement de droite à gauche. L'alphabet arabe dispose de 28 caractères différents. Parmi eux, dix caractères ont un point ; trois caractères possèdent deux points et seulement deux caractères en possèdent trois. La position des points peut être située soit en haut soit en bas du caractère. Autre caractéristique, les lettres arabes changent de forme en fonction de leur emplacement dans le mot. Qu'elles apparaissent au début, au milieu, à la fin d'un mot ou qu'elles soient isolées, les lettres arabes ne s'écrivent pas de la même manière [151]. Ainsi, elles peuvent être liées des deux côtés : à droite et à gauche ou bien seulement à droite ou bien seulement à gauche. Lors du contrôle des montants littéraux rédigés en arabe, de nombreuses ligatures comprennent l'empilement vertical des caractères. Cette propriété complique l'étape de segmentation dans les systèmes de reconnaissance conventionnels. Par conséquent, la segmentation d'un mot composé de lettres souffre de deux problèmes épineux : la sous-segmentation et la sur-segmentation [4][47][152][153]. Notre contribution principale dans ce domaine consiste donc à surmonter ces deux problèmes.

Dans ce sens, diverses approches ont été récemment proposées. Parmi elles, Alma'adeed et al. [151] ont utilisé des caractéristiques locales extraites à partir de 4700 images de mots écrits par 100 scripteurs différents et qui sont disponibles dans la base de données AHDB. La moitié de ces mots a été utilisée pour l'apprentissage des classificateurs de type Markov cachés (*Hidden Markov Model* : HMM) où le taux de reconnaissance a atteint 45%.

Dans une autre étude [4] publiée en 2004, le taux de reconnaissance a été amélioré pour atteindre 60% en regroupant des images de mots dans huit sous-groupes et en utilisant un classificateur à base de règles avec trois caractéristiques globales. Huit classificateurs HMM séparés ont été ainsi utilisés pour chaque sous-groupe. Snoussi-Maddouri et al. [154] ont utilisé une combinaison d'une modélisation de la vision globale et locale du mot. Les caractéristiques ont été extraites à travers 5000 images de mots écrits par 100 scripteurs différents. Un ensemble de 2500 mots ont été utilisés pour entraîner un classificateur neuronal transparent (*Transparent Neural Network* : TNN), le taux de reconnaissance a atteint 97%. Souici-Meslati et Sellami [155] ont utilisé des caractéristiques structurelles extraites de 1200 images de mots écrites par 25 scripteurs différents. Un ensemble de 480 mots écrits par 10 scripteurs a été utilisé pour entraîner un classificateur neuro-symbolique dans lequel un taux de reconnaissance de 93% a été obtenu. Farah et al. [156] ont utilisé des caractéristiques structurelles holistiques extraites de 4800 images de mots écrites par 100 scripteurs différents. Parmi cette base, un ensemble de 2400 mots a été utilisé pour l'apprentissage d'un système multi-classificateur basé sur trois classificateurs individuels : Perceptron multi-couche (*Multi*

Layred Perceptron : MLP), K-plus proches voisins, et K-plus proches voisins floue. Le taux de reconnaissance a atteint 89,17%. Les mêmes auteurs ont utilisé, dans un autre travail [7], des caractéristiques perceptuelles de haut niveau et des caractéristiques holistiques structurelles, ce qui a permis d'avoir un taux de reconnaissance égale à 94,93%. Boukharouba et Bennia [157] ont utilisé des caractéristiques structurelles et statistiques extraites de 7200 images de mots écrites par 50 scripteurs différents. L'étape de classification a été réalisée en combinant les classificateurs HMM et MLP, ce qui a donné un taux de reconnaissance de 97,20%. Al-Nuzaili et al. [146] ont combiné deux catégories de caractéristiques : composants et points, boucles et formes de caractères. Il s'agissait de concevoir un modèle amélioré d'extraction de caractéristiques perceptives structurelles (PFM). Ils ont utilisé pour cela la base de données AHDB pour valider leur modèle et ont obtenu un taux de reconnaissance de 92,13%.

Dans d'autres références telles que les travaux référencés dans [158] à [160] et les références qui s'y trouvent, l'extraction des caractéristiques (structurelle aussi bien que statistique) a été basée sur la segmentation des mots arabes. Malheureusement, toutes ces méthodes basées sur la segmentation recèlent un inconvénient majeur celui de la sur- ou sous-segmentation.

Dans ce présent travail, un modèle amélioré d'extraction des caractéristiques, basé sur l'analyse des PAWs (*Part of Arabic Word*), est d'abord introduit. Deuxièmement et au lieu d'utiliser un seul classificateur multi-classe, nous proposons de diviser le problème multi-classe en quatre sous-problèmes afin de les résoudre indépendamment les uns des autres en utilisant quatre ensembles séparés de SVM. Finalement, notre troisième proposition vise à résoudre le problème des PAW en contact qui se produit de manière courante dans l'écriture arabe.

En outre, nous introduisons également dans cette étude, un nouveau système de reconnaissance des montants littéraux arabes basés sur trois approches :

En premier lieu, nous proposons un modèle amélioré d'extraction des caractéristiques basé sur l'analyse des parties de mots arabes, PAWs. Il combine, à la fois, et les caractéristiques statistiques extraites du mot entier et les caractéristiques structurelles extraites de ses PAWs.

Deuxièmement, nous partitionnons le lexique utilisé en montants littéraux en quatre sous-groupes en fonction du nombre de PAWs. Chaque sous-groupe est indépendamment instruit par un ensemble de machines, à vecteurs de support (SVM), correspondant.

Troisièmement, nous proposons une étape de pré-classification permettant de résoudre le problème des PAWs en contact. La classification de chaque mot est ensuite simultanément effectuée avec et sans correction des PAWs en contact. Dans les deux cas, le mot candidat est

présenté à l'ensemble des SVM correspondants, où les deux scores de classification sont calculés. Ensuite, la décision finale est donnée en fonction du meilleur score.

Les résultats expérimentaux obtenus sur la base de données AHDB confirment l'efficacité et la robustesse du système de reconnaissance proposé en termes de taux de classification qui a atteint 95,91%.

IV.2 Principales étapes du système de reconnaissance de l'écriture

Les systèmes de reconnaissance des mots arabes manuscrits comportent trois étapes principales : le prétraitement, l'extraction de caractéristiques et la classification.

IV.2.1 Prétraitement

L'objectif principal attendu de l'étape de prétraitement est l'amélioration de la qualité de l'image d'un mot candidat, qui sera ensuite utilisé dans les étapes d'extraction des caractéristiques et de classification [154],[161][162]. Parmi les objectifs assignés à cette étape, nous citons un bon lissage, une suppression du bruit, une bonne binarisation afin de bien caractériser les mots. Dans ce travail, l'étape de prétraitement est réalisée en trois sous-étapes : binarisation, normalisation et extraction de contour.

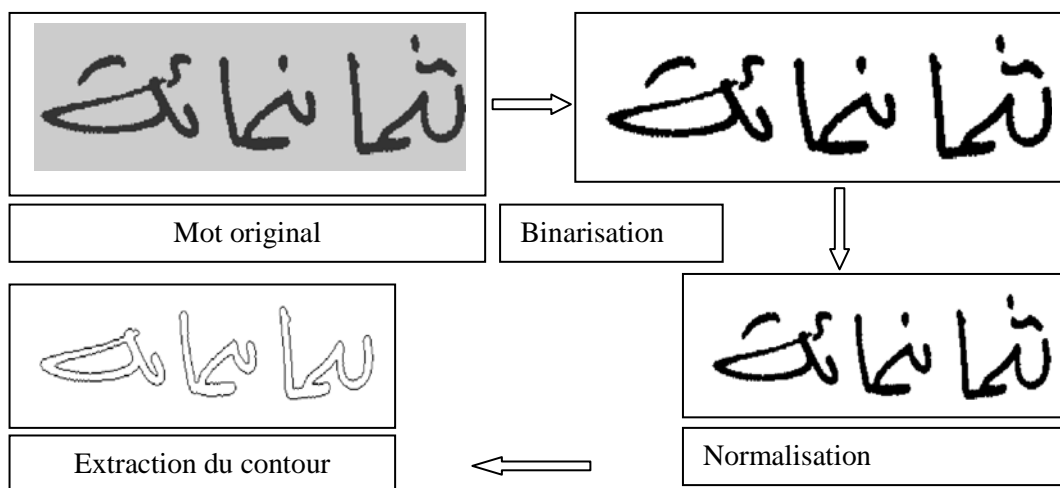


Figure IV.1 Différentes étapes de prétraitement

IV.2.2 Extraction des caractéristiques

Dans les systèmes de reconnaissance de l'écriture, les caractéristiques utilisées sont généralement de deux types : structurelles et statistiques. Les caractéristiques structurelles sont basées sur les propriétés topologiques et géométriques d'un mot candidat telles que les proportions, les points de croisement, les boucles, les points de branchement, les traits et leurs

directions, l’inflexion entre deux points, les courbes horizontales en haut ou en bas, ... etc. Les caractéristiques statistiques sont dérivées de la distribution des points tels que le zonage, les projections et les profils, les croisements et les distances, ... etc. Ces caractéristiques sont des mesures numériques, calculées sur des images ou des régions d’images. Ils incluent les densités de pixels, l’histogramme des directions de code en chaîne, les moments invariants et les descripteurs de Fourier[163-165].

En général, les caractéristiques structurelles et statistiques sont extraites après l’opération de segmentation, en lettres et/ou en graphèmes, des caractères d’un mot candidat. Ensuite, les vecteurs de caractéristiques concaténés, qui sont générés par l’un, l’autre ou les deux types précédents, sont utilisés comme entrée du classificateur. Malheureusement, les caractéristiques obtenues à partir des lettres segmentées ne sont pas assez performantes dans le cas des montants littéraux arabes manuscrits en raison des problèmes de sous-segmentation ou de sur-segmentation. Pour pallier à ces inconvénients, nous proposons l’application des méthodes d’extraction sur des PAWs isolés plutôt que sur des caractères segmentés. Dans cette étude, nous utilisons la base de donnée AHDB qui contient tous les mots possibles utilisés dans le contrôle des montants arabes, voir le **tableau IV.1**.



Figure IV.2: Aperçu sur la base de données AHDB

Tab IV.1 Vocabulaire des montants littéraux présents dans les chèques arabes (AHDB database)

N ^o de classe	Mots en Arabe	Code	N ^o de classe	Mots en Arabe	Code
1	احد	Ahad	32	لا	No
2	احدى	Ahda	33	واحد	One
3	ثمان	Eight	34	سبع	Seven
4	ثمانية	Eightb	35	سبعة	Sevenb
5	ثمانون	Eighty	36	سبعون	Seventy
6	ثمانين	Eightyb	37	سبعين	Seventyb
7	ثمانمئة	Eighth	38	سبعمئة	Sevenh
8	ثمانمئة	Eighthb	39	سبعمئة	Sevenhb
9	اثنى	Ethna	40	ست	Six
10	خمسون	Fifty	41	سنة	Sixb
11	خمسين	Fiftyb	42	ستون	Sixty
12	خمس	Five	43	ستين	Sixtyb
13	خمسة	Fiveb	44	ستمئة	Sixth
14	خمسمئة	Fiveh	45	ستمائة	Sixthb
15	خمسائة	Fivehb	46	عشر	Ten
16	اربع	Four	47	عشرة	Tenb
17	اربعة	Fourb	48	ثلاثون	Thirty
18	اربعون	Fourty	49	ثلاثين	Thirtyb
19	اربعين	Fourtyb	50	الف	Thousand
20	اربعمئة	Fourh	51	الاف	Thousands
21	اربعمائة	Fourhb	52	ثلاث	Three
22	غير	Gayr	53	ثلاثة	Threeb
23	مئة	Hundered	54	ثلاثمئة	Threeh
24	مائة	Hunderedb	55	ثلاثمائة	Threehb
25	مليون	Million	56	عشرون	Twenty
26	تسع	Nine	57	عشرين	Twentyb
27	تسعة	Nineb	58	اثنان	Two
28	تسعون	Ninety	59	اثنين	Twob
29	تسعمئة	Nineh	60	مئتين	Twoh
30	تسعمائة	Ninehb	61	مائتين	Twohb
31	تسعين	Ninetyb			

IV.2.2.1 Extraction des caractéristiques structurelles

Selon des travaux antérieurs [166], les méthodes d'extraction des caractéristiques structurelles sont couramment appliquées sur des caractères segmentés ou sur des PAW isolés. Dans ce travail, nous nous intéressons aux approches basées sur les PAWs car elles fournissent une segmentation naturelle des mots. En général, pour les montants littéraux arabes manuscrits, quatre caractéristiques structurelles peuvent être extraites : nombre de PAW, boucles, ascenseur et descendeur, diacritiques et points de position.

A. Nombre de PAW

Une PAW représente une partie isolée d'un mot et elles sont extraites à l'aide de l'algorithme CCL (*Connected-Component Labelling*). Plus de détails sur cet algorithme sont disponibles

dans les références [167] et [168]. Lors du contrôle des montants littéraux dans les chèques bancaires, le nombre de PAW isolés, dans tout mot de ce lexique spécifique, peut varier de un à quatre. La **figure IV.3** montre un exemple du mot «اربعون» qui contient ainsi quatre PAWs.

