

**REPUBLIQUES ALGERIENNES DEMOCRATIQUES ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE**

**ACADEMIE UNIVERSITAIRE DE CONSTANTINE
UNIVERSITE BADJI-MOKHTAR ANNABA**

MEMOIRE

Présenté par

BENOUARETH ABDALLAH

Pour obtenir le diplôme de

MAGISTER EN INFORMATIQUE

Option : Systèmes Multi-Agents

RECONNAISSANCE DE L'ECRITURE ARABE MANUSCRITE PAR UNE APPROCHE HYBRIDE

Soutenu -1999- devant le jury :

Président :	Z. Sahnoun	Prof.	Université de Constantine.
Rapporteur :	M. Sellami	M.C.	Université Badji-Mokhtar -Annaba.
Examineur :	M.C. Batouche	M.C.	Université de Constantine.
Examineur :	T. Bensbaa	C.C.	Université Badji-Mokhtar -Annaba.
Examineur :	M. Ahmed-Nacer	M.C.	Université de Constantine.
Invité	: M. Bedda	M.C.	Université Badji-Mokhtar -Annaba

SOMMAIRE

INTRODUCTION.....1

PARTIE I : ETAT DE L'ART

CHAPITRE 1 : RECONNAISSANCE AUTOMATIQUE DE L'ECRITURE ARABE.....5

1.1. INTRODUCTION	5
1.2. ASPECTS DE LA RECONNAISSANCE DE L'ECRITURE	6
1.3. CARACTERISTIQUES DE L'ECRITURE ARABE	9
1.4. UN MODELE GENERAL POUR LA RECONNAISSANCE OPTIQUE DE TEXTES ARABES (AOTR)	12
1.4.1. ACQUISITION	13
A) LES SYSTEMES DE RECONNAISSANCE ON-LINE.....	13
B) LES SYSTEMES OFF-LINE	13
1.4.2. PRETRAITEMENT	14
A)BINARISATION (SEUILLAGE)	15
B)LISSAGE ET FILTRAGE	15
C)SQUELETTISATION (THINNING).....	16
D)NORMALISATION	19
E)DETECTION DE LA LIGNE DE BASE.....	19
1.4.3. SEGMENTATION.....	19
1.4.3.1. DECOMPOSITION DE PAGE	20
1.4.3.2. SEGMENTATION DU MOT.....	21
A) SUPPOSER LES CARACTERES DEJA SEGMENTES	21
B) SEGMENTATION EN PRIMITIVES	22
C) SEGMENTATION EN CARACTERES	23
D) SEGMENTATION BASEE SUR LA RECONNAISSANCE.....	25
E) RECONNAISSANCE SANS SEGMENTATION (METHODES HOLISTIQUES).....	26
1.4.4. EXTRACTION DE CARACTERISTIQUES	26
A)CARACTERISTIQUES STRUCTURELLES.....	27
B)CARACTERISTIQUES STATISTIQUES.....	28
C)LES TRANSFORMATIONS GLOBALES.....	29
D)TEMPLATE MATCHING ET CORRELATION.....	30
1.4.5. CLASSIFICATION.....	31
A)METHODES STRUCTURELLES	31
B)METHODES STATISTIQUES	32
C)METHODES NEURONALES (CONNEXIONNISTES).....	33
1.4.6. LES APPROCHES HYBRIDES.....	34
1.4.7. APPRENTISSAGE	35
1.4.8. POST-TRAITEMENT	36
1.5. LES TENDANCES FUTURES.....	36
1.6. CONCLUSION.....	39

PARTIE II : CONTRIBUTION A LA RECONNAISSANCE DE L'ECRITURE ARABE MANUSCRITE

CHAPITRE 1 : LES MODELES DE MARKOV CACHES (HMMS) : ASPECT THEORIQUE.....41

1.1. INTRODUCTION	41
1.2. MODELES DE MARKOV MONO-DIMENSIONNELS (1 D).....	41
1.2.1. DEFINITIONS PRELIMINAIRES.....	41
1.2.2. MODELE STOCHASTIQUE.....	42
1.2.3. PROPRIETES DE MODELES DE MARKOV MCC/HMM.....	42
EXEMPLE	43
1.2.4. MODELES DE MARKOV CACHES DISCRETS	43
1.2.4.1. DEFINITION.....	44
1.2.4.2. EXEMPLE.....	44
1.2.4.3. ELEMENTS D'UN HMM	45
1.2.2.4.4. PROCEDURE DE GENERATION DES OBSERVATIONS.....	46
1.2.4.5. LES TROIS PROBLEMES FONDAMENTAUX ASSOCIES AUX HMMS...	46
1.2.4.6. EXEMPLE.....	47
1.2.4.7. SOLUTION DU PROBLEME 1.....	48
1.2.4.8. SOLUTION DU PROBLEME 2.....	51
1.2.4. 9. SOLUTION DU PROBLEME 3.....	52
1.2.4.10. TYPES D'HMMS	54
1.3. MODELES DE MARKOV PSEUDO-2D (PLANAIRE).....	55
1.3.1. DEFINITIONS DES PHMMS.....	55
1.3.2. CALCUL DE DISTORTION	57
1.4. CONCLUSION.....	58

CHAPITRE 2 : UTILISATION DES HMMS DANS LA RECONNAISSANCE AUTOMATIQUE DE L'ECRITURE.....59

2.1. INTROUDUCTION	59
2.2. APPLICATIONS DES HMMS 1D A LA RECONNAISSANCE DE L'ECRITUR .60	
2.2.1. DECOMPOSITION PROBABILISTE	60
2.2.2. DECOMPOSITION DE LA FORME PAR RAPPORT A L'ETIQUETTE.....	61
EXEMPLE 1.....	61
EXEMPLE 2.....	63
2.2.3. DECOMPOSITION PAR ASSOCIATION A UN MODELE	64
2.3. APPLICATION DES PHMMS A LA RECONNAISSANCE DE L'ECRITURE.....	70
2.4. CONCLUSION.....	73

CHAPITRE 3 : UNE APPROCHE HYBRIDE POUR LA RECONNAISSANCE DES MOTS ARABES MANUSCRITS.....75

3.1. INTRODUCTION	75
3.2. METHODOLOGIE.....	76
3.3. PRETRAITEMENT	78
3.3.1. DEFINITIONS.....	78
3.3.2. BINARISATION	80