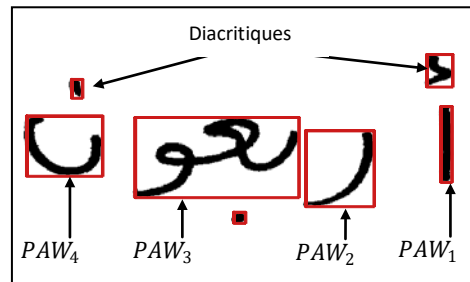


Figure IV.3 Nombre de PAWs extraits du mot «اربعون»

B. Les Boucles

Selon le **Tableau IV.1**, des boucles peuvent apparaître au début, au milieu et/ou à la fin des PAW isolés. Le nombre de boucles de chaque PAW peut être soit : aucune boucle (comme dans : «غير», «ست», «حد», etc.) ; une seule boucle (comme dans : «لف», «ستو», «ثما», etc.) deux boucles (comme dans : «مئة», «خمسو», «مليو», etc...) ou finalement trois boucles (comme dans («بعمنة», «سبعمنة», «خمسمئة», etc.)). De manière générale, le nombre de boucles est extrait en utilisant deux méthodes différentes. La première extrait tous les contours internes existants, puis détermine les objets fermés avant de décider du nombre de boucles, voir **figure IV.4** partie (a). Dans la deuxième méthode, l'image binaire du mot est inversée, de sorte que les pixels noirs deviennent blancs et inversement. Ensuite, l'algorithme CCL est utilisé pour l'extraction de tous les composants connectés et la suppression de ceux en contact avec la bordure de l'image, autrement dit l'arrière-plan. La somme des composants restants définit le nombre de boucles, voir **figure IV.4** partie (b). Dans ce présent travail, nous employons la seconde méthode. Par conséquent, nous obtenons un vecteur de caractéristiques de boucle de taille égale à 2 pour chaque PAW.

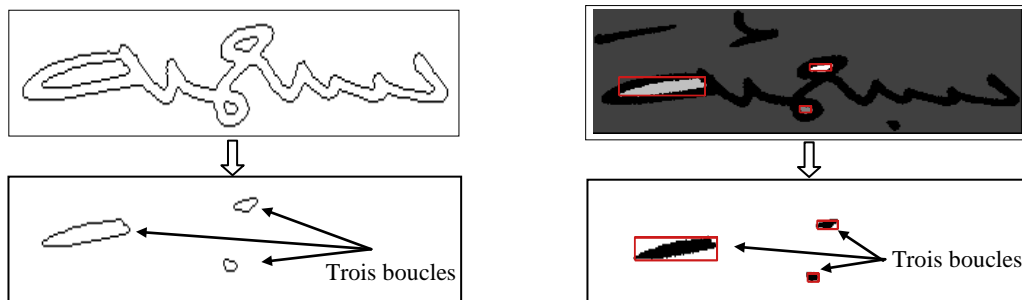


Figure. IV.4: Extraction des boucles; (a) Méthode basée sur les contours intérieurs. (b) Méthode basée sur l'algorithme CCL

C. Localisation des hampes et des jambages

En général, les performances des méthodes d'extraction des caractéristiques telles que les hampes et les jambages dépendent fortement de la qualité de détection de la ligne de base et de la zone médiane du mot. Pour le cas des chèques bancaires, les montants littéraux manuscrits sont écrits sur une ligne droite explicite, qui est préalablement pré-imprimée. Ceci simplifie la détection de la ligne de base dans les chèques bancaires. La projection horizontale basée sur la méthode de l'histogramme [169] est couramment utilisée pour découvrir les propriétés d'histogramme importantes telles que le pic central correspondant au contenu principal du mot ainsi que de la ligne de base de l'image. Deux minimums locaux apparaissent en haut et en bas de la ligne de base. Au niveau supérieur, il est déterminé les caractéristiques des hampes et au niveau inférieur les caractéristiques des jambages. La zone médiane se situe entre les lignes de base supérieure et inférieure (voir la [figure IV.5](#)). Par conséquent, nous obtenons un vecteur de caractéristiques hampes et jambages de taille égale à 2 pour chaque PAW.

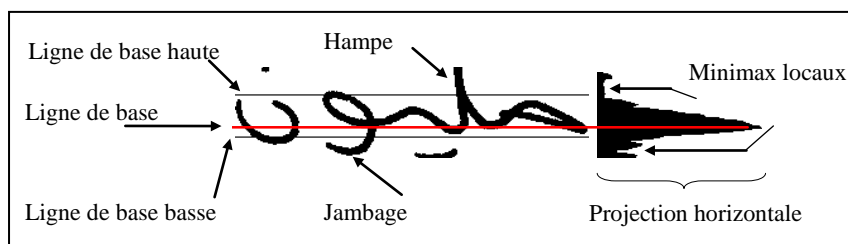


Figure IV.5 Positionnement des caractères par rapport aux lignes de base

D. Points Diacritiques

En général, les mots arabes contiennent plusieurs types de points diacritiques [69]. Selon le [tableau IV.1](#), près de 92% des PAW isolés ont de un à trois points diacritiques situés en haut ou en bas de la ligne de base.

L'écriture de ces points (simple ou multiple) est généralement respectée par les scribes, car ces points permettent de distinguer entre eux les caractères ayant le même corps principal. Les points diacritiques sont en dehors de la zone d'information principale (partie primaire du caractère), ils constituent donc les parties secondaires des caractères. Les points diacritiques sont de taille assez faible, ce que les rend très sensible aux bruits d'acquisition. Un point simple est défini comme étant la plus petite entité connexe qui peut être considérée comme information significative et non pas comme bruit. Les points multiples, par contre,

sont de forme plus complexe, car ils constituent un regroupement de points d'un même caractère. Ce regroupement est essentiellement dû au style d'écriture. Combien même les points multiples possèdent des formes assez simples dans le cas de l'écriture imprimée, ils restent inséparables dans le cas du manuscrit. Les points multiples sont de deux types : les points doubles et les points triples.

La position des points diacritiques est déterminée de la manière suivante comme cela est illustré dans les étapes suivantes. D'abord, tous les composants connectés sont extraits de chaque mot candidat. Deuxièmement, nous considérons la largeur de la zone médiane de chaque mot comme valeur de seuil de référence. Troisièmement, le seuil obtenu est comparé à chaque composant extrait pour détecter tous les points diacritiques. Finalement, la position des points diacritiques est déterminée par rapport à la ligne de base (voir [figure IV.6](#)). Les positions des points sont généralement soit en haut soit en bas de la ligne de base. Le processus de détermination de leur position s'effectue comme suit :

Tout d'abord, la coordonnée de la ligne de base (X) est extraite. Deuxièmement, les coordonnées de la zone délimitant les points sont calculées, donnant par conséquent et pour chaque point, X_{min} , Y_{min} , X_{max} et Y_{max} . Troisièmement, les coordonnées X_{min} et X_{max} sont comparées à la coordonnée de la ligne de base (X). Si $X_{min} > X$ alors, le point est considéré se situer en haut. Si $X_{max} < X$, alors le point est considéré se situer en bas.

Dans le but de déterminer le nombre des points diacritiques, nous avons utilisé des heuristiques basées sur l'épaisseur du trait et proposées par Ameur et al. [170]. L'algorithme utilisé est présenté dans la [figure IV.7](#).

Sachant que S est l'épaisseur du trait, X_{min} , Y_{min} , X_{max} et Y_{max} sont les coordonnées délimitant la composante considérée. En conséquence, nous obtenons un vecteur de caractéristiques diacritiques de taille égale à 7 pour chaque PAW.

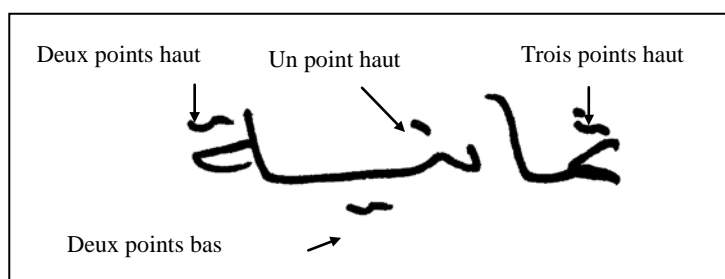


Figure IV.7 Nombre de points diacritiques et leur position dans le mot «ثمانية - Huit»

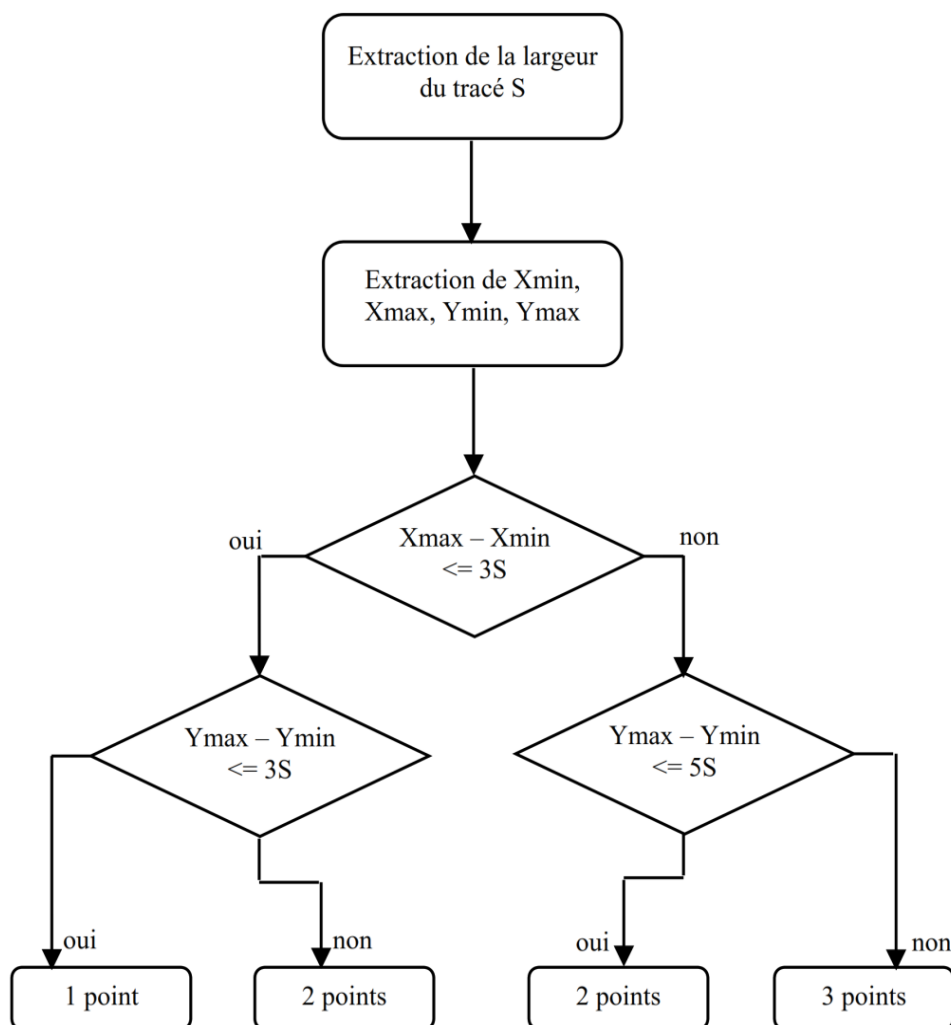


Figure IV.7 Extraction des points diacritiques

IV.2.2.2 Extraction de caractéristiques statistiques

Les caractéristiques statistiques sont extraites à partir du mot entier, qui doit être donné sans points diacritiques. Dans ce travail, quatre types de caractéristiques statistiques sont déterminés : les histogrammes de code de chaîne locale (Local Chain Code Histograms, CCH), appelés aussi les codes de chaîne Freeman, le zonage, les invariants de moments de Zernike (Zernike Moment Invariants, ZMI) et les histogrammes de profils de densité (Density Profile Histograms, DPH).

A. Histogrammes locaux de chaîne de code

L'extraction des caractéristiques CCH locales est effectuée en utilisant les étapes suivantes. D'abord, l'algorithme de suivi du contour (FCA) est appliqué sur l'image du mot afin de déterminer son contour correspondant, par lequel une séquence de pixels adjacente ordonnée est générée. Deuxièmement, le contour obtenu est partitionné uniformément en quatre zones $\{z_i\}_{i=1,2,\dots,4}$. Troisièmement, pour chaque zone, les huit directions de Freeman

$\{d_j\}_{j=0,1,\dots,7}$ sont utilisées pour coder séparément les points de contour donnés. L'histogramme normalisé $h_{(z_i,j)}$ est ensuite calculé dans chaque direction en utilisant l'équation suivante : [171]

$$h_{(z_i,j)} = \frac{n_j}{N_{z_i}} \tag{IV.1}$$

où n_j désigne le nombre total de code de chaîne de Freeman dans la direction j , et N_{z_i} indique le nombre total de pixels noirs qui apparaissent dans chaque zone z_i . Les histogrammes normalisés obtenus constituent un vecteur de caractéristiques CCH (CCH_{Vect}) de taille égale à 32 pour chaque mot (voir la figure IV.8. Par conséquent, nous obtenons :

$$CCH_{Vect} = [\{h_{(z_1,0)}, \dots, h_{(z_4,7)}\}, \{h_{(z_2,0)}, \dots, h_{(z_2,7)}\}, \{h_{(z_3,0)}, \dots, h_{(z_3,7)}\}, \{h_{(z_4,0)}, \dots, h_{(z_4,7)}\}]$$

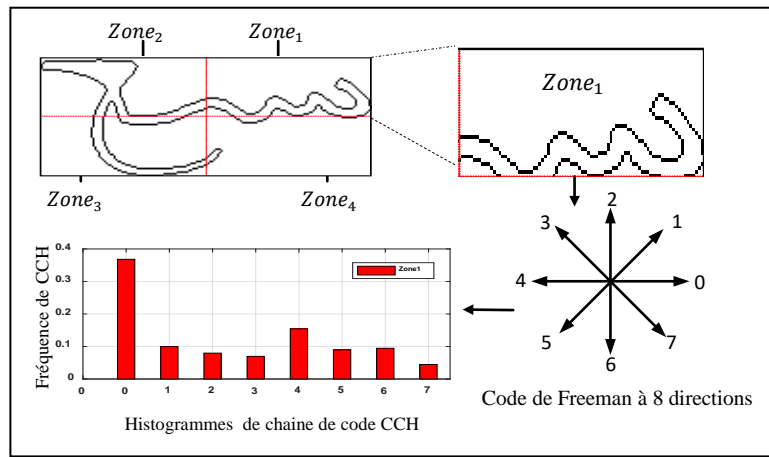


Figure IV.8 Histogrammes locaux de chaîne de code du mot « سبع - sept »

B. Zonage

Les caractéristiques de zonage sont obtenues en appliquant la méthode de zonage sur l'image du mot entier. Cette méthode divise l'image candidate en plusieurs régions en utilisant les grilles régulières. Pour chaque zone, la densité des pixels noirs est calculée puis normalisée [172]. Dans ce travail, trois types de zonage sont utilisés. Le premier est basé sur la méthode de zonage rectangulaire régulier pour fournir 9 images de sous-bloc identiques $\{Z_{rr_i}\}_{i=1,\dots,9}$, voir figure IV.9 partie (a). Le deuxième type est basé sur la méthode de zonage rectangulaire irrégulier pour fournir 3 images de sous-bloc identiques $\{Z_{ir_i}\}_{i=1,2,3}$, voir figure IV.9 partie (b). Finalement, le dernier type est basé sur la méthode de zonage triangulaire uniforme pour fournir 9 images de sous-bloc $\{Z_{rt_i}\}_{i=1,2,\dots,9}$, voir figure IV.9 partie (c) . Par conséquent, nous obtenons le vecteur de caractéristique de zonage (Z_{FV}) de taille égale à 20 pour chaque mot entier :

$$Z_{FV} = [\{Z_{rr_1}, \dots, Z_{rr_8}\}, \{Z_{ir_1}, \dots, Z_{ir_3}\}, \{Z_{rt_1}, \dots, Z_{rt_9}\}],$$

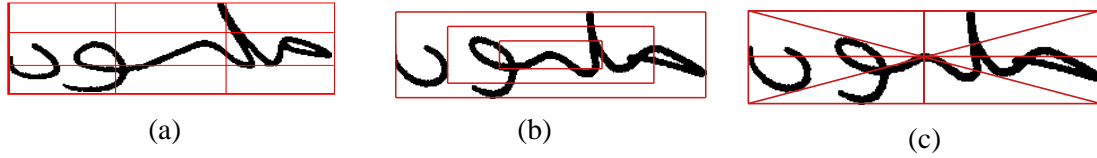


Figure IV.9 Topologies des méthodes de zonage: (a) Méthode de zonage rectangulaire régulier (b) Méthode de zonage chevauchée basée sur une grille régulière 3x3. (c) Méthode de zonage triangulaire uniforme

C. Moments Invariants de Zernike

Les ZMI (Zernike Moment Invariants) sont généralement utilisés dans la reconnaissance des formes et la reconstruction de formes d'images [42]. La forme discrète de la méthode des moments de Zernike est appliquée sur une image de mot numérique de taille $N \times M$. En conséquence, les moments de Zernike, d'ordre n avec une répétition m , sont exprimés comme suit:

$$Z_{er_{n,m}} = \frac{n+1}{\pi} \sum_{x=1}^N \sum_{y=1}^M I(x,y) \cdot V_{n,m}(x,y) \quad (IV.2)$$

où $|m| \leq n$ et $n \geq 0$, $I(x,y)$ est l'intensité du pixel avec les coordonnées (x,y) dans l'image binaire originale, à partir de laquelle les moments de Zernike sont déterminés. $V_{n,m}(x,y)$ désigne les polynômes de Zernike donnés en coordonnées cartésiennes. Ils peuvent être convertis en coordonnées polaires en utilisant les fonctions trigonométriques Sinus et Cosinus, où la coordonnée radiale r est déterminée par $r = \sqrt{x^2 + y^2}$ avec $r > 0$, et la coordonnée angulaire θ est déterminée par $\theta = \tan^{-1}\left(\frac{y}{x}\right)$ avec $-\pi \leq \theta \leq \pi$. Par conséquent, la fonction polynomiale de Zernike $V_{n,m}(r,\theta)$ est réécrite en coordonnées polaires comme suit :

$$V_{n,m}(r,\theta) = R_{n,m}(r)e^{-jm\theta} \quad (IV.3)$$

où $j = \sqrt{-1}$ est l'unité imaginaire, $R_{n,m}(r)$ est le polynôme radial orthogonal qui s'exprime ainsi :

$$R_{n,m}(r) = \sum_{k=0}^{(n-|m|)/2} \frac{(-1)^k (n-k)! r^{n-2k}}{k! ((n+|m|)/2 - k)! ((n-|m|)/2 - k)!} \quad (IV.4)$$

Dans cette étude, nous considérons les moments de Zernike d'ordre $n = 1$ avec des répétitions $m = 10$. Par conséquent, nous obtenons un vecteur de caractéristiques ZMI (ZMI_{vect}) de taille égale à 10 pour chaque mot entier qui s'exprime comme suit:

$$ZMI_{vect} = [Z_{er_{1,1}}, \dots, Z_{er_{1,10}}]$$

D. Histogrammes de la densité de profil

Les histogrammes de profil de densité sont basés sur la détermination de huit profils $\{P_{r_i}\}_{i=1,\dots,8}$ à partir de l'image du mot candidat sans points diacritiques. Ces profils peuvent être obtenus en balayant le mot candidat sur les huit différentes positions: gauche, droite, haut, bas et quatre positions correspondants aux angles d'orientation 45° , 135° , 225° et 315° (voir **figure IV.10a**). Les distances entre la zone d'image et les premiers pixels noirs sont calculées pour chaque profil. Ensuite, ils sont normalisés par rapport au nombre total de pixels dans les huit régions. Nous obtenons alors un vecteur de caractéristiques DPH (DPH_{vect}) de taille égale à 8 pour chaque mot entier par :

$$DPH_{vect} = [P_{r_1}, \dots, P_{r_8}],$$

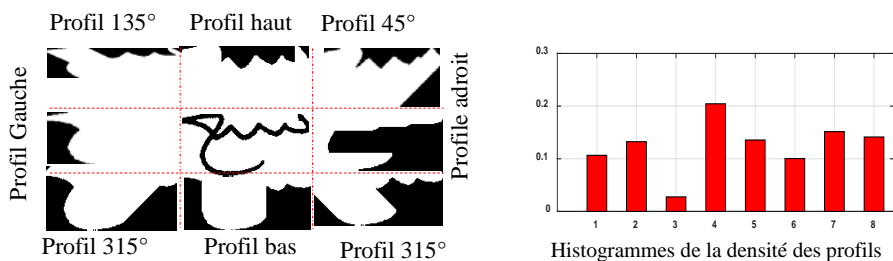


Figure IV.10 la densité de profil a) Les huit projections des profils b) Histogrammes de la densité des profils

IV.2.3 Etape de classification

Les machines à vecteurs de support (*Support Vector Machine*, SVM) sont couramment utilisées en tant que classificateurs dans les systèmes de reconnaissance de l'écriture manuscrite.

Ces classificateurs, essentiellement utilisés dans les problèmes binaires de classification, sont basés sur la détermination d'un hyper-plan optimal qui sépare autant que possible les échantillons de deux classes proches. La frontière de décision, séparant ces deux classes, est définie par un ensemble d'échantillons d'apprentissage appelés vecteurs de support.

Lorsque les échantillons ne sont pas linéairement séparables, les SVMs linéaires fournissent des performances insatisfaisantes. Ce problème pourrait être résolu en transférant chaque point de donnée dans un nouvel espace de caractéristiques dans lequel les données seront dispersées et éventuellement plus séparables.

Dans un classificateur SVM non linéaire, trois fonctions noyau sont couramment utilisées.

Ces fonctions sont:

Le noyau linéaire : (*Equation .I. 11*),

Le noyau polynomial : (*Equation .I.12*)

La fonction à base radiale : (*Equation .I.13*).

où x_i et x_j sont deux points de données et σ est l'écart type de la fonction à base radiale.

La résolution d'un problème multi-classes à l'aide de SVM est généralement effectuée en divisant l'ensemble du problème en sous-problèmes plus faciles à traiter. Il s'agit ensuite de résoudre chaque sous-problème avec un SVM binaire, et de combiner finalement toutes leurs réponses.

Les stratégies les plus communes de décomposition des problèmes multi-classes sont de types un-contre-tous (OvA, *One-versus-All*) et un-contre-un (OvO, *One-versus-One*). Dans la première stratégie OvA, chaque sous-problème consiste à classer une classe contre toutes les autres classes. Tandis que dans la deuxième stratégie, OAO, chaque sous-problème consiste à classer une classe contre une autre classe. Par conséquent, pour un problème avec K classes, OvA produit K sous-problèmes et OvO produit $(K(K - 1)) / 2$ sous-problèmes. Dans ce travail, nous utilisons l'approche OvA car elle génère moins de sous-problèmes.

IV.3 La méthode proposée:

Dans l'étape d'extraction des caractéristiques, nous proposons d'extraire les caractéristiques structurelles des PAW isolées et de les combiner avec les caractéristiques statistiques extraites directement du mot entier. Notre approche est motivée par les raisons suivantes :

- De manière générale, la façon d'écrire les PAW isolées fournit une segmentation naturelle du mot candidat. Par conséquent, l'extraction des caractéristiques structurelles des PAW isolées conduit à de meilleurs résultats que l'extraction directe à partir du mot entier.
- Les caractéristiques structurelles extraites des PAW isolées sont plus appropriées pour résoudre le problème des PAWs connectées.
- Dans certains cas, l'extraction des caractéristiques structurelles produit des résultats similaires pour différents mots candidats. Cela est dû à certaines erreurs comme la boucle manquante, les points diacritiques manquants, etc. Pour résoudre ce problème, nous combinons les caractéristiques structurelles et statistiques.

Le vecteur des caractéristiques globales est établi à travers cette combinaison pour chaque mot candidat.

Le [tableau IV.2](#) résume les différentes significations utilisées pour coder les composants du vecteur des caractéristiques concaténées.

Chaque vecteur des caractéristiques globales contient deux parties : str_vec : le vecteur de caractéristique structurelle et sta_vec : le vecteur de caractéristique statistique. Ces deux vecteurs sont donnés par:

$$str_vec = \left[O_L \mid T_L \mid As \mid Ds \mid O_DU \mid T_DU \mid Th_DU \mid O_DD \mid T_DD \mid 2T_DU \mid 2T_DD \right]$$

$$sta_vec = \left[CCH_{vect} \mid Z_{vect} \mid ZMI_{vect} \mid DPH_{vect} \right]$$

Le vecteur de caractéristique structurelle est constitué de nombres booléens (c'est-à-dire que ce nombre n'est constitué que de "1" ou "0" qu'on appelle bits) indiquant la présence ou l'absence de composants structurels. Ce vecteur est constitué de 11 bits par PAW. Parallèlement, le vecteur des caractéristiques statistiques est constitué de nombres réels indiquant des valeurs probabilistes des composantes statistiques. Il est constitué de 70 chiffres par mot.

La résolution d'un problème de classification avec K classes en utilisant la stratégie OvA implique K classificateurs. Chacun d'eux est entraîné pour classer une classe contre toutes les autres classes. Par conséquent, ces classificateurs sont entraînés en utilisant les échantillons de la classe correspondante comme échantillons positifs et tous les échantillons restants comme échantillons négatifs.

Cela conduit à des données d'apprentissage déséquilibrées. Pour pallier à cet inconvénient, nous proposons de diviser l'ensemble des données en quatre sous-groupes en fonction du nombre de PAW, puis d'utiliser 4 ensembles de SVM, chacun correspondant à un sous-groupe. En effet, le nombre de PAW isolées constituant tout mot candidat varie de 1 à 4. Le [tableau IV.3](#) résume les différents sous-groupes obtenus en fonction du nombre de PAW isolées.

En outre, notre idée vise à entraîner chaque ensemble de SVMs par différents espaces caractéristiques, ce qui permet de diversifier les données d'apprentissage. En effet, dans la stratégie OvA , tous les classificateurs partagent le même espace caractéristique, tandis que chaque ensemble de classes nécessite un espace différent. Dans ce travail, nous proposons d'utiliser un espace caractéristique différent pour chaque sous-groupe de classes composé du même nombre de PAW isolées. Le [tableau IV.4](#) donne le vecteur des caractéristiques globales de chaque classe $\{Classe\}_{SG_i}^j$ présent dans le sous-groupe correspondant SG_i . Où $[P_{i,k}]$ désigne le vecteur des caractéristiques structurelles correspondant à la $k^{ième}$ PAW isolée du $i^{ième}$ sous-groupe, $[Word_{SG_i}]$ désigne le vecteur des caractéristiques statistiques correspondant au $i^{ième}$ sous-groupe.

Tab IV.2 Différentes significations utilisées pour coder les primitives existantes.

Types de primitives	Primitives	Description	Notation	Nombre de primitives
Caractéristiques structurelles	Boucles	Une Boucle	<i>O_L</i>	2
		Deux Boucles	<i>T_L</i>	
	Hampe /Jambage	Hampe	<i>As</i>	2
		Jambage	<i>Ds</i>	
	Points diacritiques	Un point haut	<i>O_DU</i>	7
		Deux points haut	<i>T_DU</i>	
		Trois points haut	<i>Th_DU</i>	
		Un point bas	<i>O_DD</i>	
		deux points bas	<i>T_DD</i>	
		2 deux points haut	<i>2T_DU</i>	
	2 deux points bas	<i>2T_DD</i>		
<i>Le nombre Total du primitives structurelles par PAW</i>				11
Caractéristiques statistiques	Histogramme locale de Chaîne Codes		CCH	32
	Zonage		Z	20
	Moment Invariant de Zernike		ZMI	10
	Histogramme de Densite des Profils		DFH	8
	<i>Nombre Total du primitives Statistiques par mot</i>			

Tab IV.3 Différents sous-groupes obtenus en fonction du nombre de PAWs isolées

	Nombre de PAWs	Nombre de classes	Mots Candidats
Sous-groupe1 <i>SG₁</i>	1	21	ست, ستة, ستين, عشر, مئتين, سبعين, سبعة, سبع, سبعمئة, تسع, تسع, خمسين, خمس, خمسة, غير, مئة, خمسمئة, تسعمئة, ستمئة, لا, تسعين
Sous-groupe1 <i>SG₂</i>	2	25	خمسون, احد, ثمان, ثمانية, ثمانين, اثنى, تسعمائة, ستون, عشرة, مليون, تسعون, خمسمائة, مائة, سبعمائة, ثلاث, ثلاثة, عشرين, سبعون, ثلاثين, الف, ثمانمئة, ستمائة, ثلاثمئة, اثنين, مائتين
Sous-groupe1 <i>SG₃</i>	3	13	احدى, ثمانون, اربع, اربعة, اربعين, واحد, ثلاثون, اربعمئة, ثلاثمئة, ثمانمئة, الاف, عشرون, اثنان
Sous-groupe1 <i>SG₄</i>	4	2	اربعمائة, اربعون

Tab IV.4 Différents vecteurs de caractéristiques globales utilisés pour alimenter les classificateurs proposés.

Sous-groupes	Caractéristiques concaténées pour chaque mot
SG_1	$F_{SG_1} = \left[\begin{array}{c c} \overleftarrow{\text{str_vec (11)}} & \overleftarrow{\text{sta_vec (70)}} \\ \hline [P_{1,1}] & [Word_{SG_1}] \end{array} \right]$
SG_2	$F_{SG_2} = \left[\begin{array}{c c c} \overleftarrow{\text{str_vec (22)}} & & \overleftarrow{\text{sta_vec (70)}} \\ \hline [P_{2,2}] & [P_{2,1}] & [Word_{SG_2}] \end{array} \right]$
SG_3	$F_{SG_3} = \left[\begin{array}{c c c c} \overleftarrow{\text{str_vec (33)}} & & & \overleftarrow{\text{sta_vec (70)}} \\ \hline [P_{3,3}] & [P_{3,2}] & [P_{3,1}] & [Word_{SG_3}] \end{array} \right]$
SG_4	$F_{SG_4} = \left[\begin{array}{c c c c c} \overleftarrow{\text{str_vec (44)}} & & & & \overleftarrow{\text{sta_vec (70)}} \\ \hline [P_{4,4}] & [P_{4,3}] & [P_{4,2}] & [P_{4,1}] & [Word_{SG_4}] \end{array} \right]$

IV.4 Étape de reconnaissance basée sur la résolution du problème de PAWs connectées

Dans les montants littéraux manuscrits, la présence de PAWs connectées dans certains mots complique la tâche de reconnaissance. En effet, cela conduit à un nombre erroné de PAW et à une orientation vers un ensemble inapproprié de SVM. Pour surmonter ce problème, une étape de pré-classification est proposée.

La **figure IV.11** montre un exemple de deux mots inappropriés «عشرون» et «اربعون». Le mot de la partie (a) contient 3 PAWs au lieu de deux PAWs. $P_{3,1}$ touche $P_{3,2}$ donnant, par conséquent, un nouveau PAW ($P_{3,2} \cup P_{3,1}$). Le mot de la partie (b) contient 4 PAWs au lieu de trois PAWs. $P_{4,3}$ touche $P_{4,4}$ donnant, donc un PAW ($P_{4,4} \cup P_{4,3}$).

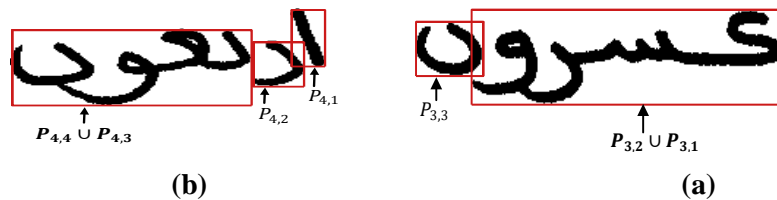


Figure IV.11. Exemple de deux mots inappropriés «عشرون-Vingt» composés de deux PAWs connectés

Le principe de l’algorithme de pré-classification proposé est basé sur les suppositions suivantes :

- Le problème des PAW chevauchées ne se pose pas dans le sous-groupe SG_1 .
- Tous les PAW chevauchées possibles sont résumées dans le **Tableau IV.5**.

Selon le [tableau IV.5](#), trois types possibles de mots incorrects peuvent apparaître dans les deux sous-groupes SG_1 et SG_3 , et cinq dans le sous-groupe SG_2 . Par conséquent, seuls ces sous-groupes sont pris en compte dans l'algorithme proposé de pré-classification. Les principales étapes de l'algorithme proposé sont résumées dans ce qui suit :

- Extraction des caractéristiques structurelles de tous les échantillons d'apprentissage, puis calcul du centre de chaque classe dans tous les sous-groupes. Ces centres seront considérés comme vecteurs de référence.
- Extraction des caractéristiques structurelles de tous les mots incorrects possibles selon le [tableau IV.5](#). Elles seront stockées et considérées comme une base de données de référence.

Tab IV.5 PAWs correctes et incorrectes possibles dans chaque sous-groupe

	PAWs Correctes	PAWs Incorrectes
Sous-groupe SG_1	$[P_{1,1}]$	$[P_{2,2} \cup P_{2,1}]$ $[P_{3,3} \cup P_{3,2} \cup P_{3,1}]$ $[P_{4,4} \cup P_{4,3} \cup P_{4,2} \cup P_{4,1}]$
Sous-groupe SG_2	$[P_{2,2}; P_{2,1}]$	$[(P_{3,3} \cup P_{3,2}); P_{3,1}]$ $[P_{3,3}; (P_{3,2} \cup P_{3,1})]$ $[(P_{4,4} \cup P_{4,3} \cup P_{4,2}); P_{4,1}]$ $[(P_{4,4} \cup P_{4,3}); (P_{4,2} \cup P_{4,1})]$ $[P_{4,4}; (P_{4,3} \cup P_{4,2} \cup P_{4,1})]$
Sous-groupe SG_3	$[P_{3,3}; P_{3,2}; P_{3,1}]$	$[P_{4,4}; P_{4,3}; (P_{4,2} \cup P_{4,1})]$ $[P_{4,4}; (P_{4,3} \cup P_{4,2}); P_{4,1}]$ $[(P_{4,4} \cup P_{4,3}); P_{4,2}; P_{4,1}]$
Sous-groupe SG_4	$[P_{4,4}; P_{4,3}; P_{4,2}; P_{4,1}]$	

Dans l'étape de reconnaissance, il est nécessaire de définir la manière d'orienter le vecteur des caractéristiques globales d'un nouveau mot candidat vers l'ensemble correct de SVM. Pour résoudre ce problème, nous proposons de classifier simultanément ce mot à l'aide de deux schémas (voir [Figure. IV.12](#)).

Le premier schéma est basé sur le système de reconnaissance standard, dans lequel le mot est considéré comme correct. Ce schéma donne le premier score de reconnaissance.

Le deuxième schéma est basé sur le système de reconnaissance proposé qui résout le problème de PAWs en contact. En se basant sur le nombre de PAWs, le vecteur de caractéristiques structurales du mot candidat est comparé à tous les vecteurs de référence possibles selon le **Tableau IV.5**. Un seul vecteur des caractéristiques structurales de référence est choisi fournissant ainsi la correction désirée pour chaque PAW. Par la suite, le vecteur des caractéristiques statistiques du mot incorrect est rajouté pour générer le vecteur global des caractéristiques. Il est enfin présenté à l'ensemble de SVM correspondant. Cela donne le deuxième score de reconnaissance. La décision finale est prise par la comparaison des deux scores de reconnaissance entre eux et le choix du meilleur d'entre les deux.

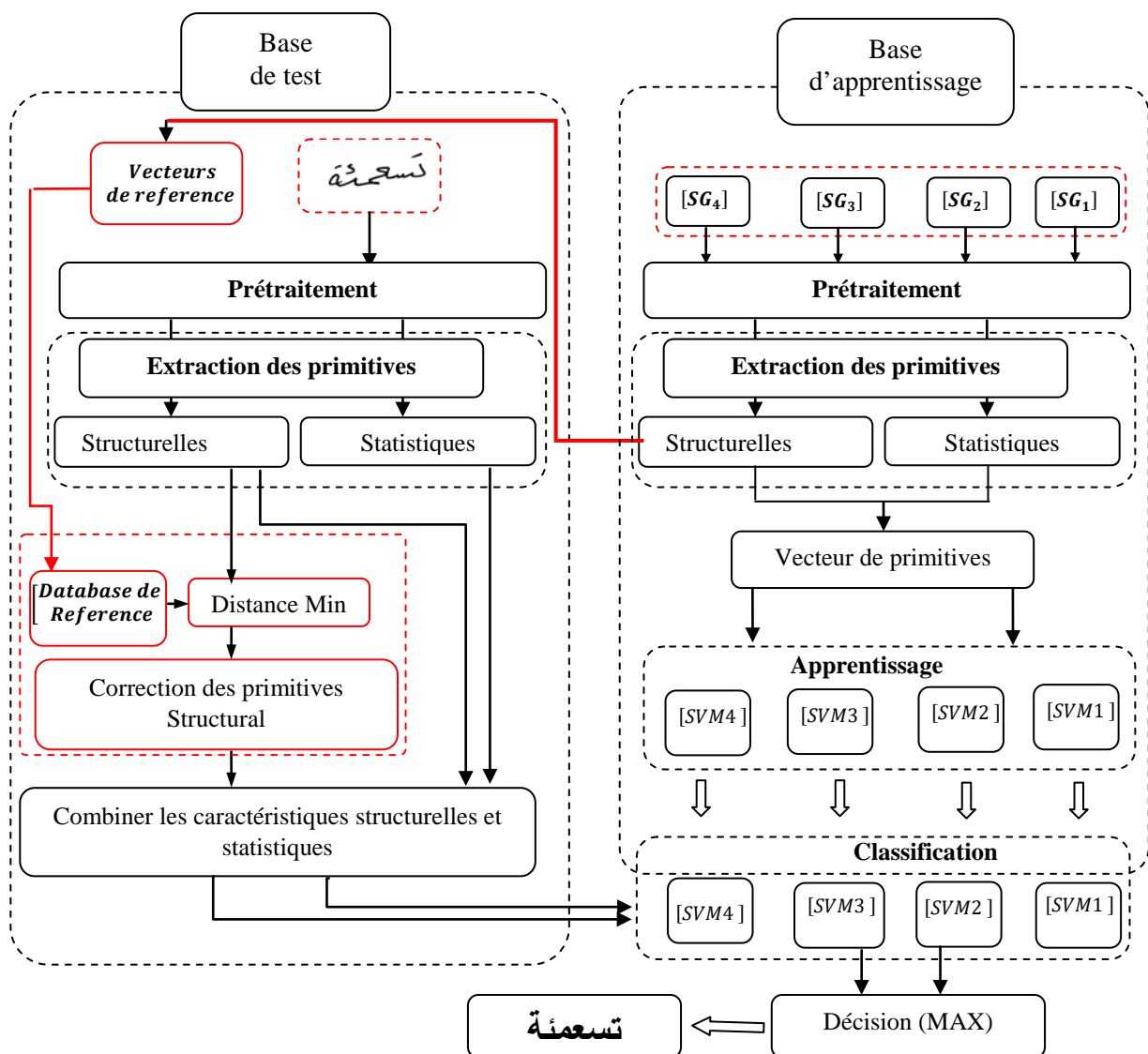


Figure IV.12 Système de reconnaissance proposé [174]

IV.5 Phase de Tests

Dans cette section, nous évaluons la méthode proposée en utilisant la base de données AHDB. Cette dernière se compose de 61 images de mots ; chaque mot est réécrit par 100 scripteurs différents, ce qui donne un total de 6405 échantillons.

Pour évaluer les performances de généralisation de la méthode proposée, nous utilisons une validation croisée d'ordre deux. Les données d'apprentissage sont alors divisées en deux parties ; l'une est utilisée pour l'apprentissage et l'autre pour les tests. Cette procédure est répétée deux fois. C'est-à-dire que, dans l'AHDB, 3204 échantillons sont utilisés pour l'apprentissage et 3201 pour les tests.

Pour chaque sous-groupe, les caractéristiques structurelles et statistiques sont combinées ce qui permet d'obtenir quatre ensembles de vecteurs de caractéristiques globaux. Les vecteurs résultants ont des tailles accrues, à savoir : 81 pour SG1, 92 pour SG2, 103 pour SG3 et 114 pour SG4. De plus, une base de données de référence contenant des vecteurs centraux de toutes les classes est établie à travers l'ensemble des échantillons d'apprentissage. La dimension de chaque vecteur central correspond au nombre de PAWs dans le sous-groupe correspondant.

De même, dans l'étape de classification proposée, 4 ensembles de SVM sont utilisés indépendamment pour chaque sous-groupe. Par conséquent, nous divisons d'abord les échantillons d'apprentissage en 4 sous-groupes, puis nous les utilisons pour l'apprentissage de l'ensemble correspondant de SVM. La décomposition de chaque sous-groupe est basée sur la stratégie *OvA*. Les nombres de SVM sont : 21 SVM pour le premier ensemble, 25 pour le deuxième, 13 pour le troisième et 2 pour le quatrième. Le [tableau IV.6](#) résume en détail les résultats obtenus pour les 4 sous-groupes.

Le [tableau IV.6](#) résume, quant à lui, les résultats obtenus avec et sans division en sous-groupes, et avec et sans correction des PAW en contact.

La [figure IV.12](#) illustre, respectivement, les taux de classification pour les premier, deuxième, troisième et quatrième sous-groupes. Pour chaque sous-groupe, le taux de classification est présenté dans deux cas : avec et sans correction des PAW en contact. Nous pouvons remarquer que, sans correction des PAW en contact, le taux de classification est supérieur à 90% pour tous les mots du 1^{er} sous-groupe ; et il est inférieur à 90% pour certains mots des autres sous-groupes. Cela inclut cinq mots dans le 2^{ème} sous-groupe qui sont: ثلاث, ثمانية, ثلاثمة, ثمانون, خمسون ; trois mots dans le 3^{ème} sous-groupe qui sont: احدى, ثمانون, اربع et finalement deux mots du 4^{ème} sous-groupe. On peut voir clairement qu'en utilisant la méthode proposée, la classification de ces mots est généralement améliorée.

Certains autres mots ont aussi un taux de classification inférieur à 90%. Ce résultat peut être lié à d'autres problèmes plutôt qu'au problème de contact. Par exemple, certaines paires de mots tels que ثلاثة / ثلاث, خمسون / تسعون, ستمائة / تسعمائة sont souvent confus. Cela est dû à la nature de l'écriture arabe, dans laquelle les scripteurs peuvent avoir la "mauvaise habitude" d'utiliser des caractères similaires pour différentes lettres.

Le [tableau IV.8](#) rassemble les meilleurs taux de reconnaissance obtenus par certains auteurs dans des travaux récents sur la base de données AHDB. Il est très clair que l'approche proposée surpasse les autres travaux.

Tab IV.7 Résultats complets de la classification.

	Sans division	Avec division (basée sur le nombre de PAWs)	
		Sans correction de PAWs connectées	Avec correction de PAWs connectées
	Taux de reconnaissance	Taux de reconnaissance	Taux de reconnaissance
Sous-groupe 1	-	95.10	95.05
Sous-groupe 2	-	92.46	94.98
Sous-groupe 3	-	91.50	96.48
Sous-groupe 4	-	80.95	97.15
Moyen	89.13%	89.98%	95.91%

Tab IV.8 Comparaison des résultats de reconnaissance sur l'ensemble de données DHTB

Auteurs	Taux de reconnaissance
Alma'adeed [4]	77.00%
Moumen T. El-Melegy [173]	86.50%
Qais Al-Nuzaili [146]	92.13%
Systeme proposé	95.91%

Finalement, le [tableau IV.9](#) illustre la moyenne du temps d'exécution pour les étapes de reconnaissance. Ces expériences ont été réalisées sur un ordinateur personnel avec un processeur Intel Dual-Core i5 3320M (2.6GHz) et 8.0G RAM sous Windows 7 en utilisant MATLAB 2016. Nous notons que le temps additionnel consommé dans l'étape proposée

(correction du problème des PAWs en contact) est court par rapport au temps nécessaire pour le prétraitement et l'extraction des caractéristiques (environ 2%).

Tab IV.9 La moyenne du temps d'exécution pour l'étape de reconnaissance.

Composition de mots	Temps de prétraitement et extraction de primitives	Temps de Correction	Temps de Classification	Temps total de reconnaissance
Mots composés de 1 PAW	248.57 ms	5.94 ms	11.23 ms	265.74 ms
Mots composés de 2 PAWs	411.87 ms	6.14 ms	13.28 ms	431.29 ms
Mots composés de 3 PAWs	768.89 ms	7.04 ms	7.45 ms	783.38 ms
Mots composés de 4 PAWs	968.70 ms	/	0.61 ms	969.31ms

IV.6 Conclusion

Dans ce travail nous proposons un nouveau système de reconnaissance des montants littéraux écrits en arabe basé sur les PAWs. Notre proposition consiste en: (i) la combinaison des caractéristiques structurelles (extraites des PAWs isolées) avec les caractéristiques statistiques extraites du mot global; (ii) la division du nombre de mots en quatre sous-groupes selon le nombre de PAWs; (iii) la correction des PAWs en contact.

L'étape de classification est effectuée en utilisant 4 ensembles indépendants de SVM au lieu d'un seul. Le but de l'étape de classification proposée est de surmonter le problème de données d'entraînement déséquilibré survenant dans la stratégie, ainsi que de simplifier la tâche de classification en réduisant le nombre de classes pour chaque ensemble de SVM.

La classification de chaque nouveau mot est effectuée simultanément avec et sans correction d'erreurs des PAWs connectées. Dans les deux cas, l'échantillon candidat est présenté à l'ensemble correspondant de SVM et deux scores de classification sont alors calculés. La décision finale est tributaire du meilleur score.

Des tests expérimentaux ont été effectués sur l'ensemble des données de la base AHDB

Les résultats obtenus sont prometteurs. Concernant le temps de traitement, notre approche est une méthode globale et, par conséquent, plus rapide que les méthodes basées sur la segmentation. D'un autre côté, comparé aux autres travaux qui utilisent le SVM comme classificateur, notre approche proposée est plus rapide car elle nécessite moins de SVM. En effet, en décomposant les données en 4 sous-groupes et en résolvant chaque sous-groupe de manière indépendante, nous réduisons considérablement le temps de formation et de reconnaissance.

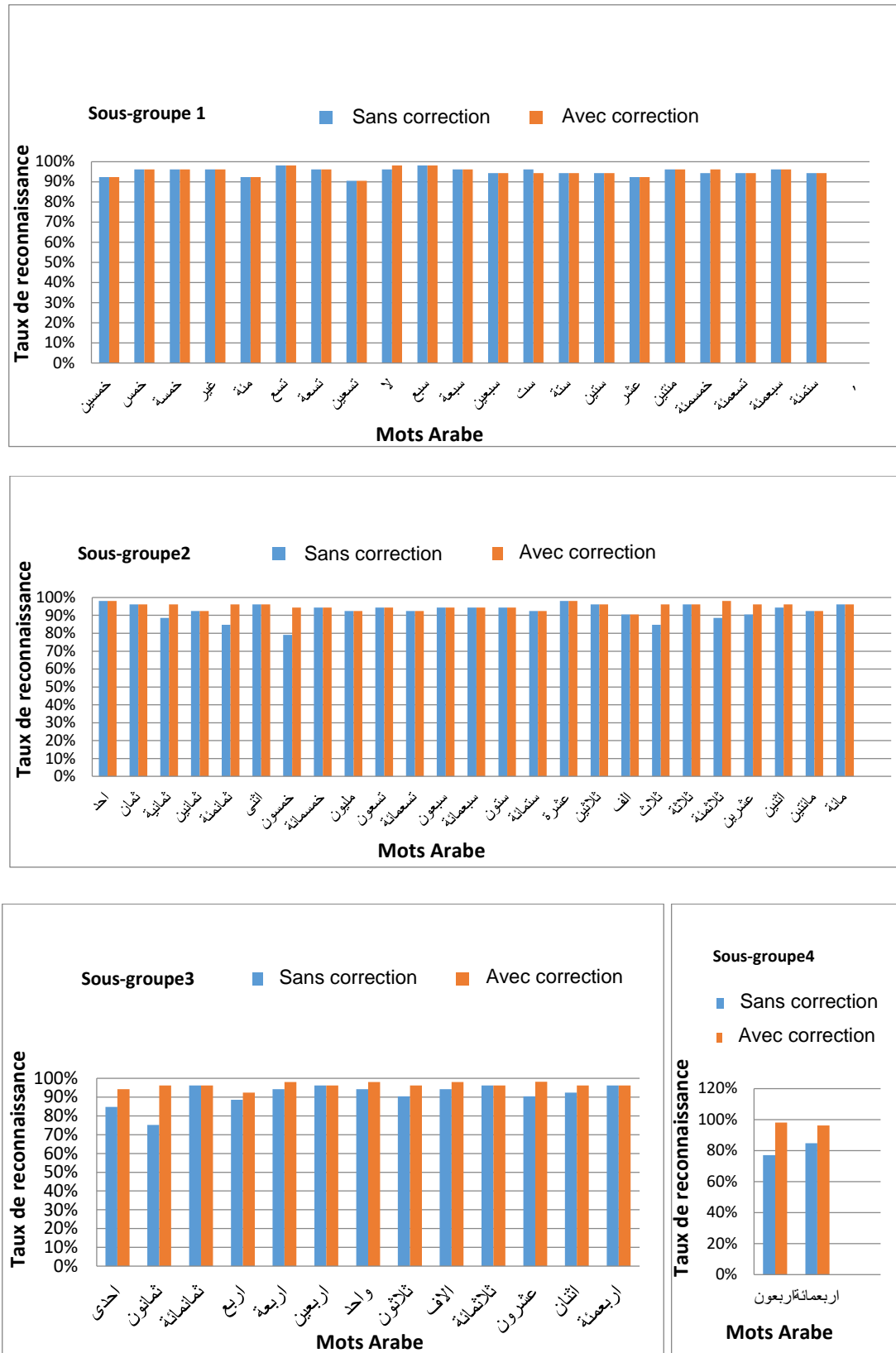


Figure IV.13 Taux de reconnaissance pour les 1^{er}, 2^{ème}, 3^{ème} et 4^{ème} sous-groupes

CONCLUSION GENERALE

Dans ce travail, nous avons proposé un nouveau système de reconnaissance pour les montants littéraux arabes utilisés dans les chèques bancaires. Le développement du système proposé est basé sur la segmentation naturelle d'un seul mot en un ensemble de pseudo-mots (PAWs) isolés. Ceci permet de faire apparaître une nouvelle approche pseudo-analytique appliquée exclusivement à l'écriture arabe. Cette approche représente une solution intermédiaire entre deux approches conventionnelles : l'approche analytique et l'approche globale. Lors du développement de notre système de reconnaissance, nous avons proposé trois nouvelles idées qui sont :

- La première idée a été proposée dans la phase d'extraction des primitives dans laquelle les caractéristiques structurelles ne sont extraites qu'à partir des PAWs au lieu que ceci ne soit fait à partir du mot candidat. La conséquence de ce choix a abouti sur des caractéristiques qui conduisent à des résultats de classification meilleurs parce que les caractéristiques structurelles extraites des PAW isolés deviennent plus appropriées que celles fournies par le mot. Néanmoins, certains mots peuvent fournir des caractéristiques structurelles similaires ce qui est susceptible de mener à une classification erronée. Pour remédier à ce problème, d'autres caractéristiques statistiques sont extraites à partir du mot candidat et qui sont ensuite combinées avec les caractéristiques structurelles fournies par ces PAWs.
- La deuxième idée a été proposée dans la phase de classification dans laquelle nous avons utilisé un multi-classificateur au lieu d'un seul. Sachant que le nombre de PAWs permet de fournir quatre sous-groupes à classifier. Pour ce faire, le problème multi-classe est subdivisé en quatre sous-problèmes et où chaque sous-problème est indépendamment résolu en utilisant quatre ensembles séparés du classificateur.
- La troisième idée a été de proposer une étape de pré-classification pour résoudre le problème des PAWs attachés entre eux. Ce dernier est causé par les écritures erronées produites par le scripteur. Dans ce cas, la classification de chaque nouveau mot est effectuée simultanément avec et sans correction d'erreurs des PAWs connectés. Dans les deux cas, l'échantillon candidat est présenté à l'ensemble correspondant de SVM et deux scores de classification sont calculés. La décision finale est donnée en fonction du meilleur score.

Le système de reconnaissance proposé est validé sur la base de données AHDB et par la comparaison des résultats obtenus par nos soins avec ceux fournis par d'autres systèmes conventionnels de reconnaissance. Ces validations ont confirmé l'efficacité de notre système proposé en termes de précision de classification ainsi que du taux de reconnaissance.

PERSPECTIVES

Le système de reconnaissance proposé a la capacité d'assurer des taux de reconnaissance très satisfaisants (sinon meilleurs) pour des montants littéraux disponibles dans la base de données AHDB et dans d'autres bases données simples. Néanmoins, pour certaines bases réelles contenant, par exemple, des montants mal écrits engendrés par le phénomène de la levée de plume, les signatures, les chiffres et autres, l'application de notre système proposé peut poser quelques problèmes. Les éventuelles solutions de ces problématiques laissent la porte entrouverte à d'autres futurs candidats intéressés par cet axe de recherche. A ce propos, nous conseillons les futurs chercheurs de proposer une approche avancée et intelligente permettant de sélectionner de façon très efficace les primitives pertinentes et cela par l'introduction d'algorithmes d'extraction devant être robustes et assurant des caractéristiques invariantes. De plus, pour la phase additionnelle de pré-classification, nous proposons aux futurs candidats d'introduire d'autres tests garantissant une bonne distinction entre les PAWs erronés et fournissant, par conséquent, les corrections appropriées. Finalement, pour la phase de classification utilisant le multi-classificateur SVM, nous conseillons les futurs candidats d'utiliser un multi-classificateur basé sur l'hybridation du SVM avec d'autres classificateurs comme ceux qui sont basés sur les réseaux de neurones, KNN, ... etc.

BIBLIOGRAPHIE

- [1] A. Belaïd et Y. Belaïd., “Reconnaissance des formes méthodes et application ” *Inter Editions*,1992.
- [2] S. Madhvanath, V. Krpasundar., and V. Govindaraju. “Syntactic methodology of pruning large lexicons in cursive script recognition”. *Pattern Recognition*, Vol.34, N°.1, pp.37–46, 2001.
- [3] 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.
- [4] 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.
- [5] H. Al-Rashaideh , “Preprocessing phase for Arabic word handwritten recognition”, *Information Transmissions in Computer Networks*, Information Process (Russian), Vol. 6, N°.1, pp. 11-19, 2006.
- [6] 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
- [7] 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.
- [8] N. Sherkat. “Influence of zoning on whole word recognition”. *Document Analysis and Recognition*. *Proceedings, Eighth International Conference on, IEEE*, Vol. 2, pp.1085-1089, 29 Aug.-1 Sept. 2005.
- [9] K.M. Sayre. “Machine recognition of handwritten words: a project report”. *Pattern Recognition*, Vol.5, pp. 213-228, 1973.
- [10] T. Pavlidis. “Algorithms for graphics and image processing”. *Hurray Hill,New Jrsay, Springer Science & Business Media*, 2012.
- [11] K.A. Berthe, D. Lamissa, and R. Stephen. "Automatic Baseline extraction based on PCA (Principal Component Analysis) method." *International Journal of Computer Science Engineering & Technology*, 2017, vol. 7, no 12.
- [12] G. Saon, “Modèles markoviens uni- et bidimensionnels pour la reconnaissance de l’écriture manuscrite hors-ligne ”. *Thèse de doctorat, Université Henri Poincaré - Nancy 1*, 1997

-
- [13] H. Fujisawa,. “Robustness design of industrial strength recognition Systems”. *In: Digital Document Processing. Major Directions and Recent Advances*, pp. 185-212. Springer, London, 2007.
- [14] A. Amin , H.B.Al-Sadoun. “A new segmentation technique of arabic text”. *IEEE. Proc ,11th IAPR, International Conference on. IEEE* ,pp. 441-445, the Hague, the Netherlands, 1992.
- [15] A. Amin, H,B, Al-Sadoun , S, Fischer . “Handprinted arabic character recognition System using an artificial network”. *Pattern recognition , Vol ,29, N°4,pp,663-675*, 1996.
- [16] H. Almuallim, S Yamaguchi. “A method of recognition of Arabic cursive handwriting”, *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol, 9, No,5, pp.715-722, Sep 1987.
- [17] I.S.I Abuhaiba., S.A Mahmoud, R.J Green., “Recognition of handwritten cursive Arabic Character”, *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol, 16, pp.664-672, June 1994.
- [18] A. Zahour, B. Taconet, A. Faure “Méthode structurelle de reconnaissance de l’écriture arabe manuscrite”. *Bigre n°68, Actes du 1^{er} Colloque national sur l’écrit et le document (CNED’90)*, pp. 148-159, Nancy, 1990.
- [19] A. Benshair., S. Mousset., P. Miche., SG. Lee. "Real-time stereo vision System for axial motion detection"; *9th Conference on Sensor Technology, Kyungpook, Korea*, november 6-7 1998; pp 50-58 (1998).
- [20] R.G. Casey and E. Lecolinet. “A survey of methods and strategies in character segmentation”. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol.18,N°7, pp.690-706, Jul 1996.
- [21] Y. Lu and M. Shridhar. “Character segmentation in handwritten words: An overview”. *Pattern recognition*, Vol.29, N°.1, pp.77-96, January 1996.
- [22] O. Hachour, “Reconnaissance hybride des caractère Arabes imprimé”, *JEP-TALN, Fès*, Avril 2004.
- [23] P.D. Gader, M.A. Mohamed, and J.H. Chiang. “Handwritten Word Recognition With Character and Inter-character Neural Networks”. *IEEE Transactions on Systems, Man, and Cybernetics*, 27(1):158-164, 1997.
- [24] S. Knerr, V. Asimov, O. Baret, N. Gorsky, and J.C. Simon. “The a2ia intercheque system : Courtesy Amount and Legal Amount Recognition for French Checks. *International Journal of Pattern Recognition and Artificial Intelligence*, 11:505- 548, 1997.
- [25] D. Ciresan. “Avoiding Segmentation in Multi-Digit Numeral String Recognition by Combining Single and Two-Digit Classifiers Trained without Negative Examples”. *In Proceedings of 10th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, pp. 225-230, Timisoara, Romania, 2008.
-

-
- [26] U. Pal, A. Belaïd, and C. Choisy. "Touching Numéral Segmentation using Water Réservoir Concept". *Pattern Recognition Letters*, 24(1-3):261-272, 2003.
- [27] A.L. Koerich, R. Sabourin, and C.Y. Suen. "Large Vocabulary Off-line Handwriting Recognition: A Survey". *Pattern Analysis and Applications*, 6(2):97-121, 2003.
- [28] R. Plamondon and S.N. Srihari. "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63-84, 2000.
- [29] A. Benouareth. "Reconnaissance de Mots Arabes Manuscrits par Modèles de Markov Cachés à Durée d'Etat Explicite", *Thèse de doctorat, Université Badji Mokhtar Annaba*, 2007
- [30] A.R. Ahmad, C. Viard-Gaudin, and M. Khalid. "Lexicon-Based Word Recognition Using Support Vector Machine and Hidden Markov Model". In *Proceedings of the 10th International Conference on Document Analysis and Recognition*, pp. 161-165, Washington DC, USA, 2009.
- [31] E. Anquetil. "Modélisation et Reconnaissance par la Logique Floue : Application a la Lecture Automatique En-ligne de l'écriture Manuscrite Omni-scripteur". *Thèse de doctorat, Université de Rennes I, France*, 1997.
- [32] E. Augustin. "Reconnaissance de mots manuscrits par systèmes hybrides Réseaux de Neurones et Modèles de Markov Cachés". *Thèse de doctorat, Université Rene Descartes Paris V*, 2001
- [33] F. Grandidier, R. Sabourin, C.Y Suen., and M. Gilloux, "Une nouvelle stratégie pour l'amélioration des jeux de primitives d'un système de reconnaissance de l'écriture". *Conférence Internationale Francophone sur l'Ecrit et le Document, CIFED*, pp. 111-120, France, 2000.
- [34] L.S Oliveira., N Benahmed., R. Sabourin, F. Bortolozzi, and C.Y. Suen, "Feature Subset Selection Using Genetic Algorithm for Handwritten Digit Recognition", *14th Brazilian Symposium on Computer Graphics and Image Processing*, (SIBGRAP'2001), pp. 362-369, Brazil, October 2001.
- [35] A.S. Britto, R.Sabourin, F. Bortolozzi and C.Y Suen., "Foreground and Background Information in an HMM-Based Method for Recognition of Isolated Characters and Numeral Strings", *9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9)*, pp.371-376, Tokyo, Japan, October, 2004.
- [36] S. Al-Ma'adeed ,C. Higgins, D. Elliman "A database for Arabic handwritten text recognition research", *Proceedings of IWFHR'02, 8th International Workshop on Frontiers in Handwriting Recognition*, pp. 485-489, Ontario, Canada, August 2002.
- [37] S. Al-Ma'adeed S., C Higgins., D Elliman., "Recognition of off-line handwritten Arabic words using hidden Markov model approach", *Proceedings of ICPR'02, 16th International*
-

-
- Conference on Pattern Recognition*, Vol. 3, pp. 481-484, Quebec City, Canada, August 2002.
- [38] M. S Khorsheed., "Automatic recognition of words in Arabic manuscripts", *PhD thesis*, University of Cambridge, Computer Laboratory, 2000.
- [39] J.Wood. "Invariant pattern recognition: A review". *Pattern Recognition*, Vol.29, N°1, pp.1-17, 1996.
- [40] L.Heutte, T. Paquet, J.V. Moreau, Y. Lecourtier and C. Olivier. "A structural/statistical feature based vector for handwritten character recognition". *Pattern Recognition Letters* 19, pp. 629 - 641, 1998
- [41] M.K. Hu. "Pattern recognition by moment invariants". In Proc. IRE, pp 1428, Sept.1961.
- [42] A. Tahmasbi, F. Saki and S. B. Shokouhi, "Classification of benign and malignant masses based on Zernike moments," *Computers in biology and medicine*. 41(8), 726-735, 2011.
- [43] M. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Amer.*, Vol. 70, no. 8, pp. 920–930, Aug. 1980.
- [44] K. Alessandrol. "Large vocabulary off-line handwritten word recognition" *Thèse de Doctorat Ecole de Technologie Supérieure, Canada, 2002*.
- [45] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.* 20(3), 273–297,1995.
- [46] F. Slimane, R. Ingold, A.M. Alimi, and J. Hennebert, "Duration models for Arabic text recognition using hidden Markov models", *Proceedings of IEEE International Conference on Computational Intelligence for Modeling, Control and Automation, Vienne-Austria, IEEE*, pp. 838-843, 2008.
- [47] A. Belaïd and Ch. Choisy. "Human reading based strategies for off-line arabic word recognition". *Summit on Arabic and Chinese Handwriting Recognition, SACH'06*, College Park, MD, United States. pp.137-149, 2006
- [48] Y. AL-Ohali, M. Cheriet et CH. Suen. "Databases for recognition of handwritten Arabic cheques". *Pattern Recognition*, Vol. 36, no 1, p. 111-121, 2003.
- [49] M. Pechwitz, S. Snoussi et Al. "IFN/ENIT Database of Handwritten Arabic Words". In *CIFED'02*, 2002.
- [50] S. N. Srihari, H. Srinivasan, P. Babu, and C. Bhole. "Handwritten Arabic word spotting using the CEDARABIC document analysis system". In *Proc. Symposium on Document Image Understanding Technology, SDIUT'05*, pp. 123–132, College Park, MD, USA, Nov, 2005.
- [51] S. A. Mahmoud, I. Abuhaiba and R. J. Green. "Skeletonization of Arabic characters using clustering based skeletonization algorithm" (CBSA). *Pattern Recognition*, Vol.24, no. 5, pp. 453-464, 1991.
- [52] S. Mozaffari, H. El Abed, V. Märgner, K. Faez and A. Amirshahi. "IfN/Farsi- Database. "A Database of Farsi Handwritten City Names". *11th International Conference on Frontiers in*
-

Handwriting Recognition, ICFHR, Canada, 2008.

- [53] S. N. Srihari, G. R. Ball and H. Srinivasan. "Versatile search of canned arabic *handwriting*. *Summit on Arabic and Chinese Handwriting Recognition*". *SACH'06*. College Park, MD, United States, pp.151-160, 2006.
- [54] S. Mozaffari, K. Faez, V. Märgner and H. El-Aded. "Strategies for Large Handwritten Farsi/Arabic Lexicon Reduction". *9th International Conference on Document Analysis and Recognition, ICDAR'07*. IEEE. Computer Society, 2007.
- [55] S.A., Mahmoud, et al. "KHATT: Arabic offline handwritten text database." *2012 International Conference on Frontiers in Handwriting Recognition (ICFHR 2012)*. IEEE, 2012.
- [56] A. Mezghani, et al. "A database for arabic handwritten text image recognition and writer identification." *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*. IEEE, 2012.
- [57] S. Al Maadeed et al. "QUWI: an Arabic and English handwriting dataset for offline writer identification." *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*. IEEE, 2012.
- [58] B. Al-Badr, S. A. Mahmoud. "Survey and bibliography of Arabic optical text recognition". *Signal Processing, vol 41*, pp. 49-77, 1995.
- [59] M.S. Khorsheed, "Off-line Arabic character recognition—a review." *Pattern analysis & applications*. Vol 5, No 1, pp.31-45, 2002.
- [60] N. E. Ben Amara, and F. Bouslama. "Classification of Arabic script using multiple sources of information: State of the art and perspectives". *IJDAR'2003*, pp. 195 – 212, 2003.
- [61] Hanene Boukerma, "Combinaison de classifieurs flous pour la reconnaissance de l'écriture arabe manuscrite". *Mémoire magister, Ecole Doctorale en Informatique de l'Est – EDI Est, 2010*
- [62] L. M. Lorigo and V. Govindaraju. "Offline Arabic Handwriting Recognition: A survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(5), pp. 712 –724, 2006.
- [63] M.T Parvez, and A.M Sabri. "Offline Arabic handwritten text recognition: a survey." *ACM Computing Surveys (CSUR)*, Vol 45 Issue 2, February 2013.
- [64] C. Olivier, H. Miled, K. Romeo, and Y. Lecourtier. "Segmentation and coding of arabic handwritten words". *In ICPR96*, pp : 264–268, 1996.
- [65] F.Farooq, G.Venu and P.Michael. "Pre-processing methods for handwritten Arabic documents." *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, on. IEEE, Vol. 1, pp .267-271, 2005.
- [66] W. Kammoun and A. Ennaji. "Reconnaissance de textes arabes à vocabulaire ouvert". *8^{ème} colloque international francophone sur l'écrit et le document, CIFED'04*, June 2004.

-
- [67] A. AL-Shatnawi, and K Omar. "Methods of Arabic language baseline detection—the state of art." *IJCSNS* 8.10 (2008): 137.
- [68] V. Märgner, M. Pechwitz, H. El Abed. ICDAR 2005 “Arabic Handwriting Recognition Competition”. In *8th International Conference on Document Analysis and Recognition, ICDAR’05*, pp.70 – 74, 2005.
- [69] F. Menasri, N. Vincent, E. Augustin, M. Cheriet. “Un système de Reconnaissance de Mots Arabes Manuscrits Hors-ligne Sans Signes Diacritiques”. Actes du dixième Colloque International Francophone sur l’Écrit et le Document, CIFED’08, 2008.
- [70] M. Ziaratban, K. Faez. “A Novel Two-Stage Algorithm for Baseline Estimation and Correction in Farsi and Arabic Handwritten Text Line”. *IEEE*, pp.1-5,2008.
- [71] M. Pechwitz, V. Maergner, H El-Abed. “Comparison of Two Different Feature Sets for Offline Recognition of Handwritten Arabic Words”.*10th International Workshop on Frontiers in Handwriting Recognition*, (France), 2006.
- [72] D. Motawa, A. Amine, and R. Sabourin. “Segmentation of Arabic cursive script”. In *Proceedings of the International Conference on Document Analysis and Recognition*. IEEE, pp.625-628, 1997.
- [73] S. Wshah, Z. Shi, and V. Govindaraju. “Segmentation of Arabic handwriting based on both contour and Skeleton segmentation”, In *proceedings of the 10 international conference on document analysis and Recognition (ICDAR)*.pp.793- 797, 2009.
- [74] A Hamid and R. Haraty "A Neuro-heuristic Approach for Segmenting Handwritten Arabic Text" In: *ACS-IEEE International Conference on Computer Systems and Applications*, pp. 0110, 2001.
- [75] A.M. Elgammal, M.A. Ismail.“A graph-based segmentation and feature extraction framework for Arabic text recognition”.*6th International Conference on Document Analysis and Recognition*, pp.622-626, 2001.
- [76] B. Ai-Badr, R.M. Haralick. “Segmentation-free word recognition with application to Arabic”. In *Proceedings of the Third international Conference on Document Analysis and Recognition*, Vol. 1, pp. 355-359,1995.
- [77] K. R. K Romeo-Parker, H Miled,., and Y. Courtier. “A new approach for Latin/Arabic character segmentation”. In *proceedings of the 3rd International Conference on Document Analysis and Recognition (ICDAR)*. pp.874-877,1995.
- [78] M. Syiam, T. M. Nazmy, A. E. Fahmy, H. Fathi and K. Ali. “Histogram clustering and hybrid classifier for handwritten Arabic characters recognition”. *Proc of the 24th IASTED international conference on Signal processing, pattern recognition, and applications*.
- [79] Yusra Osman. “Segmentation algorithm for Arabic handwritten text based on contour analysis” *International Conference on Computing, Electrica J and Electronics Engineering (ICCEEE), IEEE*. pp.447-452, 2013.
-

-
- [80] H. Miled, C. Olivier, M. Cheriet and k. Romeo-pakker. "Une méthode rapide de reconnaissance de l'écriture arabe manuscrite". *In seizième Colloque sue le Traitement du signal et des Images GRETSI, France*. pp. 857-860,1997.
- [81] H. Miled,, M. Cheriet, and C Olivier. "Multi-level Arabic handwritten words recognition." *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*. Springer, Berlin, Heidelberg, 1998.
- [82] M. S. Khorsheed, and W. F. Clocksin. "Structural Features of Cursive Arabic Script". *Proceedings of the British Machine Conference*, pp. 42.1-42.10. BMVA Press, September 1999,
- [83] H. Al Abed and V. Märgner, "Comparison of different preprocessing and feature extraction methods for off-line recognition of handwritten Arabic words", *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR)*, Curitiba-Brazil, vol. 2, pp. 974-978, 2007
- [84] P. Burrow : "Arabic handwriting recognition ". *Master of science thesis. School of Informatics, university of Edinburg, England*, 2004
- [85] A. Mubaak, A. Hussain, and Q.F. Wan. "Off-line handwritten Arabic word recognition using SVMs with normalized poly kernel. In : *International Conference on Neural Information Processing. Springer, Berlin, Heidelberg*,p. 85-91. 2012
- [86] Leila Chergui. "Combinaison de classifieur pour la reconnaissance de mots arabes manuscrits".Thèse de Doctorat, Université Mentouri Constantine, 2012.
- [87] N. Azizi, T. Sari, L. Souici-Meslati, M. Sellami."Une architecture de combinaison floue de classifieurs neuronaux pour la reconnaissance de mots arabes manuscrits". *CEFED'02, 7 Colloque International Francophone sur l'Ecrit et le Document*, pp. 89-96, Hammamet, Tunisie, Octobre 2002.
- [88] N. Farah, L. Souici, and M. Sellami."Arabic word recognition by classifiers and context". *Journal of Computer Science and Technology*, Vol. 20, no. 3, pp. 402-410, 2005.
- [89] L. Gasmi, H. Merouani, et L. Souici-Meslati."Combinaison de classifieurs". *Proceeding of the 3rd Intemational Conference: Sciences of Electronic, Technologies of Information and Telecommunication (SETIT)*, Sousse-Tunisia, 2005.
- [90] A. Benouareth, A. Enniji, and M. Sellami. "Semi-continuous HMMs with explicit State duration for unconstrained Arabic Word modeling and recognition" *Pattern Recognition Letters*, Vol. 29, pp. 1742-1752. 2008.
- [91] R. Al-Hajj, L. Likforman-Sulem, and C. Mokbel. "Combining slanted-frame classifiers for improved HMM-based Arabic handwriting recognition". *IEEE Trans. Pattern Anal. Mach. Intel.*31 (7), pp. 1165-1177, 2009.
- [92] Y. kessentini, T. Paquet, and A. Ben Hamadou."Off-line handwritten word recognition using
-

-
- multi-stream hidden Markov models”. *Pattern Recognition Letters* 31(1): pp.60-70, 2010.
- [93] S. Abdel Azeem, and H. Ahmed. “Effective technique for the recognition of offline Arabic handwritten words using hidden Markov models” *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 16, no.4, pp. 399-412, 2013.
- [94] J.H. Alkhateeb, J. Jiang, J. Ren, F.Khelifi, and S.S. Ipson. “Multiclass. Classification of Unconstrained Handwritten Arabic Words Using Machine Learning Approaches”. *The Open Signal Processing Journal*, Vol. 2, pp. 21-28, 2009.
- [95] A. Broumandnia, J. Shanbehzedeh, and M. Varnosfaderani. “Persian/Arabic handwritten word recognition using M-band packet wavelets transform”. *Image and Vision Computing, Elsevier*, Vol. 26, pp. 829-842, 2008.
- [96] J. Chen, H. Cao, R. Prasad, A. Bhardwaj and P. Natarajan, “Gabor features for offline Arabic”. *DAS '10 Proceedings of the 9th IAPR International Workshop on Document Analysis Systems .handwriting recognition*. pp. 53-58, 2010.
- [97] H. Nemmour, Y. Chibani. “Handwritten Arabic word recognition based on Ridgelet transform and support vector machines”. *In Proceedings of International Conference on high Performance Computing and Simulation (HPCS)*, pp. 357-361,2011.
- [98] M. Elzobi, A. Al-Hamadi, Z. Al Aghbari, L., Dings, and A.Saeed. “Gabor Wavelet Recognition Approach for Off-Line Handwritten Arabic Using Explicit Segmentation”. *Advances in Intelligent Systems and Computing*. p233, 2014.
- [99] B. El qacimy, M. Ait kerroum, A. Hammouch. “Word-Based Arabic Handwritten Recognition Using SVM Classifier with a Reject Option”. *In proceeding of 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015.
- [100] A Lawgali,. “A survey on arabic character recognition”, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2015, vol. 8, no 2, p. 401-426.
- [101] J. Milgram. “Contribution à l’intégration des Machines à Vecteurs de Support au sein des systèmes de reconnaissance de formes: Application à la lecture automatique de l’écriture manuscrite“. Thèse de Doctorat, Univ Québec, 2007.
- [102] R. El-Hajj, L. Likforman-Sulem, Ch. Mokbel. “Arabic Handwriting Recognition Using Baseline Dependant Features and Hidden Markov Modeling”. *In 8th International Conference on Document Analysis and Recognition, ICDAR’05*, pp.893–897.
- [103] T. Sari, L. Souici, and M. Sellami. “Off-line handwritten Arabic character segmentation algorithm: ACSA”. *In IWFHR’02: Proceedings of the Eighth International Workshop on Frontier in Handwriting Recognition, IEEE*. pp 452, 2002.
- [104] X. Ding, And H. liu. “Segmentation driven offline handwritten Chinese and Arabic script recognition”. *In Arabic and Chinese Handwriting Recognition. Lecture Notes in Computer Science*, Vol.4768. Springer,pp.196-217,2008.
- [105] M. Elzobi, A. Al-Hamadi, Z. Al Aghbari, L., Dings. “IESK-ARDB: a database for
-

-
- handwritten Arabic and an optimized topological segmentation approach". *International Journal on Document Analysis and Recognition (IJ DAR)*, pp.1-14, 2012.
- [106] H. A. AI Hamad, and R. Abu Zitar. "Development of an efficient neural-based segmentation technique for Arabic handwriting recognition". *Pattern. Recognition*. pp 2773-2798, 2010.
- [107] R.M. Abdeen. A. Affi, AB. El-Sisi. "An Enhanced Technique for Offline Arabic Handwritten Words Segmentation". *16th International Conference, CICLing 2015, Cairo, Egypt, April 14-20, 2015*.
- [108] B. M. F. Bushofa, and S. Michael. "Segmentation and recognition of Arabic characters by structural classification." *Image and Vision Computing* 15.3 (1997): 167-179.
- [109] O. Matan, C.J.C.Burges. "Recognizing overlapping hand-printed characters by centered-objects integrated segmentation and recognition". In *IJCNN*, pp. 504-511, 1991.
- [110] S.Marinai, M.Gori, and G.Soda, "Artificial Neural Networks for Document Analysis and Recognition". *IEEE Trans. Pattern Anal. Mach. Intell.* 27(1): pp.23-35 .2005.
- [111] N. Arica, F. Yarman-Vural, "An Overview of Character Recognition Focused on Off-line Handwriting", *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 31(2), pp. 216 - 233. 2001.
- [112] T. Saba, G. Sulong, A. Rehman, "A Survey on Methods and Strategies on Touched Characters Segmentation", *International Journal of Research and Reviews in Computer Science (IJRRCS)*, Vol. 1, No. 2, June 2010.
- [113] R. J. Rodrigues, A. C. G. Thomé, "Cursive Character Recognition - A Character Segmentation Method using Projection Profil-based Technique", *Proceedings of The 4th World Multi-conference on Systemics, Cybernetics and Informatics SCI 2000 and The 6th International Conference on Information Systems, Analysis and Synthesis ISAS 2000 - Orlando, USA, 2000*.
- [114] C. Fang, "Deciphering Algorithms for Degraded Document Recognition", *Ph.D thesis. State University of New York*, 1997.
- [115] D. Montana, L. Davis, "Training feed forward neural networks using genetic algorithms", in: *Proceedings of the 11th international joint conference on Artificial Intelligence*, Detroit, Vol. 1, 762-767, 20-26 August 1989.
- [116] M. Cesar, R. Shinghal, "Algorithm for segmenting handwritten postal codes", *Int. J. Man Mach Stud.*, Vol. 33, no. 1, pp. 63-80, Jul. 1990.
- [117] R.A. Wilkinson, "Comparison of massively parallel segmenters", *National Inst. of Standards and Technology Technical Report, Gaithersburg, MD*, Sept. 1992.
- [118] L.D. Harmon, "Automatic Recognition of Print and Script", *Proceeding, IEEE*, Vol. 60, No. 10, pp. 1165-1177, 1972.
- [119] A. Zramdini, R. Ingold, "Optical Font Recognition from Projection Profiles". *Electronic Publishing* 6(3):pp. 249-260 .1993.
-

-
- [120] A. Kawtrakul, W. Peerawit, and W. Yingsaeree, "The utilization of closing Algorithm and Heuristic Information for Broken Character Segmentation", *Conference on Cybernetics and Intelligent Systems & Robotics, Automation and Mechatronics*, IEEE, December 2004.
- [121] S. J. Lu, L. Li, Chew Lim Tan, "Identification of Latin-Based Languages through Character Stroke Categorization". *ICDAR 2007*, pp.352-356. 2007.
- [122] U. Pal, A. Sarkar: "Recognition of Printed Urdu Script". *ICDAR 2003*: pp.1183-1187. 2003.
- [123] S. Ouchtati, M. Bedda, and A. Lachouri, "Segmentation and Recognition of Handwritten Numeric Chains", *International journal of information technology*. Vol 4. No 1, ISSN 1305-2403. 2007.
- [124] R.G. Casey, G. Nagy, "Recursive Segmentation and Classification of Composite Patterns", *Proc. 6th Int. Conf. on Pattern Recognition*, p. 1023, 1982.
- [125] M. Gilloux, J.M. Bertille and M. Leroux, "Recognition of Handwritten Words in a Limited Dynamic Vocabulary", *Pre-Proceedings IWFHR III, Buffalo*, p. 417, 1993.
- [126] M. Gilloux, "Hidden Markov Models in Handwriting Recognition", *Fundamentals in Handwriting Recognition, S. Impedovo (Ed.), NATO ASI Series F: Computer and Systems Sciences*, Vol. 124, Springer Verlag, 1994.
- [127] S.Nebti, "Reconnaissance de Caractères Manuscrits par Intelligence Collective", *Thèse de doctorat*, Université Ferhat Abass -Sétif, mars 2013.
- [128] R. Zaghdoudi, H. Seridi, "Combination of Multiple Classifiers for Off-Line Handwritten Arabic Word Recognition". *International Arab Journal of Information Technology (IAJIT)*, 2017, vol. 14, no 5.
- [129] A.D.S.B. Jr, R.Sabourin, F. Bortolozzi, and C.Y. Suen, "A Two-Stage HMM-Based System for Recognizing Handwritten Numeral Strings". *ICDAR 2001*:pp. 396-400. 2001.
- [130] K.C. Hayes. "Reading handwritten words using hierarchical relaxation". *Computer Graphics and Image Processing*, pp. 344-364,1980.
- [131] H. Fujisawa, Y. Nakano and K. Kurino, "Segmentation Methods for character recognition from segmentation to document structure analysis", *Proc.IEEE*, Vol.80, no.7, pp.1079-1092,1992.
- [132] S. Madhvanath, V.Govindaraju, "Holistic Lexicon Reduction", *Proc.IWFHR III, Buffalo*, pp.71,1993.
- [133] M. Shridhar, A. Badreldin, "Recognition of Isolated and Simply Connected Handwritten Numerals", *Pattern Recognition*, Vol. 19, no. 1, p. 1, 1986.
- [134] X. Dupre. "Contributions à la reconnaissance de l'écriture cursive à l'aide de modèles de Markov cachés". *Phd thesis, Univ Rene Descartes - Paris V*, 2003.
- [135] B. Bushofa, M Spann, "Segmentation of Arabic characters using their contour information". *13th International Conference on Digital Signal Processing*, vol. 2, pp. 683-686
-

-
- [136] A. M. Gouda, et M. A. Rashwan,. Segmentation of connected Arabic characters using hidden Markov models. In: *2004 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2004. CIMSA*. IEEE, 2004. p. 115-119.
- [137] Y. LeCun, et al “ Handwritten Zip Code Recognition with Multilayer Networks”, in *IAPR (Eds), Proc. of the International Conference on Pattern Recognition, IEEE, pp.35-40*, 1990.
- [138] K. Koo et al “Character Segmentation and Recognition Algorithm of Text Region in Steel Images”, *ISPRA'09 Proceedings of the 8th WSEAS international conference on Signal processing, robotics and automation*,. 2009.
- [139] S.W. Lee, D.J. Lee, and H.S. Park, “A New Methodology for Gray-Scale Character Segmentation and Recognition”, *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 18, No. 10. 1996.
- [140] N. Tripathy, U. Pal., "Handwriting Segmentation of unconstrained Oriya Text" *International Workshop on Frontiers in Handwriting Recognition*, pp. 306-311,2004.
- [141] Y. M. Alginahi,. “A survey on Arabic character segmentation”. *International Journal on Document Analysis and Recognition (IJDAR)*, 2013, vol. 16, no 2, p. 105-126.
- [142] S. Madhvanath, K. Gyeonghwan, et V. Govindaraju,. “Chaincode contour processing for handwritten word recognition”. *IEEE transactions on pattern analysis and machine intelligence*, 1999, vol. 21, no 9, p. 928-932.
- [143] D. Gaceb, V. Eglin, and F. Lebourgeois, “A new mixed binarization method used in a real-time application of automatic business document and postal mail sorting,” *International Arab Journal of Information Technology*. 10(2), pp 179-188 ,2013.
- [144] I. Ahmad and S. A. Mahmoud, “Arabic bank check processing: State of the Art,” *Journal of Computer Science and Technology*. 28(2), p 285,2013.
- [145] Z. Tamen, H. Drias, and D. Boughaci, “An efficient multiple classifier system for Arabic handwritten words recognition,” *Pattern Recognition Letters*. 93, pp.123-132,2017.
- [146] Q. Al-Nuzaili et al., “Enhanced structural perceptual feature extraction model for Arabic literal amount recognition,” *International Journal of Intelligent Systems Technologies and Applications*. 15(3), pp.240-254,2016.
- [147] N. Azizi et al., “A new hybrid method combining genetic algorithm and support vector machine classifier: Application to CAD system for mammogram images,” *International Conference In Multimedia Computing and Systems (ICMCS)*, IEEE.pp.415-420,2014.
- [148] M. Elzobi et al., “IESK-ArDB: a database for handwritten Arabic and an optimized topological segmentation approach”. *International Journal on Document Analysis and Recognition (IJDAR)*, 16(3), pp.295-308,2013.
- [149] R. Palacios and A. Gupta, “A system for processing handwritten bank checks automatically,” *Image and Vision Computing*. 26(10), pp.1297-1313,2008.
-

-
- [150] M. Režnáková et al, “Forgetting of unused classes in missing data environment using automatically generated data: Application to on-line handwritten gesture command recognition,” *Pattern Recognition*. Pp. 355-367,2017.
- [151] S. Alma'adeed et al., “Recognition of off-line handwritten Arabic words using hidden Markov model approach”. In *Pattern Recognition, Proceedings. 16th International Conference on IEEE*. 3, pp.481-484, 2002. [doi: [10.1109/ICPR.2002.1047981](https://doi.org/10.1109/ICPR.2002.1047981)].
- [152] G. A. Abandah and F. T Jamour, “Recognizing handwritten Arabic script through efficient skeleton-based grapheme segmentation algorithm,” *Intelligent Systems Design and Applications (ISDA 2010)*, pp. 977-982, 2010.
- [153] S. Snoussi-Maddouri, F. Ghazouani, and F. B. Samoud. “Text lines and paws segmentation of handwritten arabic document by two hybrid methods,” In *Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 310–315, Sousse, Tunisia, 2014.
- [154] S. Snoussi-Maddouri et al., “Combination of local and global vision modelling for Arabic handwritten words recognition, ” In *Eighth International Workshop on Frontiers in Handwriting Recognition(IWFHR'02)* ,IEEE, pp. 128-135,2002.
- [155] L. Souici-Meslati and M. Sellami, “A hybrid approach for Arabic literal amounts recognition,” *Arabian Journal for Science & Engineering, Springer Science & Business Media BV*.pp 29, 2004.
- [156] N. Farah, T. Khadir, and M. Sellami, “Artificial neural network fusion: Application to Arabic words recognition,” In *ESANN*,pp. 151-156,2005.
- [157] A. Boukharouba and A. Bennis, “Recognition of handwritten Arabic literal amounts using a hybrid approach,” *Cognitive Computation*. 3(2),pp.382-393.2011.
- [158] B. M. Al-Helali and S. A. Mahmoud, “Arabic Online Handwriting Recognition (AOHR): A Survey,” *ACM Computing Surveys (CSUR)*. 50(3), p33, 2017.
- [159] Y. Chherawala, P. P. Roy, and M. Cheriet, “Feature set evaluation for offline handwriting recognition systems: application to the recurrent neural network model,” *IEEE transactions on cybernetics*. 46(12), pp.2825-2836, 2016. [doi:10.1109/TCYB.2015.2490165].
- [160] H. Ghaleb, P. Nagabhushan, and U. Pal, “Graph Modeling based Segmentation of Handwritten Arabic Text into Constituent Sub-words,” *International Journal of Image, Graphics and Signal Processing*. 8(12), p8, 2016.
- [161] A. Rehman and T. Saba, “Neural networks for document image preprocessing: state of the art,” *Artificial Intelligence Review*. 42(2), pp. 253-273, 2014.
- [162] H. Al-Rashaideh, “Preprocessing phase for Arabic word handwritten recognition,” *Information Transmissions in Computer Networks*. 6(1),pp.11-19, 2006.
- [163] A. Cardoso and A. Wichert, “Handwritten digit recognition using biologically inspired features,” *Neurocomputing*. 99, pp.575-580,2013.

-
- [164] A. Saidani, A. Kacem, and A. Belaid, "Arabic/Latin and machine-printed/handwritten word discrimination using HOG-based shape descriptor," *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 14(2), pp.1-23,2015.
- [165] F. Zaiz, M. C. Babahenini, and A Djeflal, "Puzzle based system for improving Arabic handwriting recognition," *Engineering Applications of Artificial Intelligence*. 56, pp.222-229, 2016.
- [166] A. Kacem, N. Aouiti, and A. Belaïd, "Structural features extraction for handwritten Arabic personal names recognition. In *Frontiers in Handwriting Recognition*," *International Conference on IEEE (ICFHR2012)*, pp. 268-273, 2012.
- [167] F. Hang and C. J. Chen, "A component-labeling algorithm using contour tracing technique," *International Conference on Document Analysis and Recognition, IEEE. (ICDAR 2003)*, *IEEE* ,pp. 741-745,2003.
- [168] H. Samet and M. Tamminen, "Efficient Component Labeling of Images of Arbitrary Dimension Represented by Linear Bintreees". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE. 10 (4):p 579, 1988.
- [169] S. Snoussi-Maddouri, "Baseline extraction: Comparison of six methods on IFN/ENIT database," *In the 11th International Conference on Frontiers in Handwriting Recognition*. pp 345-249, 2008.
- [170] A.Ameur, K.Romeo-Pakker,H.Miled, M.Chriet, "Approche Globale pour la Reconnaissance des Mots Manuscrits arabes", *Actes CNED94, 3ème Colloque National sur l'écrit et le document*, pp.151-156, Juillet 1994.
- [171] M. Dehghan et al., "Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM," *Pattern Recognition*, 34(5),pp. 1057-1065,2001.
- [172] D. Impedovo and G. Pirlo, "Zoning methods for handwritten character recognition: A survey," *Pattern Recognition*. 47 (3),pp. 969-981,2014.
- [173] M. T. El-Melegy et al., "Global features for offline recognition of handwritten Arabic literal amounts," *Information and Communications Technology, ICICT 2007. ITI 5th International Conference on IEEE*. pp.125-129 , 2007. [[doi: 10.1109/ITICT.2007.4475631](https://doi.org/10.1109/ITICT.2007.4475631)]
- [174] A. Menasria, A.Bennia, M. Nemissi, & M.Sedraoui "Multiclassifiers system for handwritten Arabic literal amounts recognition based on enhanced feature extraction model". *Journal of Electronic Imaging*, 27(3), 033024.2018.